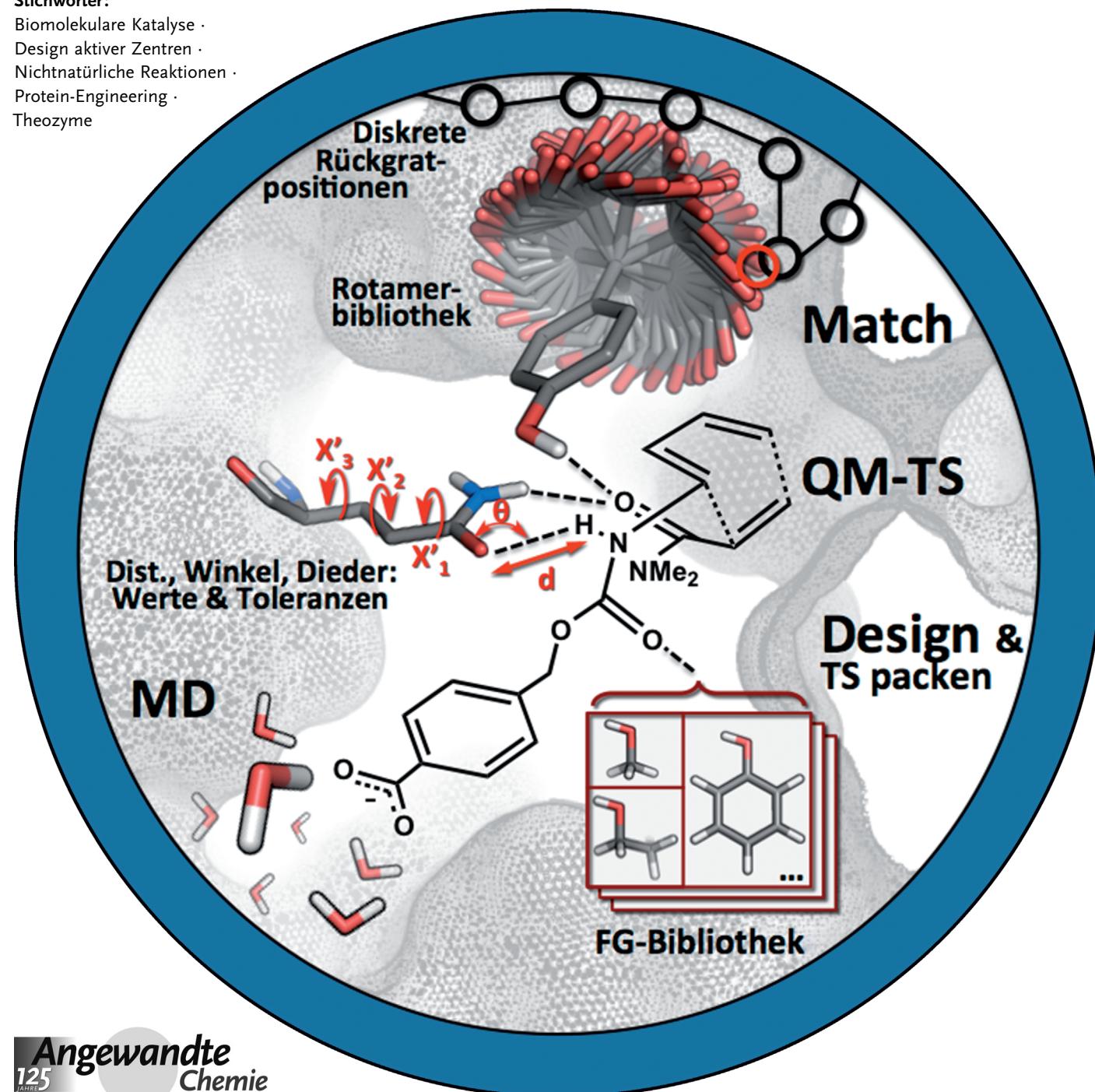


Computerbasiertes Enzymdesign

Gert Kiss, Nihan Çelebi-Ölçüm, Rocco Moretti, David Baker und K. N. Houk*

Stichwörter:

Biomolekulare Katalyse ·
Design aktiver Zentren ·
Nichtnatürliche Reaktionen ·
Protein-Engineering ·
Theozyme



Der „Inside-Out“-Ansatz des computerbasierten Enzymdesigns vereint die neuesten Entwicklungen im Bereich der Computerchemie und -biologie. In diesem Rahmen ist es möglich geworden, Proteine zur Katalyse von Reaktionen herzustellen, für die es in der Natur kein Pendant gibt. Die Erfolgsrate ist gegenwärtig gering, sodass nur ein Bruchteil dieser Konzeptproteine zum Schluss wie geplant funktioniert. Errungenschaften, aber auch Beschränkungen der gegenwärtigen Technologie werden hier behandelt und mit anderen Methoden verglichen. Auf sich alleine gestellt, ermöglicht der „Inside-Out“-Ansatz die Produktion von Proteinen mit katalytischer Aktivität und Selektivität – wenn auch mit bescheidenen kinetischen Eigenschaften im Vergleich zu natürlich vorkommenden Enzymen. Gerichtete Evolution, Proteindynamiksimulationen und Crowd-Sourcing können jedoch zu erheblich verbesserten Varianten solcher „Inside-Out“-Proteine führen.

1. Einleitung

Biologische Organismen sind auf Proteinkatalysatoren angewiesen, welche für die vielfältigen Reaktionen biochemischer Metabolismen zuständig sind. Über drei Milliarden Jahre der Evolution haben Enzyme hervorgebracht, die sogar die langsamsten, unter Normalbedingungen nicht messbaren chemischen Reaktionen soweit beschleunigen können, dass sie nun in biologischen Zeiträumen ablaufen und dadurch in biologische Prozesse integriert werden können. Die kompetentesten dieser Biokatalysatoren können Reaktanten am Diffusionslimit umsetzen. Darüber hinaus vereinen Enzyme ihre katalytische Wirkung mit einer gewissen Spezifität in Bezug auf Substrat und Reaktion. Sie sind oft auf Cofaktoren oder Coenzyme angewiesen und erzielen ihr Funktionsoptimum bei ganz bestimmten pH-Werten, Temperaturen und Lösungsmittelverhältnissen, die jedoch zusammengenommen ein erstaunliches Spektrum abdecken. Kurzum, Enzyme sind die idealen Komponenten für komplexe und fein aufeinander abgestimmte metabolische Vorgänge.

Der Mensch hat sich enzymatische Prozesse bereits früh zunutze gemacht, und so wurden natürliche Fermentationsprozesse gezielt zur Herstellung von Lebensmitteln und alkoholischen Getränken eingesetzt. Das Wort „Enzym“ – Altgriechisch „en zýmē“ für „in Hefe“ – wurde jedoch erst 1876 zum ersten Mal systematisch zur Beschreibung dieser Biokatalysatoren verwendet, hauptsächlich um „Missverständnissen vorzubeugen und lästige Umschreibungen zu vermeiden“.^[1] Die Identität und Funktionsweise von Enzymen war zu der Zeit noch spekulativ, wobei knapp zwanzig Jahre später Emil Fischer ein „Schloss und Schlüssel“-Modell als Erklärung für die beobachtete Substratspezifität von Enzymen vorschlug. Die Aufarbeitung und Kristallisation von Urease durch James B. Sumner in 1926 zeigte weiter, dass Enzyme eigenständige Proteine sind. 1946 spekulierte Linus Pauling, dass ein gegebenes Enzym nahezu komplementär zu der Übergangsstruktur einer chemischen Reaktion sei^[2] – eine bemerkenswerte Aussage in Anbetracht der Tatsache, dass es zu der Zeit noch keine Struktur gab und nur wenig

Aus dem Inhalt

1. Einleitung	5811
2. Protein-Engineering	5813
3. Computerbasiertes Inside-Out-Design von Enzymen	5819
4. Computerbasiertes Design von Enzymen – Erfolgsbeispiele	5822
5. Herausforderungen beim Enzymdesign	5830
6. Zusammenfassung und Ausblick	5832

funktionelle Information zur Verfügung stand.^[3] Die Erforschung von Struktur und Funktion mit atomistischer Auflösung wurde erst zwanzig Jahre später möglich, als die erste Kristallstruktur eines Enzyms gelöst wurde und das Gebiet der Strukturbioologie entstand.^[4]

Seither hat eine Flut an enzymkatalytischer Forschung unser Wissen über diese wundervollen „Nanomaschinen“ auf die Probe gestellt, justiert und erweitert. Paulings Hypothese kann als ein allgemeines Basismodell verstanden werden. Um jedoch der erstaunlichen katalytischen Wirkung von Enzymen näherzukommen, ist es nötig, weitere Konzepte wie etwa elektrostatische und strukturelle Präorganisation oder chemische Katalyse zu berücksichtigen.^[5] Die Faktoren erstrecken sich von nichtkovalenten Stabilisierungen des Übergangszustands (Elektrostatik, Desolvatisierung, Einschränkung der Beweglichkeit usw.) bis hin zu kovalenten Effekten (schwache Wasserstoffbrücken, Entstehung von Zwischenprodukten, Komplexverbindungen mit Metallionen usw.). Die kompetentesten^[6] dieser Katalysatoren bieten also weit-aus mehr als ein aktives Zentrum, das dem Übergangszustand komplementär ist; sie beteiligen sich an der Reaktion und verändern sowohl den Übergangszustand als auch das Energieprofil der Reaktion im Vergleich zum unkatalysierten Pendant in Lösung.^[7] Diese „kovalente Hypothese“ liefert eine Erklärung für den Befund, dass der Großteil von Enzy-

[*] Dr. G. Kiss, Dr. N. Çelebi-Ölçüm, Prof. Dr. Dr. K. N. Houk
 Department of Chemistry and Biochemistry
 University of California, Los Angeles
 607 Charles E. Young Dr. East, Los Angeles CA, 90095 (USA)
 E-Mail: houk@chem.ucla.edu

Dr. R. Moretti, Prof. Dr. D. Baker
 Department of Biochemistry and Howard Hughes Medical Institute
 University of Washington, Seattle, WA 98195 (USA)

Dr. G. Kiss
 Gegenwärtige Adresse: Department of Chemistry
 Stanford University, Stanford, CA 94305 (USA)

Dr. N. Çelebi-Ölçüm
 Gegenwärtige Adresse: Yeditepe Universität
 Department of Chemical Engineering, Istanbul (Türkei)

men den Übergangszustand um Größenordnungen stärker „bindet“ als von nichtkovalenten Wechselwirkungen zu erwarten wäre.

Katalytische Kompetenz ist formell die Bindungskonstante des Komplexes zwischen Enzym und Übergangszustand und wurde von Richard Wolfenden als $K_{tx}^{-1} = (k_{cat}/K_M)/k_{uncat}$ definiert.^[6] K_{tx}^{-1} erstreckt sich für bisher erforschte Enzyme über bemerkenswerte 21 Größenordnungen^[6,8–12] und beträgt im Durchschnitt $10^{16.0 \pm 4.0} \text{ M}^{-1}$.^[13] Dies entspricht einem mittleren ΔG von 22 kcal mol⁻¹ und reicht bis zu 38 kcal mol⁻¹; dies ist um einiges stärker als die 15 kcal mol⁻¹, die von unkovalenten Wechselwirkungen erreicht werden können.

Chemiker versuchen oft, sich das eine oder andere Merkmal von Enzymen – wie etwa Selektivität, katalytische Kompetenz, schonende, ungiftige und umweltfreundliche Reaktionsbedingungen – zu Nutzen zu machen. Häufig wird eine Kombination von Eigenschaften angestrebt, sodass der Katalysator im Idealfall sowohl einen hohen Durchsatz als auch Substratspezifität, Stereo-, Regio- oder Chemoselektivität und/oder eine erhöhte Toleranz gegenüber Temperatur, pH und Lösungsmittel aufweist. Unterfangen dieser Art

wurden an anderer Stelle bereits ausführlich behandelt. Beispiele erstrecken sich von der Bioinformatik,^[14,15] über vergleichende Mutationsexperimente innerhalb von Enzymsuperfamilien^[16] zu Wirt-Gast- und supramolekulare Chemie,^[13] gerichteter Evolution,^[17–21] katalytischen Antikörpern,^[22–24] Organokatalyse,^[25] rationalem strukturbasiertem Protein-Engineering^[26] und computerbasiertem Protein-Design.^[27]

In diesem Aufsatz beschreiben wir den computerbasierten „Inside-Out“-Ansatz zur rationalen Herstellung von Enzymen und die Anfänge einer Technologie, die wir versuchen, in eine robuste Anwendung für synthetische, biotechnologische und therapeutische Zwecke weiterzuentwickeln. Die Idee, die diesem Ansatz zugrundeliegt, beruht auf der Verwendung von biochemischen Bausteinen (Aminosäuren, Cofaktoren, Coenzymen usw.) zur Konstruktion von Katalysatoren für nicht-biologische Prozesse, die mit mikrobiologischen Verfahren hergestellt werden können. Erhöhte Rechnerleistungen haben in jüngster Zeit zu erheblichen Fortschritten bei der Entwicklung verbesserter Algorithmen zur Strukturvorhersage und Konformationssuche beigetragen. Durch die Anwendung quantenmechanischer (QM) Methoden ist es



Gert Kiss studierte Chemie in Heidelberg, forschte als Teil eines Austauschprogramms im Labor von Noah W. Allen an der UNCA und promovierte bei K. N. Houk an der UCLA. Er war NIH-CBI-Stipendiat, LLNL Lawrence Scholar und erhielt den Stauffer Research Award. Derzeit ist er Postdoktorand (NIH Simbios) im Labor von Vijay S. Pande in Stanford.



Nihan Celebi-Ölçüm erhielt ihren B.S. in Chemie von der Universität Boğaziçi, Türkei, und ihren M.S. in Computer- und Theoretischer Chemie an der Université Henri Poincaré in Nancy, Frankreich. Sie promovierte bei Viktorya Aviyente an der Universität Boğaziçi und absolvierte ein Postdoktorat in der Arbeitsgruppe von Kendall Houk an der UCLA. Seit kurzem ist sie Professorin an der Universität Yeditepe in der Türkei.



Rocco Moretti erwarb seinen B.S. in Biochemie am Worcester Polytechnic Institute und promoviert in Biochemie an der University of Wisconsin unter Betreuung von Aseem Ansari und Jon Thorson. Er ist gegenwärtig Postdoktorand im Labor von David Baker an der University of Washington.



David Baker promovierte bei Randy Schekman in Berkeley und forschte als Postdoktorand in Biophysik bei David Agard an der UCSF. Er ist gegenwärtig Professor der Biochemie an der University of Washington und Forscher am Howard Hughes Medical Institute. Er ist Mitglied der National Academy of Sciences und der American Academy of Sciences.



K. N. Houk promovierte bei R. B. Woodward in Harvard und war dann Professor an der Louisiana State University und der University of Pittsburgh, wonach er in 1986 der UCLA beitrat. Er war Direktor der Chemistry Division der National Science Foundation von 1988 bis 1990. Gegenwärtig ist er Saul Winstein Professor in Organischer Chemie. Er ist Mitglied der National Academy of Sciences, der American Academy of Arts and Sciences, und der International Academy of Quantum Molecular Sciences.

möglich, Konfigurationen von Funktionsgruppen zu ermitteln, die den Übergangszustand einer gegebenen Reaktion maximal stabilisieren. Die Wahl dieser Komponenten ist üblicherweise von natürlicher Präzedenz inspiriert. Proteinsequenzen, die solch katalytisch kompetente Konfigurationen ermöglichen, sind theoretisch dazu in der Lage, ein gegebenes nichtnatürliches Substrat in sein(e) Produkt(e) umzuwandeln. Anstelle der Synthese komplett neuer Peptidsequenzen wird ein neues aktives Zentrum in eine bereits existierende, stabile Proteinstruktur eingearbeitet.

Wir besprechen hier das „Inside-Out“-Protokoll, heben Annäherungen und Engpässe hervor, präsentieren Erfolgsbeispiele, untersuchen Errungenschaften und Herausforderungen und berichten von Fällen, in denen Variationen und Zusätze zum herkömmlichen Protokoll von Nutzen waren. Wir kommen zu dem Schluss, dass Moleküldynamik(MD)-Simulationen, gerichtete Evolution und Crowd-Sourcing zu verbesserten katalytischen Eigenschaften führen können. Wir beginnen mit Beispielen aus dem Protein-Engineering, die den Weg geebnet haben.

2. Protein-Engineering

2.1. Katalytische Antikörper

Seit den bahnbrechenden Forschungsarbeiten von Lerner und Schulz Mitte der 80er Jahre^[28,29] wurden katalytische Antikörper für eine Vielfalt chemischer Reaktionen hergestellt.^[24,30] Das Konzept basiert auf Paulings Hypothese, dass Enzyme ein Umfeld schaffen, welches der Struktur und dem elektrostatischen Charakter des Übergangszustands komplementär ist.^[2,31] Wenn Antikörper mit einem Hapten-Antigen präsentiert werden, das dem Übergangszustand einer gegebenen Reaktion ähnlich ist, so entstehen Antikörper, die nicht nur das Hapten, sondern auch den Übergangszustand binden können. Wenn der Übergangszustand stärker als das Substrat gebunden wird, kommt dies einer Verringerung der Aktivierungsenergie gleich, was sich als erhöhte Reaktionsgeschwindigkeit äußert.^[32] Die Produktion von katalytischen Antikörpern basiert auf Mutations- und Selektionsvorgängen, die charakteristisch für das adaptive Immunsystem sind. Die resultierenden Wechselwirkungen sind spezifisch und können für die Katalyse nichtnatürlicher Reaktionen, aber auch für die Umsetzung nichtnatürlicher Substrate genutzt werden. Quantenmechanische Berechnungen eignen sich für das Erzeugen von Übergangszustandsanaloga (TSAs, „transition state analogues“), die als Haptene für eine gegebene Reaktion dienen können.^[33] Für das tiefergehende Studium von Strategien und Herausforderungen bei der Entwicklung von Haptenen wird an dieser Stelle auf Literatur von Janda und Mitarbeitern verwiesen.^[24]

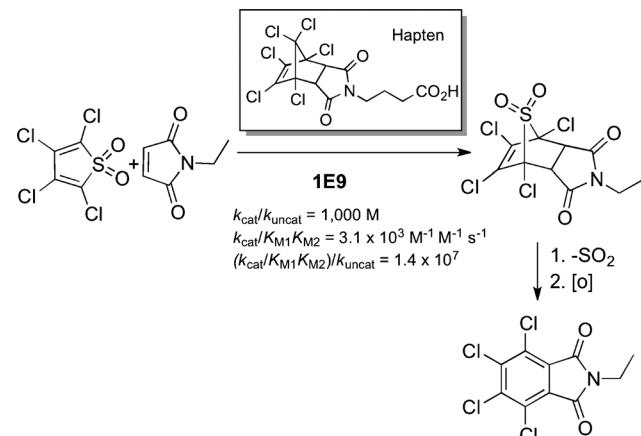
Zu den Reaktionen, die bisher von Antikörpern katalysiert werden können, zählen die Diels-Alder-Cycloaddition, Acyltransfer-Reaktionen, Oxy-Cope-Umlagerungen und Cyclisierungen. Katalytische Kompetenzen reichen von $K_{tx}^{-1} = 10^{4.6} \text{ M}^{-1}$ bis zu $K_{tx}^{-1} = 10^{8.6} \text{ M}^{-1}$, sodass die Übergangszustände der Reaktionen stärker gebunden werden als deren Substrate ($K_M^{-1} = 10^{3.5 \pm 1.0} \text{ M}^{-1}$).^[13,30] Natürlich vorkom-

mende Enzyme weisen jedoch enorme K_{tx}^{-1} -Werte auf und haben eine durchschnittliche Reichweite von 10^{12} bis 10^{20} M^{-1} . Untypisch effiziente Enzyme, wie etwa ODCase und Alkylsulfatasen, erreichen K_{tx}^{-1} -Werte von 10^{24} M^{-1} bzw. 10^{29} M^{-1} .^[8,34] Die katalytische Effizienz (k_{cat}/K_M) solcher Enzyme ist dann oft diffusionslimitiert und beträgt zwischen 10^4 und $10^9 \text{ M}^{-1} \text{ s}^{-1}$. Katalytische Antikörper sind im Vergleich dazu um etwa vier Größenordnungen und mehr langsamer ($k_{cat}/K_M = 10^2\text{--}10^5 \text{ M}^{-1} \text{ s}^{-1}$).^[13,24,34] Dieser Befund wurde einer Vielzahl von Faktoren zugeschrieben, wie etwa Produktinhibition,^[35-37] geringere Bindungskonstanten,^[13] nichtkovalente Katalyse,^[7] kleine, verborgene Oberflächen,^[13] Unterschiede in Evolutionszeitskalen^[22] und Limitationen der Immunglobulinstruktur.^[22] Die vergleichsweise geringe Stabilität von Immunglobulinen und die einhergehenden hohen Produktionskosten katalytischer Antikörper limitieren darüber hinaus ihr industrielles Anwendungspotential. Von einem therapeutischen Standpunkt jedoch, ist das Interesse an katalytischen Antikörpern am steigen: Beispiele erstrecken sich von der Neutralisierung von HIV-1^[38,39] zu Antikörper-gesteuerter Enzym-Prodrug-Therapie (ADEPT) und der Inaktivierung von Suchtmitteln.^[40]

Während die katalytische Effizienz von Antikörpern oft durch Produktinhibition beschränkt ist und nur die einfachsten chemischen Reaktionen durch Haptene imitiert werden können, so hat das Feld kritische Beiträge zur Erforschung biokatalytischer Prozesse geleistet. Im folgenden Abschnitt stellen wir Beispiele vor und diskutieren die Funktion von Simulationen und Berechnungen im Bezug auf das Design und Verständnis von Antikörerkatalysatoren.

2.1.1. Die Diels-Alder-Reaktion

Der erste katalytische Antikörper (1E9) für die Diels-Alder(DA)-Reaktion wurde 1989 von Auditor und Mitarbeitern hergestellt.^[41] Er katalysiert die Cycloaddition von Tetrachlorthiophen und N-Ethylmaleimid mit einem k_{cat}/k_{uncat} von 1000 M (Schema 1).^[42] Die Kristallstruktur zeigt ein meist hydrophobes Bindungszentrum, in welchem ein Asparagin eingebettet ist (Asn^{H85}). Chen et al. untersuchten 1E9 mit

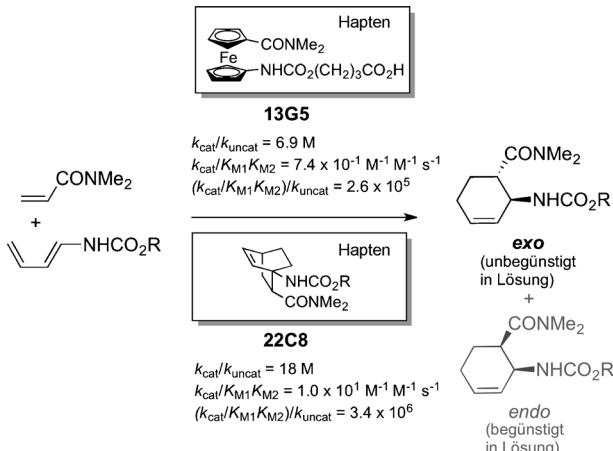


Schema 1. Die Antikörper-katalysierte Cycloaddition von Tetrachlorthiophen und N-Ethylmaleimid.

einer Kombination von QM-Rechnungen, Docking, MD-Simulationen und einer Wechselwirkungsextrapolationsmethode (linear interaction energy approach).^[43] Das aktive Zentrum in 1E9 ist komplementär zu der Geometrie und dem elektrostatischen Profil des Übergangszustands, sodass dieser stärker gebunden wird als die Reaktanten.

Die Diels-Alder-Reaktion bietet jedoch Protein-Ingenieuren die Möglichkeit, sich – über das Beschleunigen der Reaktionsgeschwindigkeit hinaus – der stereochemischen Kontrolle anzunehmen. Gouverneur et al. demonstrierten dies an der Cycloaddition von *trans*-1-*N*-Acylamino-1,3-butadien mit *N,N*-Dimethylacrylamid, die unter thermodynamischer Kontrolle zu einem Gemisch von *endo*- und *exo*-Stereoisomeren führt. Die sterischen und elektrostatischen Merkmale des Übergangszustands wurden anhand von QM-Rechnungen ermittelt, was dann zum Design von Haptenen für den *endo*- als auch den *exo*-Pfad führte.^[44]

Der DA-Antikörper 13G5 katalysiert die unbegünstigte *exo*-Diels-Alder-Reaktion zwischen Methyl-*N*-butadienylcarbamat und *N,N*-Dimethylacrylamid ($k_{\text{cat}} = 1.20 \times 10^{-3} \text{ min}^{-1}$, $k_{\text{cat}}/k_{\text{uncat}} = 6.9 \text{ M}$) und ergibt nur eines der Stereoisomere mit hohem Enantiomerenüberschuss (Schema 2).^[45] Die Kristallstruktur und QM-Rechnungen



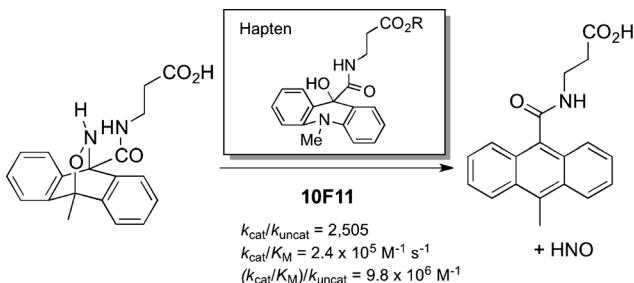
Schema 2. Die Antikörper-katalysierte, unbegünstigte *exo*-Diels-Alder-Reaktion von Methyl-*N*-butadienylcarbamat und *N,N*-Dimethylacrylamid.

zeigen, dass Asp^{H50} und Tyr^{L36} für den Großteil der katalytischen Effizienz von 13G5 verantwortlich sind, während Asn^{L91} die Reaktanten stärker bindet als den Übergangszustand und somit die Reaktion etwas verlangsamt.^[45,46] Der Befund legt nahe, dass Asn^{L91} für die korrekte relative Orientierung der Substrate zuständig ist.

Die absolute sterische Konfiguration des Produkts wurde später experimentell als *exo*-(3*S*,4*S*) bestimmt. Die hohe stereochemische Spezifität von 13G5 wurde zunächst mit Docking-Rechnungen der Übergangszustände in die Kristallstruktur untersucht.^[45] MD-Simulationen zeigten später, dass die katalytische Aminosäure Asp^{H50} in Gegenwart des *exo*-(3*S*,4*S*)-Übergangszustands zusätzlich von einem Wasserstoffmolekül koordiniert ist. In Gegenwart des *exo*-(3*R*,4*R*)-Übergangszustands beträgt diese Koordinationszahl 3.^[46]

Dies führt zu einer geringeren Wechselwirkung zwischen dem katalytischen Asp^{H50} und dem Carbamat-NH des *exo*-(3*R*,4*R*)-Übergangszustands verglichen mit dem *exo*-(3*S*,4*S*)-Fall.

Der DA-Antikörper 10F11 katalysiert eine retro-Diels-Alder-Reaktion, die HNO mit einem $k_{\text{cat}}/k_{\text{uncat}}$ von 2500 freisetzt (Schema 3).^[47] Die Untersuchung der Kristallstruk-



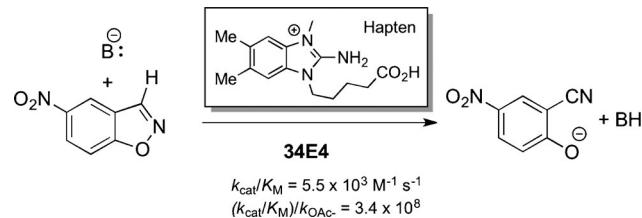
Schema 3. Antikörper-katalysierte retro-Diels-Alder-Reaktion.

tur ergab, dass das aktive Zentrum dem Übergangszustand komplementär ist, wobei Trp, Phe und Ser als die katalytischen Aminosäuren agieren.^[48] DFT-QM-Rechnungen wurden für eine genauere Untersuchung der katalytischen Wechselwirkungen eingesetzt.^[49]

Der Übersichtsartikel von Kim et al. untersucht die unkovalente Katalyse von Diels-Alder-Reaktionen und stellt einen Vergleich unterschiedlicher Katalysatoren auf, wie etwa Cyclodextrine, selbstorganisierende Kapseln, Antikörper und RNAsen. Der Rückchluss der Studie ist, dass – anders als natürliche Enzyme – keines dieser Systeme den jeweiligen Übergangszustand substantiell und spezifisch bindet.^[50]

2.1.2. Die Kemp-Eliminierung

Der erste Antikörper (34E4), der die ringöffnende Reaktion von 5-Nitrobenzisoxazolen (Kemp-Eliminierung) katalysiert, wurde 1996 von Hilvert und Mitarbeitern hergestellt (Schema 4).^[51] 34E4 weist einen k_{cat}/K_M von $5.5 \times 10^3 \text{ M}^{-1} \text{s}^{-1}$ auf.



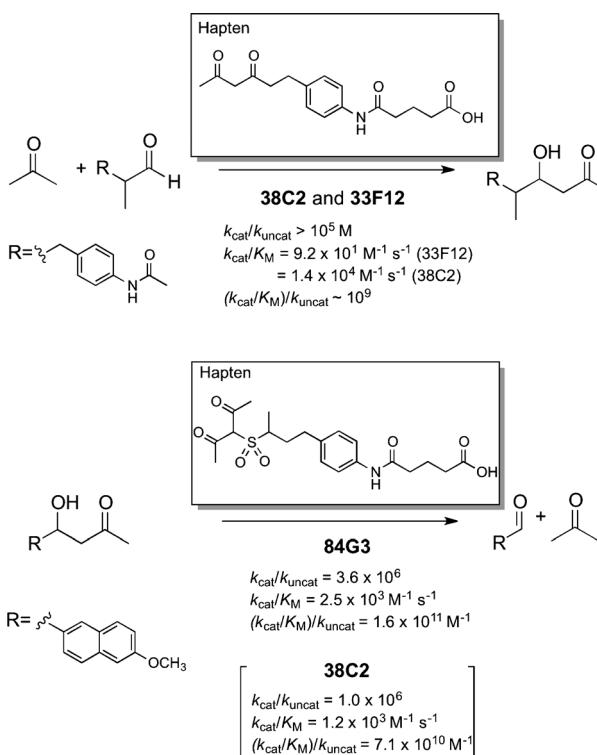
Schema 4. Antikörper-katalysierte ringöffnende Reaktion von 5-Nitrobenzisoxazol.

auf und hat einen $k_{\text{cat}}/k_{\text{uncat}}$ von 2.1×10^4 im Vergleich zur unkatalysierten Hintergrundreaktion ($k_{\text{uncat}} = 3.1 \times 10^{-5} \text{ s}^{-1}$). Das aktive Zentrum ist hydrophob und wird von einer katalytischen Base (Glu^{H50}) dominiert. Verglichen mit der Wirkung von Acetat in wässriger Lösung (was einer theoretischen Translokation der katalytischen Base aus dem aktiven Zentrum heraus und in das Lösungsmedium hinein entspricht und

eine Geschwindigkeitskonstante k_{OAc^-} von $1.6 \times 10^{-5} M^{-1} s^{-1}$ aufweist), hat der Antikörper eine Beschleunigung von $(k_{cat}/K_M)/k_{OAc^-} = 3.4 \times 10^8$. Experimentelle Befunde und theoretische Studien zeigen, dass die Kemp-Reaktion – ähnlich wie andere Protonentransfer-Reaktionen auch – stark von der geometrischen Orientierung zwischen Substrat und Base abhängt.^[52,53] Darüber hinaus, und vor allem wenn eine Carboxygruppe als katalytische Base fungiert, ist die Kemp-Eliminierung stark vom polaren Charakter des Lösungsmittels abhängig.^[54,55] Diese Eigenschaften sowie auch die Leichtigkeit, mit der Produktbildung gemessen werden kann (UV/Vis), haben die Kemp-Eliminierung zu einer beliebten Modellreaktion für die Untersuchung biochemischer Basenkatalyse gemacht.

2.1.3. Aldol/Retro-Aldol-Reaktion

QM-Rechnungen von Aldol-Reaktionen reichen bis in die 1980er Jahre zurück,^[56] und so standen den Forschern die entsprechenden Übergangszustände schon relativ früh zur Verfügung. Die katalytischen Antikörper ab38C2, ab84G3 und ab33F12 wurden später mithilfe der jeweiligen Haptene hergestellt und erzielten ähnliche Reaktionsgeschwindigkeiten wie natürliche Aldolasen – mit jedoch wesentlich geringerer Substratspezifität (Schema 5).^[57–61] Diese Antikörper ähneln Klasse-1-Aldolasen, welche die ε-Aminogruppe eines Lysins zur Ausbildung einer Schiffschen Base mit dem Substrat nutzen. Das größtenteils hydrophobe aktive Zentrum ist zwischen 5 und 7 Å vom katalytischen Lysin-ε-N mit einer Ansammlung polarer Aminosäuren dekoriert. Arno und Domingo untersuchten deren potentielle katalytische Funk-

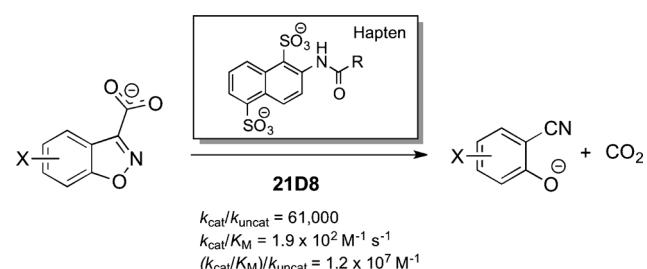


Schema 5. Beispiele von Antikörper-katalysierten Aldol- und retro-Aldol-Reaktionen.

tion vor allem mit Bezug auf den C-C-bindungserzeugenden Schritt der Reaktion.^[62]

2.1.4. Decarboxylierungsinduzierte Ringöffnungsreaktionen

Die Geschwindigkeit, mit der 5-Nitro-3-carboxybenzisoxazol decarboxyliert, hängt vom Lösungsmittel ab und variiert um bis zu acht Größenordnungen.^[54,55] Aprotische polare Lösungsmittel beschleunigen die Reaktion durch Desolvatisierung der Carboxygruppe und somit durch Stabilisierung des Übergangszustands durch Dispersionswechselwirkungen.^[63–66] Antikörper 21D8 katalysiert die Decarboxylierung von 5-Nitro-3-carboxybenzisoxazol^[67] um ein 61000-faches im Vergleich zur Hintergrundreaktion in Wasser (Schema 6).^[68]

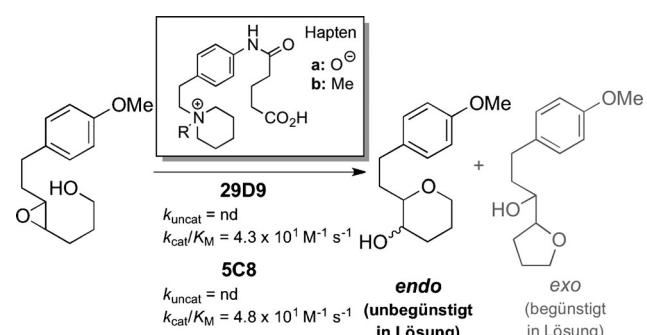


Schema 6. Antikörper-katalysierte Decarboxylierung von 5-Nitro-3-carboxybenzisoxazol.

Die Ursache der katalytischen Wirkung von 21D8 wurde mit QM-Rechnungen, MD-Simulationen und Gibbs-Energie-Perturbationsrechnungen (FEP) weiter untersucht.^[69] Dabei stellte sich heraus, dass Lösungsmittelkontakte mit der Carboxygruppe die katalytische Leistungsfähigkeit beeinträchtigen können, während stabilisierende Wasserstoffbrücken mit dem Isoxazol-Sauerstoff dem Lösungsmittelleffekt entgegenwirken.

2.1.5. Cyclisierung von trans-Epoxyalkoholen

Insgesamt vier Antikörper wurden anhand zweier Haptene hergestellt, um die unbegünstigte *endo*-tet-Cyclisierung von *trans*-Epoxidalkoholen (Schema 7) zu katalysieren.^[70,71] Quantenmechanische Rechnungen zeigten, dass der *endo*-tet-



Schema 7. Antikörper-katalysierte, unbegünstigte *endo*-tet-Cyclisierung von *trans*-Epoxidalkoholen.

Übergangszustand S_N1-Charakter hat und elektrostatisch durch eine Carboxygruppe destabilisiert werden kann.^[72–74] Ähnliche Rechnungen wurden für die 6-exo- sowie die 7-endo-Cyclisierung durchgeführt und ergaben, dass letztere bevorzugt ist. Das Ergebnis wurde daraufhin experimentell verifiziert^[75] und deckt sich ebenfalls mit der später für Antikörper Fab 5C8 gelösten Kristallstruktur.^[71] Die katalytische Komponente des aktiven Zentrums besteht aus der Asp^{H95}-His^{L89}-Dyade, welche für Säure/Base-Katalyse geeignet positioniert ist.

2.1.6. Hydrolyse von aromatischen Amiden und Estern

Der Antikörper 43C9 katalysiert die Hydrolyse von aromatischen Amiden und Estern mit einer ungewöhnlich hohen Effizienz ($k_{\text{cat}}/k_{\text{uncat}}$ of 2.5×10^5).^[30,76–78] Getzoff und Mitarbeiter erzeugten eine homologe Struktur zur variablen Region des Antikörpers. Die Forscher ermittelten, dass Arg^{L96} höchstwahrscheinlich zur Stabilisierung des entstehenden Oxyanions dient, während His^{L91} als das katalytische Nukleophil fungiert.^[79] Die später gelöste Kristallstruktur von 43C9 untermauerte die Vorhersagen und zeigte weiterhin auf, dass ein Wasser-vermitteltes Wasserstoffbrückennetzwerk im aktiven Zentrum für die beobachtete katalytische Effizienz unentbehrlich ist.^[80] Kollman und Mitarbeiter führten später QM-Rechnungen, MD-Simulationen und FEP-Rechnungen an Antikörper 43C9 aus und schlugen einen direkten hydrolytischen Angriff als Alternative zum vorangehenden nucleophilen Postulat vor.^[81]

2.1.7. Chorismat-Prephenat-Umlagerung

Die Claisen-Umlagerung von Chorismat zu Prephenat wird sowohl von natürlich vorkommenden Chorismat-Mutassen^[82,83] als auch von den Antikörpern IF7 und IIF1-2E11 katalysiert.^[84,85] Wiest et al. verwendeten QM-Methoden, um das aktive Zentrum von Chorismat-Mutase zu untersuchen und ermittelten, dass im Falle des Antikörpers IF7 spezifische Wasserstoffbrücken für die ca. 200-fache Beschleunigung verantwortlich sind.^[86]

2.2. Gerichtete Evolution

Evolution im Labor (gerichtete Evolution) ist zu einer der ausgereifteren Formen des Proteingenieurwesens geworden und ist mittlerweile Bestandteil von modernen industriellen Prozessen.^[87] Die Methode ist weit verbreitet und beruht auf mehreren Runden der Mutation und Selektion.^[17,18] Anwendungsbeispiele erstrecken sich von verbesserter Thermostabilität^[88] zu erhöhter Lösungsmitteltoleranz,^[89] verstärkten Protein-Protein-Wechselwirkungen,^[90] modifizierter Substrattoleranz und -spezifität,^[91,92] erhöhter enzymatischer Aktivität^[93] und Inversion der Enantioselektivität.^[94,95] Durch Mutagenisieren eines Startgens wird eine Ansammlung von Sequenzen (Variantenbibliothek) zusammengestellt. Diese wird daraufhin nach Proteinen filtriert, welche die gewünschte Eigenschaft verstärkt aufweisen. Die Auswirkungen einer einzelnen Evolutionsrunde sind im Allgemeinen

relativ gering, sodass dieser Prozess einige Male wiederholt werden muss, bis die gewünschte Steigerung erzielt wird. Variantenbibliotheken können auf unterschiedlichem Wege zusammengestellt werden. Drei klassische Beispiele sind stochastiche PCR-Mutagenese (Abbildung 1a), gezielte

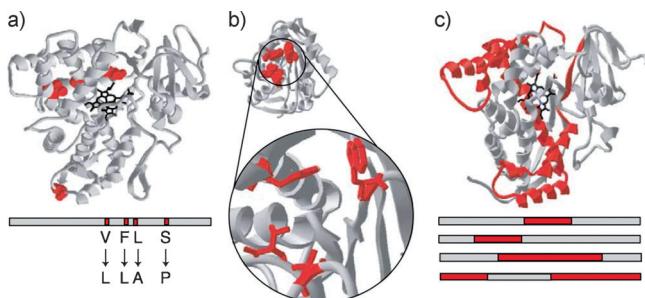


Abbildung 1. Drei Strategien für das Zusammenstellen von Variantenbibliotheken in der gerichteten Evolution. a) Stochastische Mutagenese der gesamten Proteingehälfte. b) Gezielte, ortsspezifische Mutagenese. c) Sequenzrekombination gesamter Proteingeschäfte. Nachdruck von Bloom et al.^[19]

ortsspezifische Mutagenese (CASTing, ISM)^[95] (Abbildung 1b) und Sequenzrekombination (Abbildung 1c).^[19] Die Stärken der stochastischen Mutagenese liegen darin, dass keine strukturellen oder mechanistischen Vorkenntnisse nötig sind und dass unerwartete Mutationen weit vom aktiven Zentrum aufgedeckt werden können.^[96]

Die ortsspezifische Mutagenese (Abbildung 1b) zielt hingegen auf ganz bestimmte Bereiche des Proteins ab (wie etwa das aktive Zentrum), sodass strukturelle und/oder biochemische Vorkenntnisse über das Protein nötig sind. Durch den kleineren Sequenzbereich ist es jedoch möglich, dass mehrere positive Mutationen innerhalb des Zielbereichs aufgedeckt werden können.^[97] Dies ist vor allem dann von Vorteil, wenn eine Kombination von Mutationen für die Verbesserung der gewünschten Funktion nötig ist.

Positive Mutationen innerhalb einer Variantenbibliothek können beispielsweise durch eine statistische Auswertung des Zusammenhangs zwischen Aktivität und Sequenz (ProSAR)^[98] identifiziert werden. Diese können dann durch Gen-Shuffling kombiniert und eingearbeitet werden.^[99] Die strukturellen und funktionellen Eigenschaften können durch Neutral-Drift-Variationen weiter diversifiziert werden. Dabei kommt es zur Anreicherung von Mutationen, welche die gewünschte Eigenschaft des Proteins nicht beeinträchtigen.^[100,101]

Eine der Hauptherausforderungen in der gerichteten Evolution ist die Ermittlung derjenigen Varianten – aus einer großen Zusammenstellung mutierter Sequenzen –, welche die gewünschten Merkmale aufweisen.^[20] Selektionsbasierte In-vitro-Verfahren, wie etwa mRNA-Display^[102] und emulsionsbasierte Mikrofluidik-FADS (fluoreszenzaktivierter Tröpfchensorierer),^[103] weisen einen beachtlichen Durchsatz auf. Screening-Verfahren, die auf der Bestimmung der Substrat- oder Produktkonzentration beruhen, sind vergleichbar vielseitig anwendbar, haben jedoch einen relativ begrenzten Durchsatz.^[104] Wenn eine Verbindung zwischen Genotyp und

Phänotyp hergestellt werden kann, sind gerichtete Evolutionsmethoden generell auf alle biologisch herstellbaren Proteine anwendbar, auch auf solche, die nichtnatürliche Aminosäuren oder anderweitige prothetische Modifikationen enthalten.^[105]

Es ist zwar möglich, von einem völlig naiven Startpunkt aus positive Mutationen zu finden, gerichtete Evolutionsversuche sind jedoch am effektivsten, wenn das in Frage kommende Enzym bereits ein geringes Maß an Aktivität gegenüber der Zielreaktion – oder zumindest einer eng verwandten Reaktion – aufweist.^[109] Während der Erfolg von gerichteten Evolutionsexperimenten auf einen eindeutigen, bergauf gerichteten Pfad vom Startpunkt zu einer hochaktiven Variante angewiesen ist,^[105] so haben Proteinsequenzen vielmals nicht die gewünschte anfängliche Aktivität. Solche Herausforderungen können zum Teil mithilfe von Neutral-Drift-Variantenbibliotheken bewältigt werden, oder durch ein graduelles Ändern des selektiven Zwangs von der existierenden Funktion zu der gewünschten hin.^[20,21]

In den letzten Jahrzehnten wurden viele Versuche unternommen, um strukturelle und/oder funktionelle Merkmale von Proteinen und Enzymen zu modifizieren. Die erfolgreichen beruhten meist auf gerichteten Evolutionstechniken, wie etwa stochastischer Mutagenese, halb-rationalen oder gezielten Strategien für das Zusammenstellen von Variantenbibliotheken und auch anspruchsvollen statistischen Selektionsverfahren wie ProSAR. Zwei der jüngsten Anwendungsbeispiele sind die asymmetrische Synthese von chiralen Aminen für die industrielle Produktion des Typ-2-Diabetes-Medikaments Sitagliptin (Januvia)^[106] und die oxidative Desymmetrisierung des prochiralen Amines für die Produktion des Hepatitis-C-Medikaments Boceprevir.^[107] In anderen Fällen haben Computerberechnungen zu einem erheblich besseren mechanistischen Verständnis künstlich evolvieter Enantioselektivitäten beigetragen.^[108,109] In den letzten fünf Jahren alleine wurden über 60 Publikationen veröffentlicht, die sich mit der Verbesserung von Thermostabilität, Substrat- und Cofaktorspezifität, Enantioselektivität und Reaktionsgeschwindigkeit natürlich vorkommender Enzyme befassten. Viele dieser Proteine wurden für Anwendungen in der asymmetrischen organischen Synthese hergestellt und setzen sich aus Transaminasen, Enoat-Reduktasen, Esterasen, Monoaminoxydases, Dehalogenasen, Aldolasen, aber auch Cytochrome-P450-Systemen (für Oxidationen und Epoxidierungen) und Baeyer-Villiger-Monooxygenasen zusammen. Das Thema ist jüngst Gegenstand einiger ausgezeichnetner Übersichtsartikel gewesen.^[87,110-113]

2.3. Natürliche Evolution und Enzym-Redesign

Milliarden Jahre der Evolution haben in der Natur zu einer enormen Vielfalt katalytischer Funktionen geführt. Die Untersuchung der sich dabei herauskristallisierten Strategien gewährt Einblicke, die es ermöglichen, gezielte Mutationen innerhalb der aktiven Zentren existierender Proteine vorzunehmen und diese umzuprogrammieren. Enzyme, die zu mechanistisch diversen Superfamilien gehören, sind wertvolle Startpunkte für solche Redesign-Bemühungen, vor allem

wenn sie einander strukturell sehr ähnlich sind. Elemente solcher Superfamilien sind häufig promisk, sodass ein Enzym oft eine Reihe chemischer Reaktionen katalysiert – obgleich mit geringerer Geschwindigkeit.^[116,114]

Einer der bisher erfolgreicheren Redesign-Ansätze beruht auf dem Abgleichen der Sequenz und Struktur von Enzymen innerhalb einer Superfamilie. Hierfür muss ein Enzym gewählt werden, das die gewünschte Reaktion zumindest zu einem geringen Maße katalysiert. Mechanistische Kenntnisse können dann von einem Enzym geborgt werden, welches die Zielreaktion bereits katalysiert. Zum Beispiel verglichen Fersht und Mitarbeiter die Sequenz von *N*-Acetylneuraminat-Lyase (NAL) mit der Sequenz der homologen Dhydrodipicolat-Synthase (DHDPS) und identifizierten eine Leu-Arg-Diskrepanz im aktiven Zentrum.^[115] Der Leu142Arg-Mutant wurde hergestellt, wobei ebenfalls an stabilisierende Mutationen für das neue Arg gedacht wurde. Das Ergebnis war eine NAL mit ausgetauschter Aktivität: Anstelle seiner nativen Retroaldol-Reaktion (*N*-Acetylneuraminat zu Pyruvat und *N*-Acetyl-d-mannosamin) katalysiert der Leu142Arg-Mutant die Kondensation zwischen Pyruvat und L-Aspartat- β -semialdehyd (die native Reaktion von DHDPS). Die Retroaldol-Aktivität von NAL verschwand, während die Geschwindigkeit der NAL-Kondensationsreaktion achtfach höher war als die des DHDPS. In einem ähnlichen Fall tauschte eine Leu-zu-Arg-Mutation die physiologische Aktivität von 4-Oxalocrotonat-Tautomerase (4-OT) zu der von *trans*-3-Chloracrylat-Dehalogenase (CaaD).^[116] Strukturelle Untersuchungen zeigten relativ geringe geometrische Änderungen. Der k_{cat} für die CaaD-Aktivität von 4-OT wurde um den Faktor 9 gesteigert, und k_{cat}/K_M stieg 50-fach. Ein etwas ambitionierteres Projekt änderte vier Aminosäuren im aktiven Zentrum von Keto-L-gulonat-6-phosphat-Decarboxylase (KGPDc) und erhöhte die Geschwindigkeit seiner Nebenaktivität für die Reaktion von D-Arabinose-hex-3-ulose-6-phosphat-Synthase (HPS) 170-fach.^[117,118]

Das Redesign von Enzymen zur Katalyse von Reaktionen, für die sie absolut keine Aktivität aufweisen, ist eine größere Herausforderung. Der Ansatz ist hier jedoch ähnlich wie zuvor und beginnt mit dem Abgleichen der Sequenz und Struktur von Enzymen innerhalb einer Superfamilie. Gerlt und Mitarbeiter vereinten dazu rationale Mutationen mit gerichteter Evolution und schufen L-Ala-D,L-Glu-Epimerase (AEE) und Muconat-lactonisierendes Enzym (MLE).^[119] Es ging dabei darum, die Funktionen von AEE und MLE so umzuschalten, dass beide die Reaktion von OSBS (*o*-Succinylbenzoat-Synthase) katalysieren, wobei weder AEE noch MLE eine messbare Anfangsaktivität für die OSBS-Reaktion aufweisen. Dieses scheinbare Kunststück wurde vollbracht, indem die Forscher sich auf die Substratspezifität konzentrierten: Asp-zu-Gly für AEE und Glu-zu-Gly für MLE erlaubte es beiden aktiven Zentren, das OSBS-eigene Substrat zu akzeptieren. Dieses reagiert bereitwillig mit den vorhandenen katalytischen Komponenten und ergibt *o*-Succinylbenzoat. Ohta und Mitarbeiter gingen einen Schritt weiter und entlockten der Arylmalonat-Decarboxylase (AMD) eine racemisierende Aktivität gegenüber α -Arylpropionat, indem sie eine katalytische Säure/Base in das aktive Zentrum ein-

führten (Gly74Cys).^[120] Weitere Versuche ergaben eine Doppelmutante (Gly74Cys, Cys188Ser) mit invertierter Enantioselektivität.^[121] Diese produziert (*R*)- α -Thienylpropionat mit einer Ausbeute von 60 % und einem Enantiomerenüberschuss (*ee*) von 84 %, erleidet jedoch einen ca. 600-fachen Aktivitätsverlust im Vergleich zur natürlich vorkommenden AMD. Stochastische Mutagenese verbesserte den k_{cat} 10-fach, sodass der Abstand zur natürlichen AMD daraufhin nur noch einem 60-fachen Aktivitätsverlust entspricht.^[122] Dunaway-Mariano und Mitarbeiter zeigten eindrucksvoll, dass Funktionen innerhalb der Crotonase-Superfamilie transplantiert werden können, indem man eine His-Asp-Dyade durch ein Glu-Glu-Säure/Base-Paar ersetzt.^[123] Die zwei Glutamat-Aminosäuren, die in das aktive Zentrum der 4-CBA-CoA-Dehalogenase eingeführt wurden, erforderten sechs weitere stabilisierende Mutationen. Der resultierende Octomutant zeigt zwar mit einem k_{cat} von 0.064 s^{-1} eine wesentlich geringere Aktivität als die Wildtyp-Crotonase ($k_{\text{cat}} = 1000 \text{ s}^{-1}$), die Übung veranschaulicht jedoch, dass „mit einer begrenzten Zahl von Mutationen ein völlig neuer katalytischer Pfad auf Kosten eines präexistenten generiert werden kann“.

Das Redesign von Enzymen innerhalb einer Superfamilie ist hilfreich, wenn es darum geht, Evolutionsstrategien und -prinzipien zu entziffern. Die chemische Vielfalt, die dem Proteiningenieur durch diesen Ansatz jedoch zur Verfügung steht, wird von der Superfamilie diktiert und ist im Allgemeinen stark begrenzt. Die Crotonase-Superfamilie ist dabei eine interessante Ausnahme, „in welcher gemeinsame strukturelle Merkmale durch natürlich Evolutionsprozesse zu einer bemerkenswerten funktionellen Vielfalt geführt haben“,^[124] sodass Reaktionen aller sechs Klassifizierungen der Enzyme Comission (EC) vertreten sind.

2.4. Rationales und De-novo-Proteindesign

2.4.1. Design und Vorhersage von Proteinstrukturen

Drastischere, computerbasierte Ansätze haben gezeigt, dass das Redesign ganzer Proteine möglich ist. Anfängliche Versuche befassten sich mit dem Redesign von Helix-Bündeln.^[125] Die Vorgehensweise beschränkte sich dabei auf die hydrophile Eigenschaft von Aminosäuren und ihre relative Position zueinander innerhalb der Proteinsequenz. Die Ausrichtung der Seitenketten muss dabei nicht ausdrücklich beschrieben werden.^[126–128] Computerbasierte Ansätze zum Design von Proteinen wurden auch auf weniger regelmäßige Strukturen erweitert,^[129] und die allgemeine Anwendbarkeit solcher Protokolle wurde unter anderem am Beispiel des Redesigns von kleinen globulären Proteinen getestet.^[130] Computermethoden wie „Dead End Elimination“ (DEE) und Monte-Carlo(MC)-Protokolle ermöglichen dabei das virtuelle Packen von Seitenketten zum Energieminimum hin. Der Anwendungsbereich solcher Ansätze ist jedoch nicht auf das Redesign existierender Topologien begrenzt. Beispielsweise iterierten Kuhlman et al. zwischen Sequenzdesign und Strukturvorhersage und erschufen mit Top7 eine neue Tertiärstruktur.^[131] Die Errungenschaft ist teilweise einem geometrischen Optimierungsschritt zuzuschreiben, der das Pep-

tidrückgrat während der Designsequenz den Umgebungswechselwirkungen anpasst.

2.4.2. Protein-Protein-Wechselwirkungen

Computerbasierte Methoden kamen auch beim Design von Grenzflächen für das Binden von Proteinen zum Einsatz. Huang et al. erzielten mikromolare Bindungsaffinitäten durch Variieren der Aminosäure und derer Seitenkettenorientierung an der Protein-Protein-Kontaktfläche.^[132] Stärkere Wechselwirkungen (130 nm) können erreicht werden, wenn Eigenschaften bereits vorhandener Grenzflächen als Anhaltspunkte genommen werden.^[90] Schlüsselpositionen, die für den Großteil der Bindungswechselwirkung verantwortlich sind, werden als „Hotspots“ ausgewählt und für ein Maximum der attraktiven Wechselwirkung in die Grenzfläche eingebaut. Die restliche Kontaktfläche wird dann mit stabilisierenden und packenden Aminosäuren ausgekleidet. Die um eine Stufe schwierigere Aufgabe, ein Bindungspeptid für ein spezifisches, biologisch relevantes Protein zu entwickeln, resultierte aus einem computerbasierten Design mit einer Affinität von 200 nm.^[133] Die Bindungskonstante konnte durch gerichtete Evolution auf 4 nm und 180 pm verbessert werden. Eine Untersuchung der resultierenden Mutationen suggerierte weiter, dass diese Verbesserungen durch ein fortgeschrittenes Computermodell – vor allem mit Bezug auf elektrostatische Wechselwirkungen, Solvatationsmodelle und ein flexibles Peptidrückgrat – ebenfalls erkannt werden können.^[133] DeGrado und Mitarbeiter benutzten jüngst ihr CHAMP-Protokoll (computed helical antimembrane protein method),^[134] um ein helikales β -Peptid herzustellen, das eine Transmembranhelix des $\alpha_{IIb}\beta_3$ -Integrins bindet.^[135] Die Gruppe benutzte ebenfalls computerbasierte Methoden für die Entwicklung eines helikalen Proteinaggregats entlang von Nanoröhren.^[136]

2.4.3. DNA-bindende Proteine

Die Herstellung von DNA-bindenden Proteinen ist ein weiteres interessantes Anwendungsgebiet computerbasierter Methoden. Eine der Strategien beruht auf der Kombination bereits existierender DNA-bindender Module, sodass die Problemstellung auf die Entwicklung von Protein-Protein-Wechselwirkungen umformuliert wird.^[137] Ein gezielterer Ansatz ergab das Redesign von Homing-Endonukleasen, welche einzelne Basenpaarunterschiede innerhalb einer DNA-Sequenz voneinander unterscheiden können.^[138,139] Die gleichzeitige Erfassung mehrerer Basenpaarunterschiede ist ebenfalls möglich.^[140] Erfahrungsgemäß ist dabei das Design solcher Mehrfach-Erfassungen mit einer höheren Erfolgswahrscheinlichkeit versehen. Proteine dieser Art wurden eingesetzt, um die Rolle der DNA-Sequenz bei Bindungs- und Katalyseereignissen am Beispiel der Homing-Endonuklease I-AnI zu untersuchen.^[141]

2.4.4. Protein-Ligand-Wechselwirkungen

Anfängliche Arbeiten im Bereich der Protein-Ligand-Wechselwirkungen waren vielversprechend und generierten

Berichte über das Design von Metall-,^[142,143] Lactat-,^[144] Serotonin-,^[144] TNT-^[144] und Nervengift-bindenden Proteinen.^[145] Es kamen jedoch Zweifel auf, als sich herausstellte, dass für die periplasmatischen Bindungsproteine, die für Lactat, Serotonin, TNT und Organophosphate entwickelt wurden, anhand von isothermischer Kalorimetrie (ITC) und NMR-Spektroskopie keine Aktivität gemessen werden konnte.^[146] Es wird vermutet, dass die anfänglichen Erfolgsberichte auf indirekte umgebungssensitive Fluoreszenzversuche zurückzuführen sind. In der Zwischenzeit wurden in diesem Bereich jedoch Fortschritt gemacht. Beispielsweise erforschten Harbury und Mitarbeiter periplasmatische Bindungsproteine mit computerbasierten Methoden und fanden heraus, dass das aktive Zentrum nur durch den Einsatz hochauflösender Rotamer-Suchen, kontinuierlicher Minimierung und akkurate Elektrostatischer Rechnungen rekapituliert werden kann.^[147] DeGrado und Mitarbeiter waren in der Lage, ein α -Helixbündel zu konstruieren, das einen Häm-ähnlichen Cofaktor binden kann.^[148] Jüngste Versuche, ein Dipeptid-bindendes Protein zu produzieren, schlugen jedoch fehl.^[149] Es wird vermutet, dass einer der Hauptgründe dafür darin liegt, dass die Flexibilität des aktiven Zentrums nicht ausreichend in Betracht gezogen wurde.

2.4.5. Katalytische Peptide und Proteine

Im Folgenden sind einige der frühen Arbeiten aufgeführt, die sich mit der Entwicklung funktioneller Proteine beschäftigten.

Johnsson et al. betteten ein reaktives Amin, welches als Schiff'sche Base fungieren kann, in eine amphiphile α -Helix ein und stellten somit eine metallfreie Oxalacetat-Decarboxylase (Oxaldie) her.^[150] Designte Oxaldies katalysieren die Decarboxylierung von Oxalacetat mit einem k_{cat}/K_M von $0.63 \text{ M}^{-1} \text{s}^{-1}$. Die Bildung des Imins ist hierbei um drei bis vier Größenordnungen schneller als bei einfachen Amin-Katalysatoren – und dadurch vergleichbar zu katalytischen Antikörpern (10^3 – 10^6).^[151]

Sasaki und Kaiser verknüpften vier amphiphile α -Helices mit einem Porphyrin-Ring und stellten somit das künstliche Häm-Protein „Helichrom“ her. Der Fe^{III}-Helichrom-Komplex erwies sich als hydrolytisch und katalysiert die Reaktion von Anilin zu *p*-Aminophenol in Gegenwart von NADPH mit einem k_{cat}/K_M von $1.67 \text{ M}^{-1} \text{s}^{-1}$.^[152]

Broo et al. entwickelten das Helix-Schleife-Helix-Peptid KO-42, das in ein antiparalleles Vier-Helix-Bündel dimerisiert und durch die kooperative Wirkung von Histidinen die Umesterung aktiverter Ester katalysiert.^[153] Rossi et al. betteten zwei und vier künstliche Triazacyclonanon-Aminosäuren in drei unterschiedliche Helix-Schleife-Helix-Peptide ein.^[154] Die somit generierten Zn^{II}-Bindungsstellen sind in der Lage, die Umesterung eines RNA-Modellsubstrats bis auf ein 380-faches zu katalysieren.

Dutton und Mitarbeiter untersuchten anhand von Drei-Helix-Bündeln (α_3W^1 und α_3Y^1),^[155] wie Seitenkettenradikale generiert, kontrolliert und für katalytische Zwecke genutzt werden können.^[156] Pecoraro und Mitarbeiter verwendeten ein α_3D -Bündel für die Platzierung dreier Cysteine, die Cd^{II}-, Hg^{II}- und Pb^{II}-Ionen binden können.^[157]

DeGrado und Mitarbeiter entwickelten eine O₂-abhängige Phenoloxidase, die auf einem heterotetrameren Helixbündel beruht.^[158] Die aktivste Variante katalysiert die Oxidation von 4-Nitrophenolacetat 1000-fach.

Bolon und Mayo verwendeten eine „Compute and Build“-Strategie und verliehen einem Thioredoxin-Protein anhand eines nukleophilen Histidins hydrolytische Eigenschaften. Die resultierende PZD2-Variante katalysiert die Hydrolyse von *p*-Nitrophenolacetat 180-fach.^[159]

3. Computerbasiertes „Inside-Out“-Design von Enzymen

In den letzten Jahren ist es mit zunehmender Verlässlichkeit möglich geworden, Sequenzen und dreidimensionale Strukturen miteinander zu verbinden. Beispielsweise war die Entwicklung von Top7^[131] ein Meilenstein im Bereich der inversen Proteinfaltung.^[160–163] Dessen Produktion zeigte, dass es möglich ist, Sequenzen zu konzipieren, die eine gewünschte dreidimensionale Struktur annehmen können. Während noch vieles in diesem Bereich zu tun ist, so stellt sich mit der Entwicklung funktioneller Proteine die nächste große Herausforderung.

Die Forschungsgruppen von Baker und Houk begannen dazu eine Reihe kooperativer Projekte, die zu der Entwicklung des „Inside-Out“-Protokolls führten (Abbildung 2). Im Kern des Protokolls steht ein theoretisches aktives Zentrum (Theozym, Abbildung 2, oberes Feld), das die passenden ka-

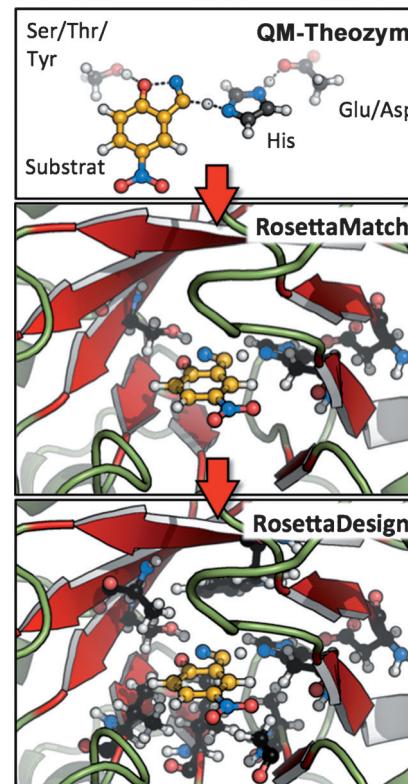


Abbildung 2. Hauptschritte des computerbasierten „Inside-Out“-Designprotokolls (am Beispiel der Kemp-Eliminierung).

talytischen Funktionsgruppen beinhaltet. Diese werden anhand von QM-Rechnungen so gewählt und ausgerichtet, dass der Übergangszustand maximal stabilisiert wird. Proteinstrukturen aus der PDB (<http://www.rcsb.org>)^[164] dienen als Vorlagen, in welche die QM-Geometrien der Übergangszustände eingebettet werden (RosettaMatch, Abbildung 2, mittleres Feld). Aminosäuren, die sich in der direkten Umgebung des QM-Theozyms befinden, werden mutiert und gepackt, sodass sie den geometrischen und elektrostatischen Eigenschaften des Übergangszustandes komplementär sind (RosettaDesign, Abbildung 2, unteres Feld).

3.1. Theozyme

Der erste Schritt des Inside-Out-Designprotokolls besteht aus QM-Rechnungen, anhand welcher dreidimensionale Anordnungen von Funktionsgruppen generiert werden, die den Übergangszustand der Zielreaktion stabilisieren.^[165] Ein solches Theozym (theoretisches Enzym) wird üblicherweise aus Aminosäureseitenketten und Proteinrückgrat-Amiden zusammengestellt, aber auch nichtnatürliche Aminosäuren und Cofaktoren können eingebettet werden. Dies führt im Normalfall zur Konstruktion mehrerer Theozyme, die sich in ihrer katalytischen Zusammensetzung voneinander unterscheiden. Die Energieprofile dieser theoretischen aktiven Zentren werden dann quantenmechanisch berechnet und anhand des daraus projizierten katalytischen Potentials bewertet. Die effektivsten Konstrukte werden anhand geometrischer Operationen (Spiegelung und Rotation) weiter diversifiziert, sodass für den nächsten Schritt ein Ensemble katalytisch kompetenter Theozym-Konformere zur Verfügung steht.

3.2. Theozym-Einbettung

Die aktiven Zentren existierender Proteine können anhand des RosettaMatch-Algorithmus^[166] nach solchen Rückgratpositionen durchsucht werden, die mit den dreidimensionalen Anforderungen eines Theozyms kompatibel sind. Das Programm versucht, eine Funktionsgruppe nach der anderen mit dem Proteinrückgrat zu verbinden, sodass zum Schluss das gesamte Theozym-Konstrukt in das aktive Zentrum einer existierenden Proteinstruktur eingebaut wird. Dazu werden Seitenkettenrotamere an denjenigen Positionen innerhalb des aktiven Zentrums generiert, die für die Verankerung des Theozyms in Frage kommen. Abweichungen von der idealen QM-Geometrie um auch nur einige Zehntel Angstrom und einstellige Winkelgrade können dabei Einbußen von bis zu 5 kcal mol⁻¹ bzw. vier Größenordnungen der Reaktionsgeschwindigkeit (k_{cat}) bewirken. Eine möglichst präzise Einbettung des Theozyms ist daher äußerst anstrengend, sodass im Idealfall dessen dreidimensionale Struktur vollständig erhalten bleibt. In der Praxis ist jedoch ein solcher Idealfall noch nicht verwirklicht worden. Dies hat mit der diskreten Natur des Proteinrückgrats, aber auch mit der des primären Abgleichsalgorithmus zu tun, und ist darüber hinaus durch den rechnerischen Aufwand limitiert,

der für ein vollständiges Abdecken aller möglichen Konformationen nötig wäre. Das Einbetten des Theozyms wird dann schnell zum Engpass des computerbasierten Designprotokolls, vor allem wenn mehr als drei katalytische Aminosäuren voneinander sind. Demzufolge führt eine exakte Behandlung im Normalfall zu keinem Ergebnis, sodass die Zuweisung von Toleranzwerten für die Abstände und Winkel, die ein Theozym beschreiben, erforderlich ist. Die eingebetteten Konstrukte entsprechen somit einer Anzahl deformierter Versionen der ursprünglichen QM-Geometrie und werden im nächsten Schritt nach ihrer Theozym-Ähnlichkeit bewertet. EDGE (enzyme design geometry evaluation) ist ein Programm, das eine geometrische Hash-Funktion verwendet, um Theozym-Atome mit einer Zielstruktur zu vergleichen, und wurde für diese Aufgabe entwickelt.

SABER (selection of active/binding sites for enzyme redesign), ein Programm das im Houk-Labor entwickelt wurde, stellt eine interessante Alternative zu RosettaMatch dar. Anstelle des Einbettens von Theozym-Konstrukten durchsucht SABER die PDB-Datenbank nach Proteinen, die bereits die katalytischen Theozym-Aminosäuren in der gewünschten Anordnung bereitstellen – auch wenn diese im ursprünglichen Protein nicht die gewünschte Funktion haben. Im nächsten Schritt werden nur diejenigen Aminosäuren mutiert, die der korrekten Bindungsorientierung des gewünschten Substrats im Wege stehen. Im Vergleich zum RosettaMatch-Ansatz ist hierbei im Allgemeinen eine geringere Anzahl von Mutationen nötig.

3.3. Design des aktiven Zentrums

Nach Einfügen des Theozyms in eine existierende Proteinstruktur – ob durch RosettaMatch oder SABER – wird das neue aktive Zentrum um die katalytischen Aminosäuren herum aufgebaut. Die Identität und Position der nicht-katalytischen Aminosäuren wird dabei anhand eines Monte-Carlo-Algorithmus in RosettaDesign optimiert, sodass stabilisierende Theozymkontakte und Eigenwechselwirkungen entstehen. Der Ausbau des neuen aktiven Zentrums erfordert mehrere Runden der Rotamersuche, Seitenketten- und Rückgratoptimierung sowie auch Starrkörperbewegungen des Übergangszustandes. Die Geometrie des Theozyms wird während des gesamten Ablaufs konserviert und nur zum Schluss gelockert.^[167] Im Idealfall ergibt sich daraus ein aktives Zentrum, das die Orientierung der katalytischen Aminosäuren unterstützt, deren pK_s-Werte abstimmt und den Übergangszustand stärker bindet als den Grundzustand. In der Praxis jedoch ist jedes Theozym bereits nach dem Einbetten um ein nicht unerhebliches Maß von der idealen QM-Geometrie deformiert. RosettaDesign hat dann die Aufgabe, dieses suboptimale Konstrukt zu stabilisieren und – mit den geometrischen Vorgaben von der Quantenmechanik – eventuell zu verbessern. Zum Schluss weichen jedoch auch die besten Designs erheblich von der idealen Vorlage ab. Abbildung 3 zeigt dies anhand von vier Designs für die Kemp-Eliminierung (KE), deren katalytische Atome mit denen des Theozyms überlagert sind. Die jeweiligen Seitenketten dieser Designs sind zwar zusammen gruppiert (Abbildung 3a),

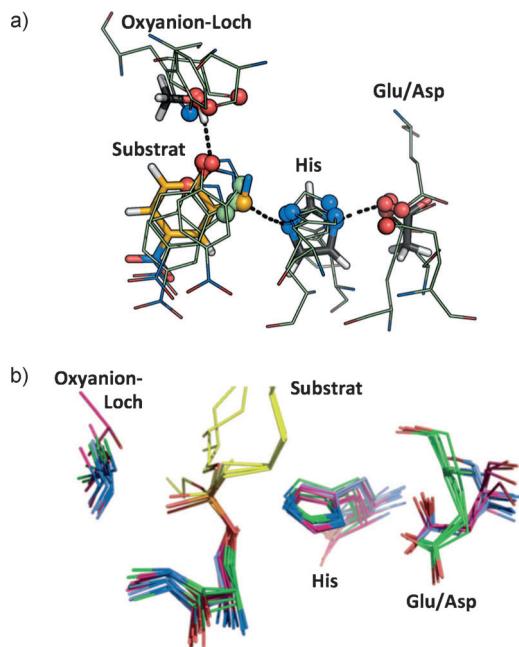


Abbildung 3. Geometrische Überlagerung der katalytischen Atome. a) Theozym (schwarz/orange) über vier fertiggestellte Designs in der TIM-Barrel-Falte (hellgrün). Katalytische Atome (Kugeln). RMSD-Werte der KE-Designs mit einer His-Glu/Asp-Dyade im Vergleich zum Theozym: 1.2 Å für KE70 (aktiv), 0.8 Å für KE38 (inaktiv), 0.8 Å für KE54 (inaktiv), 0.5 Å für KE66 (inaktiv). b) Katalytische Triade von Esterase-Kristallstrukturen; RMSD = 0.45 Å.^[168]

zeigen jedoch bei weitem nicht die Präzision auf, die für natürlich vorkommende Enzyme innerhalb einer katalytischen Klasse typisch ist (Abbildung 3b).^[168]

3.4. Filtern, Ranking und Auswerten von computergenerierten Designs

Bevor die experimentelle Aufarbeitung beginnt, wird die katalytische Kompetenz der fertigen Designs rechnerisch bewertet. Dazu werden empirische Faktoren herangezogen wie etwa Rosetta-Energien, Bindungswerte, Wasserstoffbrückenbindungsrate, die Geometrie des aktiven Zentrums und Packwerte. Diese Faktoren werden mit denen des nativen Proteins verglichen, sodass ein gegebenes Design im ungefähren Kontext betrachtet werden kann. Bisher hing die Qualitätskontrolle fertiger Designs stark von der chemischen Intuition des Forschers ab, sodass eine humane Komponente vor allem oft dann als ein kompensierendes Element zum Einsatz kam, wenn Rosetta nicht glänzte. Beispielsweise sind die aktiven Zentren natürlich vorkommender Enzyme in der Regel durch Netzwerke von Wasserstoffbrücken, hydrophobes Packen, π-π-Wechselwirkungen, eingeschränkte Bewegungen und begrenzte Lösungsmittelzugänglichkeit gekennzeichnet. Gegenwärtig werden einige dieser Effekte bei den Berechnungen implizit miteinbezogen (Rosetta ist ein teilweise wissensbasiertes Programm), sodass günstige Wechselwirkungen automatisch gefördert und ungünstige benachteiligt werden. Das explizite Einfügen positiver Wechselwir-

kungen kann als zweite Stufe betrachtet werden, für welche wir erst jetzt mit der Aufstellung automatischer Vorgehensweisen beginnen. Werkzeuge wie Foldit, EDGE, interne Skripte und gründlichere computerbasierte Tests, die das dynamische Verhalten der Designs proben, werden gegenwärtig entwickelt und verfeinert, mit dem Ziel, die Erfolgsquote zu maximieren, vor allem wenn anspruchsvollere Reaktionen verfolgt werden sollen.

So kam in der Entwicklung der Kemp-Eliminasen und Retro-Aldolasen die erste Version des RosettaMatch- und RosettaDesign-Algorithmus zum Einsatz, wobei nur 87 handkurierte Proteinstrukturen als Vorlagen für die neuen aktiven Zentren dienten. Darüber hinaus stand für das Einbetten der Theozyme lediglich ein diskreter Algorithmus zur Verfügung, und die Proteinrückgrate wurden als vollständig starr behandelt. Die Qualitätskontrolle beruhte hauptsächlich auf manueller Inspektion. Bei der Entwicklung der bimolekularen Diels-Alderasen kam eine leistungsfähigere Version von Rosetta zum Einsatz. Es konnten nun 227 Proteinstrukturen als Vorlagen verwendet werden, wobei die fertigen Designs – über manuelle Inspektionen hinaus – auch anhand von MD-Simulationen untersucht wurden.

Mit der gegenwärtigen Kollektion von Rosetta-Modulen (Rosetta3) können Proteinstrukturen aus der gesamten PDB-Datenbank als Vorlagen verwendet werden. Ein sekundärer, kontinuierlicher Einbettungsalgorithmus ergänzt den primären und verleiht dem Proteinrückgrat ein geringes Maß an Plastizität, sodass eine gegebene Proteinstruktur sich dem neuen aktiven Zentrum besser anpassen kann. MD-Simulationen sind besonders hilfreich, wenn es darum geht, die strukturelle Integrität eines fertigen Designs zu testen und eventuelle Mängel innerhalb des aktiven Zentrums aufzuspüren, die sich statischen Untersuchungen entziehen.^[169]

Die Sequenz eines fertigen Designs unterscheidet sich oft um 10% oder mehr von der des Vorlagenproteins. Es kann dabei davon ausgegangen werden, dass im neuen Protein die Seitenkettenpackung und das Wasserstoffbrückennetzwerk weniger ideal ausgeprägt ist als im Vorlagenprotein. Das schrittweise Optimieren und Neupacken der Seitenketten und des Rückgrats stellt während des Designschrittes sicher, dass sich die fertige Struktur in einem Minimum seiner Konformationslandschaft befindet. Die nicht unerheblichen Perturbationen, die im Verlauf des Designvorgangs eingearbeitet werden, können jedoch dazu führen, dass benachbarte Minima thermodynamisch begünstigt werden. Der tatsächliche Konformationszustand des aktiven Zentrums kann sich demnach erheblich von dem des computerbasierten Designs unterscheiden. Ob und zu welchem Ausmaß dies geschieht, kann anhand von MD-Simulationen getestet werden, sodass entsprechende Auswertungen fertiger Designs mittlerweile routinemäßig durchgeführt werden, damit strukturelle und funktionelle Mängel identifiziert und beseitigt werden können.

3.5. Experimente

Abgesehen von der Herkunft der Gene (chemische Synthese anstelle des Klonens) kommt die experimentelle Be-

wertung von computerbasierten Designs herkömmlichen Aktivitätsmessungen natürlicher Enzyme gleich.

3.5.1. Synthese und Expression

Die optimierten genetischen Sequenzen wurden im Fall der Retro-Aldolasen,^[170] der Kemp-Eliminasen^[170,171] und der Diels-Alderasen^[172] kommerziell synthetisiert, Codon-optimiert und in einen herkömmlichen His-Tag-*E. coli*-Expressionsvektor kloniert. BL21(DE3)-Zellen aus *E. coli* wurden daraufhin mit dem Plasmid transformiert und unter herkömmlichen IPTG- oder Autoinduktionsbedingungen exprimiert.^[173] Lösliche Proteine wurden daraufhin mittels konventioneller IMAC-Reinigung^[174–176] und Gelfiltration erhalten.

3.5.2. Enzymbestimmung

Aktive computerbasierte Enzyme katalysieren die Zielreaktion anfänglich meist nur geringfügig, sodass deren Detektion oft von der Empfindlichkeit der Messmethode abhängig ist. Die Wahl eines geeigneten Testsubstrats kann dabei diesem Sachverhalt entgegenkommen. So wurden die Kemp-Eliminasen von Röthlisberger et al. und die Retro-Aldolasen von Jiang et al. auf UV/Vis- bzw. Fluoreszenzaktive Substrate hin entwickelt, sodass die Produktbildung über den Reaktionszeitraum hinweg relativ unkompliziert verfolgt werden konnte. Die Diels-Alderasen von Siegel et al. wurden hingegen auf eine Reaktion hin entwickelt, die spektrophotometrisch inaktiv ist. Produktbildung wurde hier anhand von LC/MS-Datenpunkten über Tage hinweg gemessen. Chirale LC/MS-Bestimmungen ermöglichen jedoch, dass die entwickelten Enzyme auf ihre Stereoselektivität hin untersucht werden konnten.

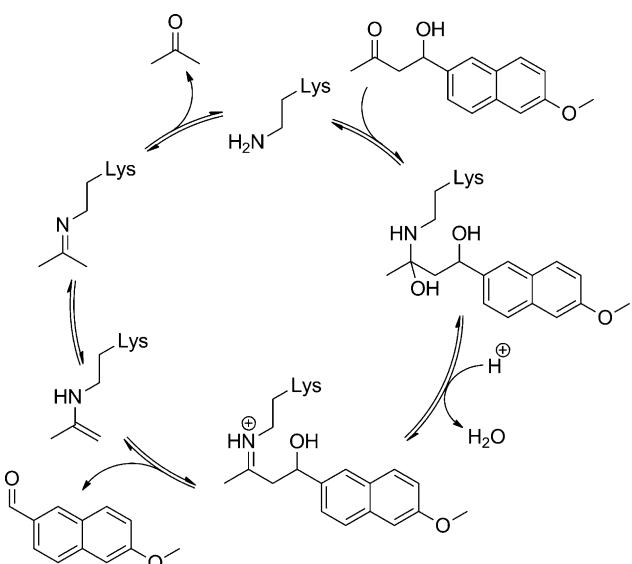
3.5.3. Gerichtete Evolution

Die typischerweise geringe Aktivität computerbasierter Enzyme kann im Normalfall durch mehrere Runden stochastischer Mutagenese und gezielter Diversifizierung weiter verbessert werden. Rechnerische Analysen können ebenfalls bei der Wahl der Mutationen (und Insertionen) der anschließenden Runden miteingebracht werden. Spezifische Beispiele sind in Abschnitt 4.2.3 näher beschrieben.

4. Computerbasiertes Design von Enzymen – Erfolgsbeispiele

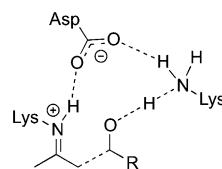
4.1. Retro-Aldolasen

Das „Inside-Out“-Enzymdesignprotokoll kam das erste Mal bei der Entwicklung von Retro-Aldolasen zum Einsatz. Deren aktive Zentren katalysieren den C-C-Bindungsbruch in 4-Hydroxy-4-(6-methoxy-2-naphthyl)-2-nutanon.^[170] Der Mechanismus beruht – ähnlich wie bei Klasse-1-Aldolasen – auf einem nukleophilen Lysin, das mit dem Substrat eine Schiff'sche Base als Intermediat bildet (Schema 8).^[177]

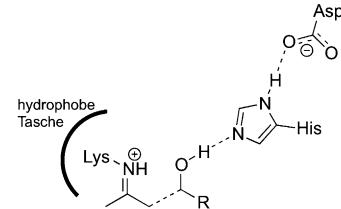


Schema 8. Schritte der Amin-katalysierten Retro-Aldolreaktion von 4-Hydroxy-4-(6-methoxy-2-naphthyl)-2-butanon.

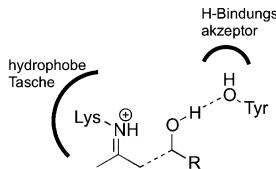
Motiv I: 12 Designs / 0 aktiv



Motiv II: 14 Designs / 8 aktiv



Motiv II: 10 Designs / 0 aktiv



Motiv IV: 38 Designs / 22 aktiv

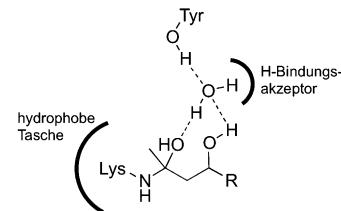


Abbildung 4. Die Theozyme der computerbasierten Retro-Aldolasen.

Der Designprozess beruht auf vier verschiedenen Theozymen (Abbildung 4). Diese enthalten sowohl ein Lysin als Schiff'sche Base als auch einen Säure/Base-Katalysator (I: Lys/Asp-Dyade; II: Tyr; III: His/Asp-Dyade; IV: H₂O) für die Deprotonierung des β-Alkohols. Der Lys-Asp-Lys-vermittelte Protonentransfer in Motiv I ähnelt dem aktiven Zentrum von D-2-Deoxyribose-5-phosphat-Aldolase.^[178] Motive II, III und IV imitieren das aktive Zentrum von katalytischen Antikörpern, in welchen ein Lysin zur Reduktion seines pK_S-Wertes in eine hydrophobe Tasche eingebettet ist.

Die räumlichen Anordnungen der vier Motive stammen von QM-Theozym-Berechnungen, die für jeden Schritt entlang des Reaktionspfades durchgeführt wurden. Die Übergangszustände der einzelnen Schritte wurden daraufhin miteinander verbunden, sodass ein einzelnes Theozym entstand,

das die geometrischen und elektrostatischen Eigenschaften des gesamten Reaktionspfades vereint. Die resultierenden Konsensus-Theozyme wurden durch das Scannen interner Freiheitsgrade sowie der Ausrichtungen und Konformationen der katalytischen Aminosäuren diversifiziert, sodass 10^{13} – 10^{18} einzelne Anordnungen generiert wurden. Diese Theozym-Motive wurden anhand der Hash-Funktion in Rosetta-Match^[166] in 71 Proteinvorlagen eingebettet, sodass rund 180 000 Lösungen entstanden. Die aktiven Zentren wurden mit RosettaDesign auf eine maximale Stabilisierung des Übergangszustands und des katalytischen Lysins hin ausgebaut. 72 dieser Proteine wurden für die experimentelle Aufarbeitung und Charakterisierung ausgewählt. Die Kriterien der Qualitätskontrolle beruhten auf a) der berechneten Bindungsenergie des Übergangszustands, b) der geometrischen Übereinstimmung zwischen fertigem Design und Theozym, c) der Seitenkettenpackung um das katalytische Lysin herum und d) der Beständigkeit der Seitenkettenkonformationen.

70 der 72 Proteine waren löslich, nachdem sie exprimiert und gereinigt wurden, und beachtliche 32 zeigten messbare Retro-Aldol-Aktivitäten. Die Produktbildung wurde anhand von Fluoreszenzmessungen verfolgt. Die aktiven Enzyme erstrecken sich über fünf verschiedene Proteinvorlagen (1mw4, 1f5j, 1thf, 1i4n, 1a53) der Triosephosphat-Isomerase-(TIM)-Barrel und der Jelly-Roll-Falte und beruhen auf den Theozym-Motiven III und IV. Die relativ offenen aktiven Zentren der Jelly-Roll-Designs sind durch eine einfache lineare Kinetik charakterisiert, während für die eingeschlosseneren TIM-Barrel-Designs ein komplexeres kinetisches Profil gemessen wurde. Dies kann auf einen eingeschränkten Substratzugang und/oder auf Produktinhibition hinweisen. Die Kristallstrukturen zweier Varianten (RA22_S210A und RA61_M48K) wurden mit einer jeweiligen Auflösung von 2.2 und 1.9 Å bestimmt und decken sich mit den computergenerierten Modellen.^[170] $k_{\text{cat}}/k_{\text{uncat}}$ -Werte reichen zwar bis 10^4 , liegen aber um zwei bis drei Größenordnungen unter denen katalytischer Antikörper.^[60,61] Die katalytische Effizienz (k_{cat}/K_M) der computerbasierten Retro-Aldolasen reicht von 0.02 bis $0.74 \text{ M}^{-1} \text{s}^{-1}$ und ist ebenfalls relativ bescheiden, vor allem im Vergleich zu natürlichen Enzymen.

Ruscio et al. untersuchten den Einfluss struktureller Fluktuationen auf die Präorganisation des aktiven Zentrums in RA22^[179] und beobachteten, dass das Substrat in einer alternativen Orientierung gebunden werden kann. Die Autoren fanden weiterhin, dass die His233-Asp53-Dyade aufgrund der Solvatisierung des Aspartats unterbrochen ist, was wiederum Konsequenzen für die Wechselwirkung zwischen Substrat und His233 hat. Die relativ geringe Aktivität von RA22 wurde dieser dynamischen Verzerrung des Deprotonierungsschrittes zugeschrieben.

Lasilla et al. zeigten jüngst, dass die designierten Wasserkontakte mit Tyr78 und Ser87 in der Retro-Aldolase RA61 keinen Beitrag zur Katalyse leisten.^[180] Die Autoren schrieben die Aktivität dieses Designs stattdessen sowohl dem stark nukleophilen Charakter des katalytischen Lysins ($\text{p}K_s = 6.8$ –7.5) als auch den günstigen Bindungswechselwirkungen zwischen der Naphthylgruppe des Substrats und dem hydrophoben aktiven Zentrum zu.

4.2. Kemp-Eliminasen

4.2.1. Computerbasiertes Design

Die Kemp-Eliminierung (Abbildung 5 a) ist eine ausgiebig untersuchte Ringöffnungsreaktion, die durch Deproton-

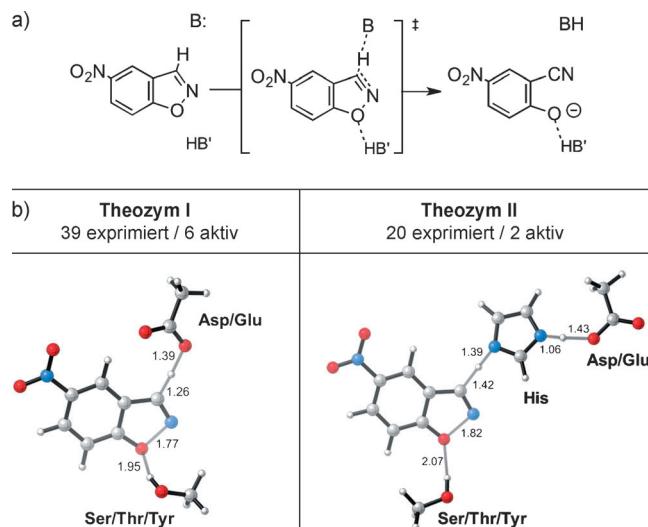


Abbildung 5. Die Kemp-Eliminierung. a) Reaktionsschema für 5-Nitrobenzisoxazol. b) Die zwei verwendeten Theozyme.

nierung des Substrats eingeleitet wird. Obwohl es zu der Reaktion kein natürliches Gegenstück gibt, so dient sie dennoch als weitläufiges Modell für die biochemisch relevante Deprotonierung von Kohlenstoffatomen. Die Reaktion ist in der Vergangenheit zu einem attraktiven Ziel in der Entwicklung von Säure/Base-Katalysatoren geworden, mit Beispielen, die von katalytischen Antikörpern^[51] bis zu „Synzymen“ reichen.^[181] DeGrado und Mitarbeiter gelang es jüngst, durch einen minimalistischen Designansatz Calmodulin in eine Ca^{2+} -abhängige Kemp-Eliminase umzuwandeln.^[182]

Die Reaktionsgeschwindigkeit der Kemp-Eliminierung ist stark lösungsmittelabhängig – vor allem wenn die katalytische Base ein Carboxylat ist –, sodass mit Acetat alleine (in einem polaren, unprotischen Lösungsmittel wie Acetonitril) Geschwindigkeitserhöhungen von bis zu 10^7 erreicht werden können.^[183] Eine zusätzliche Beschleunigung von bis zu 10^6 kann durch das präzise Positionieren von Donor und Akzeptor erzielt werden^[53,183,184] sodass die Kemp-Eliminierung ein theoretisches Limit von 10^{13} hat.

Röthlisberger et al. entwickelten anhand des „Inside-Out“-Protokolls das erste aktive Zentrum für eine nichtnatürliche Reaktion und stellten acht aktive Enzyme für die Ringöffnungsreaktion von 5-Nitrobenzisoxazol (5-NBZ) her.^[171] Das Protokoll verwendete zwei Theozyme (Abbildung 5b) und erzielte Geschwindigkeitserhöhungen von bis zu 10^5 . Die Kristallstruktur des aktiven KE07-Designs ($k_{\text{cat}}/K_M = 12.2 \text{ M}^{-1} \text{s}^{-1}$) ist in guter Übereinstimmung mit dem Computermodell.

Die kinetischen Kennzahlen dieser acht Kemp-Eliminasen sind vergleichbar mit denen von katalytischen Antikörpern. Die acht Geschwindigkeitserhöhungen ($k_{\text{cat}}/k_{\text{uncat}}$) reichen von 10^3 bis 10^5 , wobei die beiden katalytischen Anti-

körper 34E4 und 35F10 ein $k_{\text{cat}}/k_{\text{uncat}}$ von 10^4 aufweisen.^[51] Die beiden Antikörper (K_M von 0.6 bis 0.1 mm) übertreffen die Computerdesigns jedoch 10-fach in ihrer Fähigkeit, das Substrat zu binden (K_M von 4.2 bis 0.6 mm). Die kinetischen Charaktere „Inside-Out“-Eliminasen wurden durch gerichtete Evolutionsexperimente weiter verbessert. k_{cat}/K_M des KE07-Designs konnte um einen Faktor 200 erhöht werden,^[185] das von KE70 über 400-fach,^[186] und im Falle des KE59-Designs konnte für k_{cat}/K_M eine über 2000-fache Verbesserung erreicht werden.^[187] Dabei wird deutlich, dass computerbasiertes Design und gerichtete Evolutionstechniken Hand in Hand effektiv eingesetzt werden können, um substantielle katalytische Werte zu erzeugen.

4.2.2. Computerbasierte Untersuchungen von de novo entworfenen Kemp-Eliminasen

4.2.2.1. PDDG/PM3 Monte Carlo

Alexandrova et al. analysierten vier der acht aktiven Röthlisbergerschen Kemp-Eliminasen, KE07 (258 Aminosäuren), KE10 (253), KE15 (258) und KE16 (258), durch Berechnung der Reaktionskoordinaten mit einem semiempirischen PDDG/PM3-QM/MM-Monte-Carlo-Ansatz.^[188] Die rechnerische Zusammenstellung der Systeme bestand aus 200 Aminosäuren großen Ausschnitten der vier Proteine. Der semiempirische QM-Bereich setzte sich aus dem 5-NBZ-Substrat und der katalytischen Base (Glu/Asp) zusammen. Wassermoleküle waren nicht Teil des PM3-Bereichs. Seitenkettenkonformationen wurden abgetastet, das Proteinrückgrat wurde jedoch als starr behandelt. Der Versuch, eine Korrelation zwischen den berechneten und experimentell gemessenen Aktivierungsenergien aufzustellen und dadurch die Ursache hinter den beobachteten Aktivitäten zu untersuchen, war nur begrenzt erfolgreich. Die berechneten Werte weisen große absolute Ungenauigkeiten auf, und sogar der relative Trend ist dem der experimentell gemessenen Werte entgegengesetzt. Es sollte dabei jedoch bemerkt werden, dass die Aktivitäten dieser vier Kemp-Eliminasen sich lediglich über einen ΔG^\ddagger -Bereich von $0.9 \text{ kcal mol}^{-1}$ erstrecken – zu eng für aussagekräftige Untersuchungen mit den meisten modernen QM-Methoden.

4.2.2.2. Quantenmechanische DFT-Rechnungen

Sechs aktive und vier inaktive Kemp-Eliminasen, mit ΔG^\ddagger -Werten zwischen 18.3 und $20.6 \text{ kcal mol}^{-1}$ (aktiv), aber auch mit $\Delta G^\ddagger \geq 23.2 \text{ kcal mol}^{-1}$ (inaktiv) wurden mit DFT-Methoden untersucht.^[169] Drei Ansätze wurden auf ihre Anwendbarkeit hin zur Auswertung von De-novo-Designs getestet. Diese erstreckten sich von einer minimalistischen Repräsentation der katalytischen Einheit (Abbildung 6, rechts oben) zu QM am ganzen aktiven Zentrum bis zu QM/MM-Rechnungen am kompletten Protein (Abbildung 6, links) – nach kurzen MD-Simulationen –, wobei das aktive Zentrum und einzelne Wassermoleküle mit QM behandelt wurde (Abbildung 6, rechts unten).

Qualitativ betrachtet, rekapituliert der MD-QM/MM-Ansatz die experimentell bestimmten katalytischen Reaktionsgeschwindigkeiten am besten; die berechneten Aktivie-

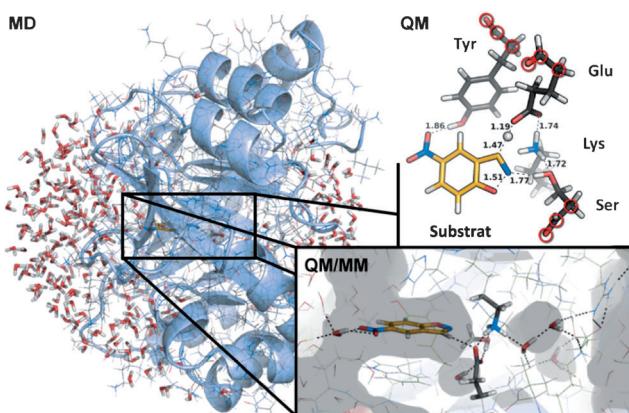


Abbildung 6. Ansätze zur Untersuchung von de novo entworfenen Kemp-Eliminasen. Rechts oben: QM an der katalytischen Einheit, wobei die Rückgratatome (eingekreist) festgehalten wurden. Links: ONIOM-QM/MM nach 2 ns MD am gesamten Enzym, wobei das aktive Zentrum mit QM behandelt wurde (rechts unten). Modifizierter Nachdruck von Kiss et al.^[169]

rungsenergien der inaktiven Proteine sind bedeutend höher als die der aktiven Proteine. Abgesehen davon jedoch zeigt der Ansatz nur eine schwache Korrelation ($R^2 = 0.58$) mit den experimentell bestimmten Energien, was darauf hindeutet, dass auch diesem Ansatz einige nicht unbedeutende katalytische Beiträge entgehen. Eine Lehre aus diesem und vorangehenden Unterfangen ist, dass bei Berechnungen von basenkatalysierten Reaktionen – wie der Kemp-Eliminierung – der Miteinbezug von polaren Gruppen und Wassermolekülen als Teil des QM-Bereichs unumgänglich ist.

4.2.2.3. Empirische Valenzbindungs(EVB)-Rechnungen

Warshel und Mitarbeiter untersuchten in zwei Instanzen fünf aktive de novo entworfene Kemp-Eliminasen, KE07,^[189] KE70, KE59, HG-2 und 1A53-2,^[190] mit einer doppelschichtigen EVB-Methode, wofür Gibbs-Energie-Perturbationsrechnungen mit Umbrella-Sampling gekoppelt wurden (FEP/US). Die EVB-Konfiguration wurde für eine möglichst genaue Reproduktion von Ab-initio-Rechnungen der Reaktionsfläche innerhalb eines Lösungsmittelkäfigs kalibriert und daraufhin zur Berechnung von Aktivierungsenergien (ΔG^\ddagger) benutzt. Für einige der Kemp-Eliminasen stimmen diese mit experimentell hergeleiteten Werten außergewöhnlich gut überein, für andere jedoch (z. B. KE59, HG-2) sind die Abweichungen erheblich (Tabelle 1).

Tabelle 1: Mit EVB berechnete Gibbs-Aktivierungsenergien für computerbasierte Kemp-Eliminasen.^[189, 190]

System	PDB-Eintrag	Base	ΔG^\ddagger ^[a]	ΔG^\ddagger ^{EVB} ^[a]
HG-2 (S265T)	NA	Asp127	17.7	18.2
34E4-Antikörper	1vol	Glu ^{H50}	17.9	17.3
KE59	NA	Glu231	18.3 ^[b]	31.7
HG-2	3nyd	Asp127	18.5 ^[c]	24.3
KE70	3npu	His16-Asp44	18.5	19.3
1A53-2	3nyz	Glu178	20.0	20.7
KE07	2rkx	Glu231	20.1	19.5

[a] In kcal mol^{-1} . [b] Berechnet mit $k_{\text{cat}} = 0.29 \text{ s}^{-1}$.^[171] [c] Berechnet mit einem extrapolierten $k_{\text{cat}} = 0.22 \text{ s}^{-1}$.^[191]

4.2.2.4. MD-Simulationen und die Dynamik neuer aktiver Zentren

MD-Simulationen an 23 Kemp-Eliminasen (14 aktiv, 9 inaktiv) waren, wenngleich von eher qualitativer Natur, aufschlussreicher als vorangehende Untersuchungen. Die Auswertung der Simulationen zeigte, dass die aktiven Zentren der fehlgeschlagenen computerbasierten Designs nicht dazu in der Lage sind, katalytisch essenzielle Kontakte aufrechtzuerhalten.^[169] Abbildung 7c verdeutlicht dies am Beispiel des inaktiven KE38. Verglichen mit dem katalytischen His-Asn-Kontakt im aktiven Zentrum des natürlich vorkommenden Cathepsin K (Abbildung 7a) und der katalytischen His-Asp-Dyade des aktiven KE70-Designs (Abbildung 7b) zeigt die KE38-Simulation keine nennenswerte Population, in welcher die His-Glu-Dyade intakt ist. Histidin alleine ist ohne diesen Kontakt eine zu schwache Base, als dass ein katalytischer Effekt zu erwarten wäre.

In den untersuchten Systemen konnte der Zerfall der eingebauten katalytischen Kontakte einer Kombination zweier Gründe zugeschrieben werden: exzessive Lösungsmittelzugänglichkeit und alternativ gepackte Seitenketten, wobei beide Ereignisse ihr eigenes Verteilungsmuster haben. Die Beobachtungen sind von allgemeiner Bedeutung für das Konzeptionieren neuer Enzyme, aber umso mehr, wenn eine Carboxylat-Base als zentrales katalytisches Element fungieren soll. Polare, protische Lösungsmittel, die in direkten Kontakt mit den Carboxylat-Sauerstoffatomen treten, können deren Basizität erheblich reduzieren (bis zu sechs Größenordnungen, was k_{cat} betrifft). Abbildung 8a zeigt diesen Trend für einen Querschnitt der untersuchten Systeme. Die aktiven Zentren funktionstüchtiger Kemp-Eliminasen sind im Durchschnitt weniger stark hydriert als die der inaktiven (Abbildung 8b). Jedoch stehen selbst die aktivsten dieser computerbasierten Designs natürlichen Enzymen, wie etwa Cathepsin K, auch in dieser Hinsicht nach.

Aktive De-novo-Enzyme können also qualitativ von inaktiven Proteinen unterschieden werden, wenn, wie am Beispiel der Kemp-Eliminasen, ein multidimensionales Problem in ein vereinfachtes zweidimensionales Modell reduziert werden kann (Abbildung 9).

Aus dieser Studie geht somit hervor, dass die dynamische Erkundung eines Protein-Substrat-Komplexes in Gegenwart expliziter Lösungsmittelmoleküle (wenn gekoppelt mit der Abfrage intuitionsbasierter Größen) eine Fülle aussagekräftiger Daten hervorbringen kann, die es ermöglichen, Zusammenhänge zwischen experimentellen Beobachtungen herzuleiten. Auf dieser Basis ist es zu einer nützlichen Praxis geworden, MD-basierte Untersuchungen mit der Herstellung und Verbesserung neuer Enzyme sowie der Interpretation gerichteter Evolutionsversuche zu kombinieren.

4.2.3. Gerichtete Evolution der Kemp-Eliminasen KE07, KE70 und KE59

Tawfik und Mitarbeiter vereinten gerichtete Evolutionsmethoden mit rationalem Design und waren in der Lage, die katalytischen Aktivitäten dreier Kemp-Eliminasen zu verbessern.

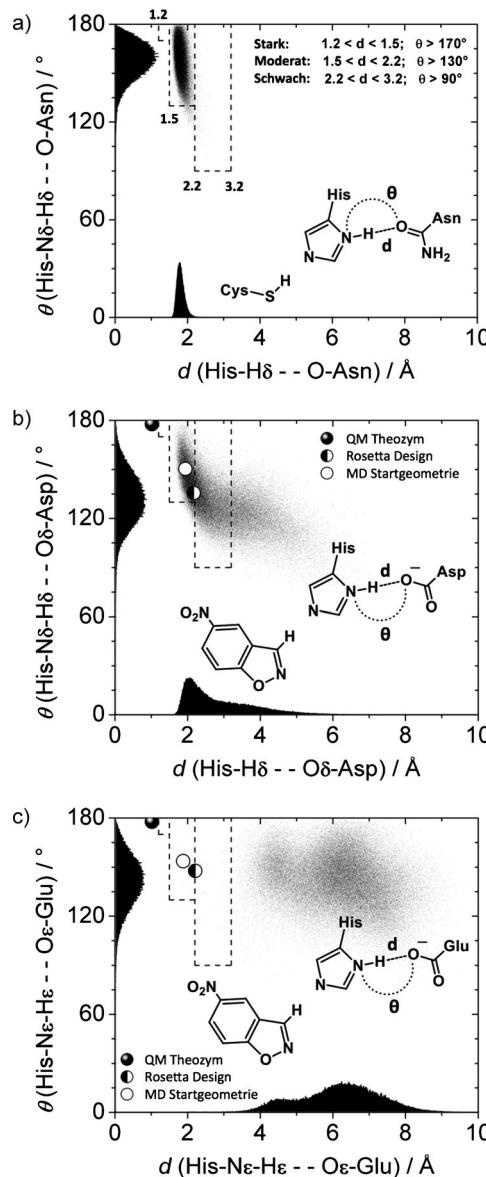


Abbildung 7. Streudiagramm der katalytischen Kontakte, wobei Winkel (θ) gegen Distanzen (d) aufgetragen sind. a) His-Asn-Kontakt der katalytischen Triade im natürlichen Enzym Cathepsin K; b) His-Asp-Kontakt des aktiven KE70-Designs; c) His-Glu-Dyade des inaktiven KE38-Designs. Aufgetragen sind jeweils 100000 Datenpunkte von 20 ns MD-Simulationen. Drei Wasserstoffbrückenkategorien^[192] sind als Umrisse dargestellt. Die Komponenten der jeweiligen Verteilungen sind als Achsenprojektionen dargestellt. Die Entwicklung der jeweiligen Kontakte als Teil des Designprotokolls (von Theozym zu fertigem Design und MD-Startgeometrie) ist anhand von Kreisen aufgetragen. Modifizierter Nachdruck von Kiss et al.^[169]

4.2.3.1. KE07

Sieben Runden der stochastischen Mutagenese und Selektion führten zu einer 200-fachen Zunahme des k_{cat}/K_M -Wertes im Vergleich zum dem des computerbasierten „Wildtyp“-KE07-Designs.^[185] Der Aktivitätsanstieg setzte sich aus einem 2.6-fach verminderten K_M und einem 76-fach erhöhten k_{cat} zusammen, was zum Großteil der Ile7Asp-Mutation nahe dem aktiven Zentrum zuzuschreiben ist. Asp7

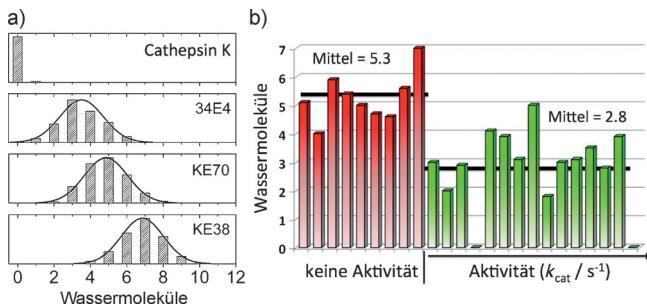


Abbildung 8. Verteilung der Wasseroordinationszahlen in MD-Simulationen ($d < 3.2 \text{ \AA}$ von den katalytischen Carboxylat-Sauerstoffatomen). a) Asn182 im natürlichen Enzym Cathepsin K ($k_{cat} = 4.2 \text{ s}^{-1}$), Glu^{H50} im katalytischen Antikörper 34E4 ($k_{cat} = 0.7 \text{ s}^{-1}$), Asp44 im aktiven KE70-Design ($k_{cat} = 0.2 \text{ s}^{-1}$) und Glu170 im inaktiven KE38-Design. Nachdruck von Kiss et al.^[169] b) Maxima dieser Verteilungen für die 23 untersuchten Kemp-Eliminasen, einschließlich des katalytischen Antikörpers 34E4 und des natürlichen Enzyms Cathepsin K.

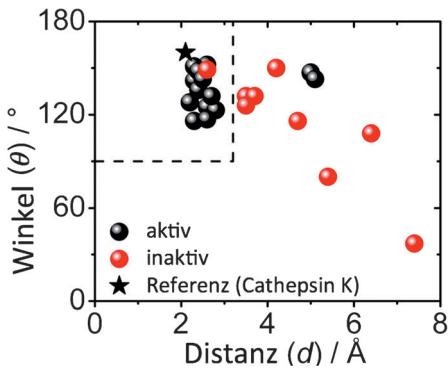


Abbildung 9. Streudiagramm der katalytischen Kontakte, wobei Winkel (θ) gegen Distanzen (d) aufgetragen sind. Der Datensatz besteht aus 23 Kemp-Eliminasen und Cathepsin K. Die Datenpunkte entsprechen den Maxima der jeweiligen Winkel-Distanz-Verteilungen. Die zwei falsch-negativen Datenpunkte gehören zu zwei der drei Kemp-Eliminasen mit der geringsten Aktivität. Die gestrichelte Linie markiert den Umriss der schwachen Wasserstoffbrückenkategorie.^[192] Atompaare mit einem Verteilungsmaximum über diesen Bereich hinaus werden als ungebunden betrachtet. Nachdruck von Kiss et al.^[169]

schwächt die partielle Salzbrücke zwischen dem katalytischen Glu101 und Lys222 (Abbildung 10 a) in doppelter Hinsicht: Das Aspartat konkurriert mit Glu101 um das positiv geladene Lys222 und rekrutiert darüber hinaus weitere Wassermoleküle (Abbildung 10 b), sodass die katalytische Wirkung des Glu101 nicht weiter von der Salzbrücke mit Lys222 geschwächt wird.^[169,185]

4.2.3.2. KE70

Eine Kombination von computerbasierter Sequenzoptimierung und neun Runden der gerichteten Evolution ergab einen > 400 -fachen Anstieg des k_{cat}/K_M -Wertes durch eine bis zu 12-fache Verringerung von K_M und eine 53-fache Erhöhung von k_{cat} .^[186] Der Aktivitätsanstieg wurde einer verstärkten Substratbindung, einer besser abgestimmten Elektrostatik (Abbildung 11a,b) und einer stabilisierten katalyti-

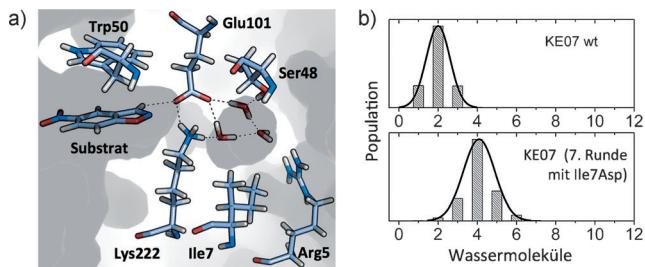


Abbildung 10. a) Das aktive Zentrum des computerbasierten „Wildtyp“-KE07-Designs. Wassermoleküle dringen von rechts ein und koordinieren die Glu-Lys-Salzbrücke. b) Die Ile7Asp-Mutation erhöht die durchschnittliche Wasseroordinationszahl des Glu101 von 2 im „Wildtyp“-KE07-Design auf 4 (hier am Beispiel der 7.-Runde-Variante). Nachdruck von Kiss et al.^[169]

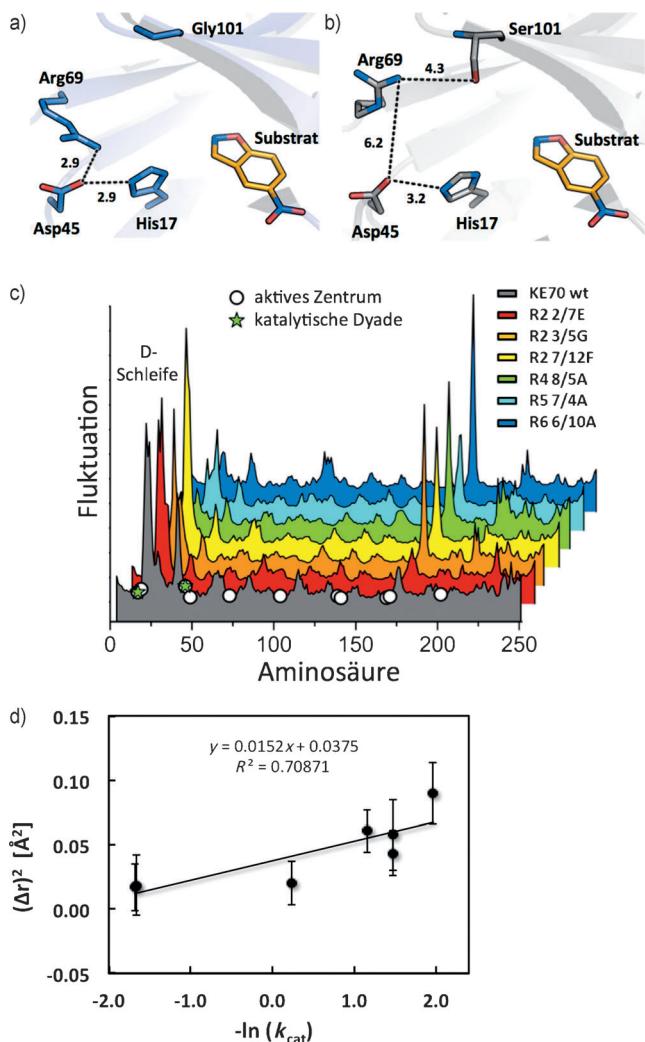


Abbildung 11. a) Kristallstruktur des „Wildtyp“-KE70-Designs. b) Kristallstruktur der 6.-Runde-Variante R6 6/10A. Gly101Ser stabilisiert Arg69 in einer alternativen Ausrichtung, die nicht mit der His17-Asp45-Dyade interferiert. c) Fluktuationsprofile von MD-Simulationen. Die Aminosäuren des aktiven Zentrums (weiße Kreise) und die katalytische Dyade (grüne Sterne) sind gekennzeichnet. Signalspitzen zeigen Regionen mit erhöhter Flexibilität. d) Das Quadrat der Abweichung von der idealen Wasserstoffbrückendistanz (1.8 Å für diesen Kontakt) ist gegen $-\ln(k_{cat})$ aufgetragen. Modifizierter Nachdruck von Khersonsky et al.^[186]

schen Dyade (Abbildung 11 c,d) zugeschrieben. Neun Runden der gerichteten Evolution führten zu einer reduzierten Flexibilität der „D-Schleife“ (Abbildung 11 c) zwischen His17 und Asp45, sodass die katalytische Wasserstoffbrücke dieser beiden Aminosäuren verstärkt wurde (Abbildung 11 d).

Das aktive Zentrum in KE70 basiert auf Theozym II in Abbildung 5b. Das Wechselwirkungspotential des His17-Asp45-Kontaktes kann nahe dem Potentialminimum (mit Distanz r_0) als annähernd harmonisch beschrieben werden. Abweichungen von r_0 bewirken ein Abschwächen der Wechselwirkung, was proportional zu $(\Delta r)^2$ ist. Unter der Annahme eines einfachen Übergangszustandsmodells kann dann anhand der Eyring-Theorie davon ausgegangen werden, dass $\ln(k_{\text{cat}})$ proportional zur Aktivierungsenergie (ΔG^\ddagger) ist. Ein lineares Verhältnis zwischen $(\Delta r)^2$ und $\ln(k_{\text{cat}})$ (Abbildung 11 d) deutet dann darauf hin, dass der Anstieg des k_{cat} -Wertes im Zuge der gerichteten Evolution zum Großteil auf eine Verbesserung des His17-Asp45-Kontakts zurückzuführen ist.

Im Gegensatz zur gerichteten Evolution von KE07, wo positive Mutationen nur in der zweiten und dritten Schale um das aktive Zentrum herum gefunden wurden, konnten diese im Fall der KE70-Optimierung auf die erste Schale erweitert werden.

4.2.3.3. KE59

Die funktionellen Mutationen, die für die relativ hohe Anfangsaktivität dieser Kemp-Eliminase verantwortlich sind, bewirkten ebenfalls, dass KE59 zu einem der thermisch instabilsten „Inside-Out“-Enzyme wurde. Im Gegensatz zu KE07 und KE70 war es hier nötig, zunächst stabilisierende Konsensusmutationen einzuarbeiten. Die resultierenden Varianten wurde daraufhin 16 Runden der gerichteten Evolution unterworfen, was zu einem >2000-fachen Anstieg des k_{cat}/K_M -Wertes führte – hauptsächlich aufgrund einer substantiellen Erhöhung des k_{cat} .^[187] Die kompetenteste Variante zeigte einen K_M -Wert von 37 μM , einen k_{cat}/K_M -Wert von 0.6 \times $10^6 \text{ s}^{-1}\text{M}^{-1}$ und einen $k_{\text{cat}}/k_{\text{uncat}}$ von ca. 10^7 – kinetische Größen, die denen natürlicher Enzyme sehr nahekommen.

KE59 ist darüber hinaus die einzige de novo entwickelte Kemp-Eliminase, die außer 5-NBZ eine Reihe unterschiedlicher Benzisoxazol-Substrate umsetzen kann. Der größte Aktivitätsanstieg wurde für 5,7-Dichlorbenzisoxazol (ein weniger aktiviertes Substrat als 5-NBZ) gemessen. Die drastische Erhöhung der k_{cat} -Werte wurde hauptsächlich einer Verstärkung der katalytischen Base zugeschrieben, da fortgeschrittene Varianten besser in der Lage zu sein scheinen, Wassermoleküle aus dem aktiven Zentrum auszuschließen (Abbildung 12 a). Substituenten an Position 5 und 7 des Benzisoxazol-Substrats sind dazu am besten geeignet (Abbildung 12 b). Das katalytische Glu230 kann – durch solche Wasserkontakte induziert – eine alternative, katalytisch suboptimale Konformation annehmen (Abbildung 12 c).

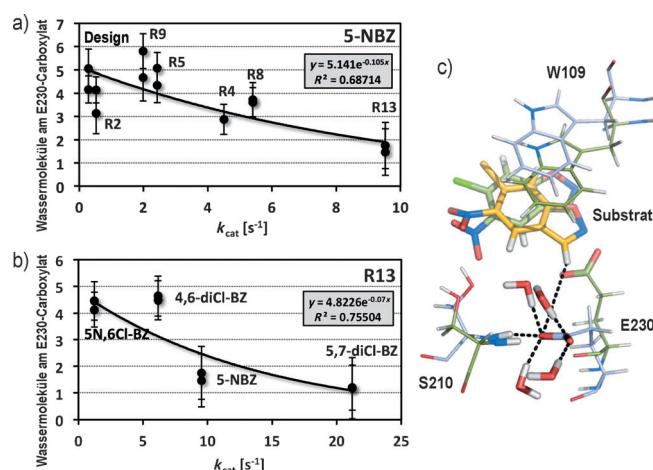


Abbildung 12. Wassermoleküle innerhalb eines 3.2 Å Radius um die Glu230-Carboxylat-Sauerstoffatome. a) Aufgetragen für alle KE59-Varianten mit k_{cat} -Werten für 5-Nitrobenzisoxazol und b) für alle Benzisoxazol-Substrate, für die, in Kombination mit der R13-Variante, ein k_{cat} gemessen werden konnte. Die Fehlerbalken entsprechen der Standardabweichung $+$ / $-$ der MD-basierten Verteilungen. c) Die in den MD-Simulationen vorherrschende Ausrichtung des katalytischen Glu230 am Beispiel der Variante R13 3/11H mit 5-Nitro-6-chlorbenzisoxazol (in Blau mit dem Substrat in Orange) im Vergleich zu der Ausrichtung, die in der Kristallstruktur beobachtet wurde (in Grün). Modifizierter Nachdruck von Khersonsky et al.^[187]

4.3. Diels-Alderasen

Siegel et al. beschrieben das Design und die experimentelle Charakterisierung von De-novo-Enzymen für die Katalyse einer bimolekularen Diels-Alder-Reaktion (Abbildung 13 a), die ein hohes Maß an Stereoselektivität und Substratspezifität aufwies (Abbildung 13 b).^[172] Bisher sind keine natürlichen Enzyme mit einer ähnlichen katalytischen Funktion bekannt. Das Theozym-Motiv war vom aktiven Zentrum eines katalytischen Antikörpers inspiriert, in welchem die katalytische Einheit aus einem Asp, Asn und Tyr besteht.^[44,46] Das Theozym-Motiv der computerbasierten Diels-Alderasen setzt sich aus einer Gln- und zwei Tyr-Aminosäuren zusammen, die für die bevorzugte Stabilisierung des 3R,4S-endo-Übergangszustands positioniert wurden. Das Unterfangen ergab zwei aktive Proteine: DA20, das in eine β -Propellerstruktur eingebaut wurde, und DA42 in einer KSI-Struktur (Abbildung 14).

Computerbasierte Untersuchungen (QM und MD) halfen bei der Erklärung der experimentellen Beobachtungen und waren richtungsgebend bei der Wahl aktivitätssteigernder Mutationen. Ein Beispiel, das dabei heraussteht, ist die Entwicklung der DA_20_10-Variante aus DA_20_00. MD-Simulationen an DA_20_00 ($k_{\text{cat}} = 0.1 \text{ h}^{-1}$) zeigten, dass das katalytische Tyr121 eine unproduktive Ausrichtung einnehmen kann, in welcher es das Peptidcarbonyl der Position 271 kontaktiert (Abbildung 15, rot). Ein erhöhter sterischer Anspruch an der benachbarten Position 272 kann dem entgegenwirken (Abbildung 15 a, blau), sodass die OH-Gruppe des Tyr121 nun dauerhaft in einer katalytisch kompetenten Orientierung zu finden ist, in welcher das Dien-Substrat entsprechend gebunden werden kann. Abbildung 15 b gibt

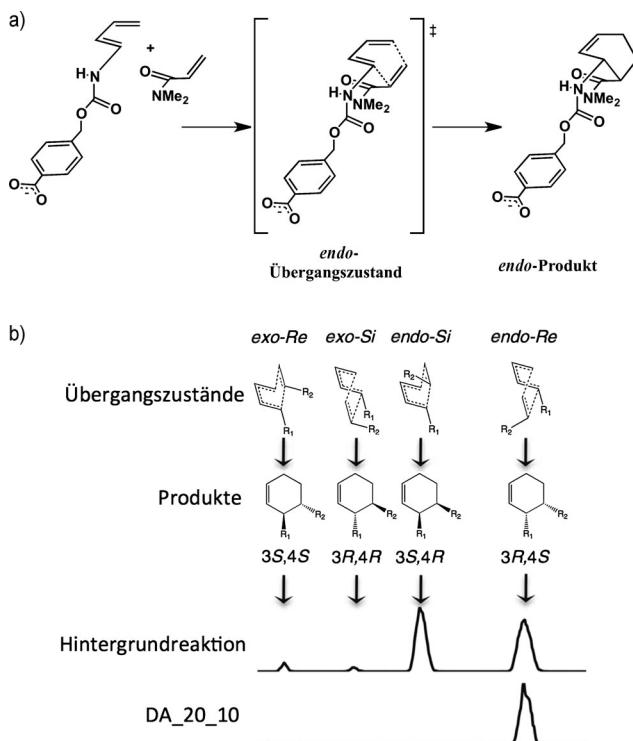


Abbildung 13. Diels-Alder-Reaktion von 4-Carboxybenzyl-trans-1,3-butadien-1-carbamat und *N,N*-Dimethylacrylamid (a), wobei nur das 3*R*,4*S*-endo-Produkt entsteht (b). Teil (b) Nachdruck von Siegel et al.^[172]

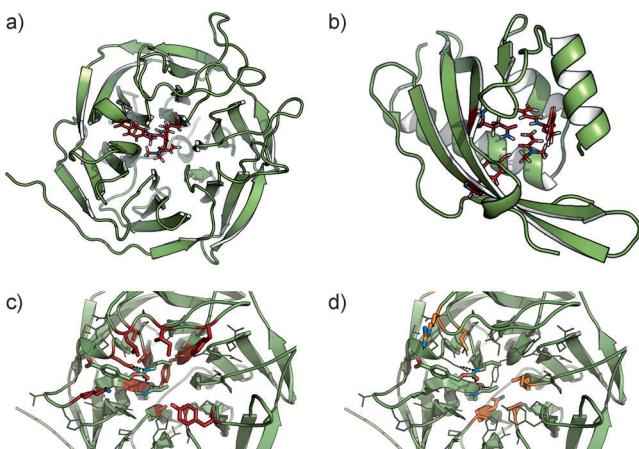


Abbildung 14. Computerbasierte Diels-Alderasen. a) DA_20_10. b) DA_42_04. Gln, Tyr und Substrat im aktiven Zentrum sind in Rot gekennzeichnet. c) Das aktive Zentrum von DA_20_00 und d) DA_20_10. Die jeweiligen Mutationen sind in Rot und Orange hervorgehoben. Teile (b) und (c) sind Nachdrücke von Siegel et al.^[172]

diesen Sachverhalt im Vergleich zur QM-Theozymgeometrie wieder. DA_20_10 wurde mit einer Geschwindigkeitskonstante (k_{cat}) von 2.1 h^{-1} charakterisiert.

Die Kristallstruktur deckt sich mit dem Computerbasierten Modell und die katalytische Effizienz dieser Diels-Alderasen ist vergleichbar mit der von katalytischen Antikörpern. Dabei glänzen die de novo Designs mit Hinblick auf die erreichte Geschwindigkeitserhöhung, stehen den Anti-

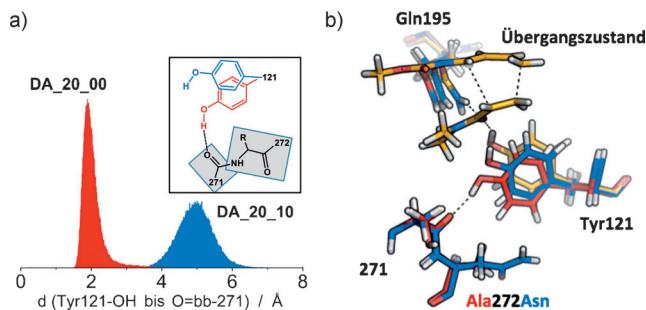


Abbildung 15. a) DA_20_00 zeigt eine geringe Streuung um 2 \AA (Kontakt zwischen Tyr121 und dem Carbonyl an Position 271), während DA_20_10 eine breite Verteilung um 5 \AA aufweist (kein Kontakt). b) Überlagerung des QM-Übergangszustands (orange) mit äquilibrierten MD-Geometrien von DA_20_00 (rot) und DA_20_10 (blau). Nachdruck von Siegel et al.^[172]

körpern jedoch in Sachen Bindungsvermögen der Substrate nach.

4.4. Iterativer Ansatz zum „Inside-Out“-Enzymdesign

Die Arbeitsgruppen von Houk und Mayo erkundeten zusammen eine iterative Variante^[191] zum „Inside-Out“-Designansatz nach Baker/Houk.^[171] Anstelle der Expression und Charakterisierung einer Vielzahl von Proteinen lag der Fokus hier auf einem einzelnen computerbasierten Design (mit dem Vorlagenprotein 1gor;^[193] Abbildung 16 a). Wie auch bei den Röthlisbergerschen Kemp-Eliminasen, wurde hier ein neues aktives Zentrum aufgebaut, das den Übergangszustand der Kemp-Eliminierung mit 5-NBZ (Abbildung 5 a) stabilisiert. Theozym I (Abbildung 5 b) diente dabei als Vorlage und Phoenix^[194] fungierte als das Designprogramm für den Zusammenbau des aktiven Zentrums in silico. Das hierbei verwendete Designprotokoll ist dem der Röthlisberger-Studie sehr ähnlich. Dessen Funktionstüchtigkeit wurde anhand zweier Vorlagenproteine getestet, die zuvor für das Design der Röthlisberger-Eliminasen verwendet wurden. Auf das 1gor-Vorlagenprotein angewendet, war es jedoch nicht möglich, eine aktive Kemp-Eliminase zu produzieren. HG-1 (Abbildung 16 b), das inaktive De-novo-Protein, dass dabei herauskam, unterscheidet sich vom Vorlagenprotein um sieben Mutationen und ist unter den Bedingungen der Aktivitätsmessung stabil. Strukturelle und dynamische Untersuchungen zeigten jedoch einige Problemregionen auf. Das inhärente bewegliche aktive Zentrum des Vorlagenproteins 1gor konnte nicht angemessen stabilisiert werden, um die Theozymgeometrie in HG-1 zu unterstützen (vergleiche Abbildung 18a,b). Zusätzlich dazu stellte sich heraus, dass eine erhebliche Anzahl von Wassermolekülen das aktive Zentrum fluten können, was sowohl die Substratbindungsstärke als auch die Basizität des Glu237 beeinträchtigt (Abbildung 18c). Bemühungen, die darauf hinausgingen, den hydrophoben Charakter des aktiven Zentrums in HG-1 zu erhöhen, waren erfolglos, sodass eine invasivere Strategie verfolgt wurde. Anstelle sich von 1gor abzuwenden und die RCSB nach einem Vorlagenprotein zu durchsuchen, das zur

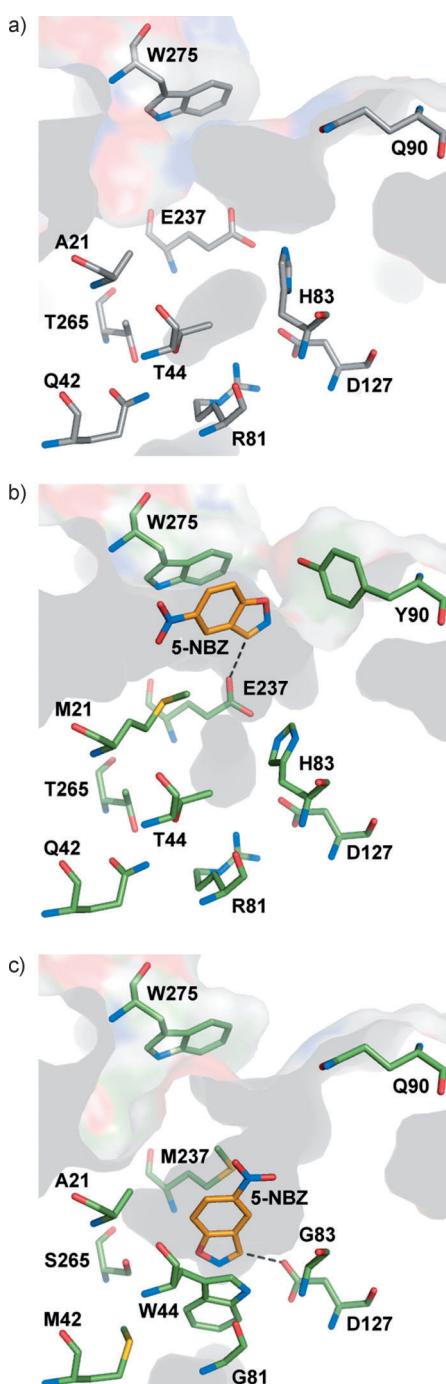


Abbildung 16. Variation des aktiven Zentrums. a) Unmodifiziertes 1gor-Vorlagenprotein, b) HG-1-Design und c) HG-2-Design. Substrat in Orange. Nachdruck von Privett und Kiss et al.^[19]

Einbettung von Theozym I eventuell besser geeignet ist, wurden die Designbemühungen auf eine bereits vorhandene kleine Öffnung im Inneren des Proteins verlagert (sichtbar in der unteren Mitte in Abbildung 16 a,b).

Das daraus resultierende HG-2-Design unterscheidet sich vom 1gor-Vorlagenprotein um 12 Mutationen und macht sich Asp127 als katalytische Base zunutze (Abbildung 16c). Das native katalytische Zentrum wurde um 7 Å vertieft und zum Eingang hin verengt, sodass zum Schluss ein komplett neues

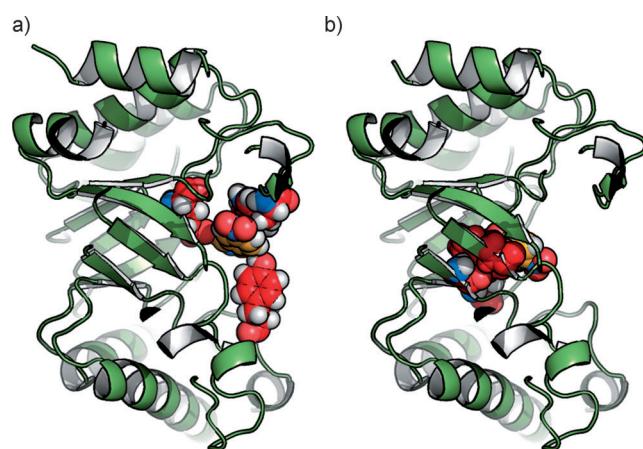


Abbildung 17. Aktives Zentrum, verschoben um 7 Å. a) HG-1. b) HG-2. Ausschnitt mit den Aminosäuren des aktiven Zentrums in Rot und dem Substrat in Orange.

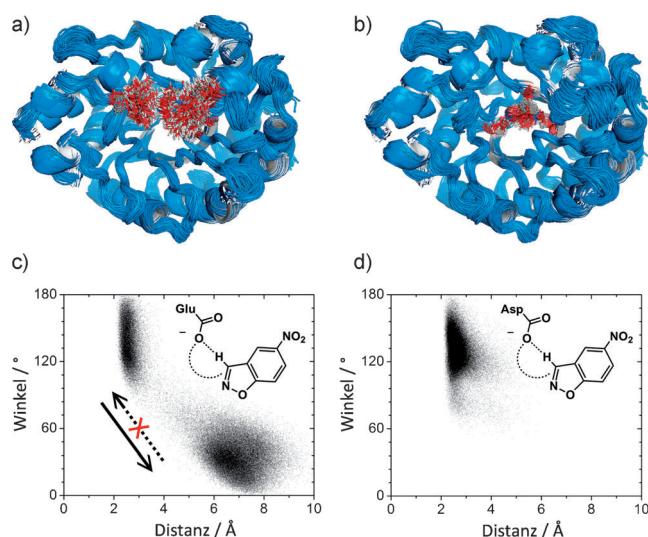


Abbildung 18. Dynamischer Charakter von HG-1 (a) und HG-2 (b). Äquidistante Momentaufnahmen von MD-Trajektorien. Die Beweglichkeit des Rückgrats (blau) ist bei beiden Designs vergleichbar, wogegen sich die aktiven Zentren (rot) drastisch unterscheiden. c) und d) zeigen Streudiagramme der katalytischen Kontakte, wobei in c) das Substratsbindungsproblem von HG-1 deutlich wird. Modifizierter Nachdruck von Privett und Kiss et al.^[19]

aktives Zentrum innerhalb des β -Barrels entstand (Abbildung 17).

MD-Simulationen zeigten, dass die Problemregionen von HG-1 durch den Umbau beseitigt werden konnten und dass HG-2 in der Lage zu sein scheint die Theozymgeometrie zu unterstützen, den Übergangszustand zu binden (Abbildung 18d) und den Zufluss von Lösungsmittelmolekülen zu limitieren. HG-2 wurde daraufhin exprimiert, und die kinetische Charakterisierung ergab einen k_{cat}/K_M -Wert von $123.2 \text{ M}^{-1} \text{s}^{-1}$, was vergleichbar zu dem der aktivsten Röthlisberger-Eliminase ist (KE59, mit $k_{\text{cat}}/K_M = 163 \text{ M}^{-1} \text{s}^{-1}$).

Die Kristallstruktur (Auflösung 1.2 Å) wurde in der Gegenwart eines cokristallisierten Übergangszustandsanalogons (TSA, transition state analogue) gelöst und zeigt das

TSA in zwei unterschiedlichen, jedoch katalytisch kompetenten Orientierungen. Die Struktur validierte zwar den MD-gesteuerten Designansatz, betonte darüber hinaus jedoch auch die Notwendigkeit, während des Designprozesses alternative Substratorientierungen zu berücksichtigen. Weitere Variationen des aktiven Zentrums wurden getestet und ergeben die Ser265Thr-Variante, die einen dreifach erhöhten k_{cat}/K_M -Wert aufweist.

4.5. Strukturvorhersage und -design durch Crowd-Sourcing

Moderne Computeralgorithmen sind mittlerweile äußerst wirksam bei der Beschreibung physikalischer Größen, die molekularen Wechselwirkungen zugrunde liegen. Ein effektives Abtasten des Konformationsraumes ist jedoch nach wie vor eine der schwierigsten Herausforderungen. Die Energielandschaften von Biomolekülen sind so enorm, dass deren Navigation einen der Hauptengpässe in der Proteininfaltung, der Strukturvorhersage und dem Proteindesign darstellt. Verschiedene Ansätze – etwa durch Strukturvorhersage (z.B. Rosetta) oder Simulation (z.B. Markov-State-Modelle in Kombination mit Folding@home), aber auch hochspezialisierte Hardware (z.B. Anton) – wurden jüngst entwickelt, die sich mit diesem numerischen Problem befassen. Strukturvorhersage durch Crowd-Sourcing ist eine der neuesten Formen und hat sich bisher als erstaunlich effektives Addendum zu statistischen und deterministischen Suchalgorithmen erwiesen.

Cooper et al. stellten Foldit vor, eine graphische Benutzeroberfläche zu einigen der Rosetta-Funktionalitäten, welche darüber hinaus als Online-Mehrspielerplattform dienen kann.^[195] Das Konzept hinter der Anwerbung des „homo ludens“, des spielenden Menschen, zur Lösung wissenschaftlicher Rätsel geht auf Beobachtungen an menschenbasiertem Computing zurück. Hier werden bestimmte Aufgaben innerhalb eines Protokolls, wie etwa das Erkennen von Formen, von Menschen wesentlich schneller und effizienter durchgeführt als von Rechnern. Die Foldit-Studie zeigte, dass fundamentale Fähigkeiten der Mitspieler, wie etwa räumliche Erkennung, Intuition und Entscheidungsfindung der stochastischen Komponente, den modernsten Konformations-suchalgorithmen überlegen sind. Die Studie zeigt weiter, dass solche elementaren menschlichen Fertigkeiten zur Problemlösung in der Strukturvorhersage eingesetzt werden können.

Der gemeinschaftliche Charakter des Spiels erlaubt den Teilnehmern, sich in Gruppen zu organisieren und gegen andere Gruppen anzutreten, aber auch ihre Erfahrungen und Strategien miteinander zu teilen und in der Form von „Rezepten“ weiterzuentwickeln. Erfolgreiche „Rezepte“ werden

häufiger verwendet und optimiert als weniger erfolgreiche, sodass interessante Entwicklungsmäßige Prozesse entstanden, die durch das gemeinschaftliche Wirken wissenschaftlicher Laien zu neuen und verbesserten Vorhersagealgorithmen führten. Beispielsweise hatten die beiden beliebtesten „Rezepte“ eine erstaunliche Ähnlichkeit mit einem verbesserten Strukturvorhersagealgorithmus, der zu der Zeit noch unveröffentlicht war.^[196]

Zwei jüngste Berichte veranschaulichen die Anwendbarkeit von Crowd-Sourcing und Foldit über die Demonstrationsphase hinaus. Khatib et al. generierten Modelle für Molecular Replacement in der Kristallographie, anhand welcher die seit langem ungeklärte Struktur der M-PMV-Retrovirusprotease endlich gelöst werden konnte.^[197] Eiben et al. erzielten durch ein umfangreiches Loop-Design eine insgesamt 18-fach verbesserte Substratbindung für eine der beiden De-novo-Diels-Alderases (Abbildung 19).^[198]

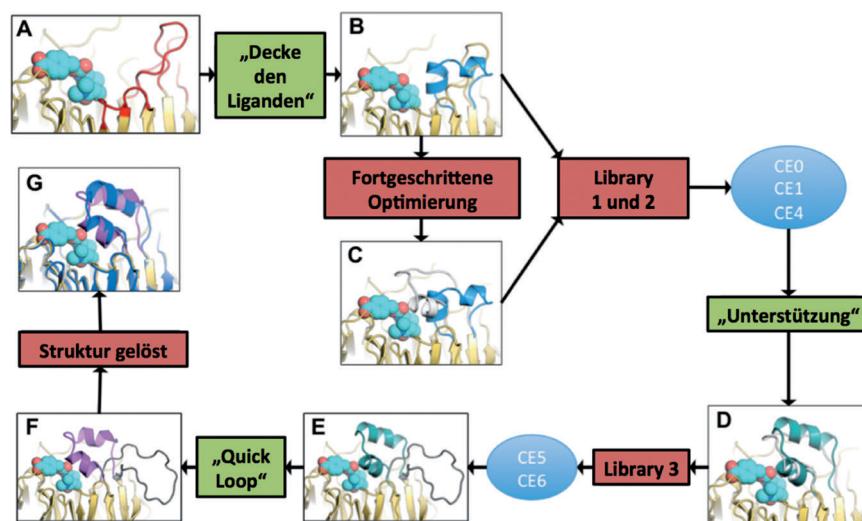


Abbildung 19. Ablauf des Foldit-unterstützten Loop-Designs von DA_20_10. Aufgaben, die den Online-Spielern gestellt wurden, sind in den grünen Feldern dargestellt. Die besten Designs in der Rangliste wurden exprimiert und getestet. Die besten davon wurden dann in Form einer weiteren Aufgabe und die Spieler weitergeleitet. Nachdruck von Eiben et al.^[198]

Während die evolutionsmäßige Dynamik von Crowd-Sourcing und die Anwendbarkeit nicht-wissenschaftlicher Gedankengänge auf wissenschaftliche Fragestellungen in sich faszinierende Themen sind, so hängt die Qualität der resultierenden Vorhersagen stark davon ab, wie geschickt die wissenschaftliche Zielsetzung in eine allgemein verständliche Sprache für die Online-Spieler gelegt werden kann.^[197] Nichtsdestotrotz kann in der nahen Zukunft in diesem Bereich mit der einen oder anderen äußerst interessanten Entwicklung gerechnet werden.

5. Herausforderungen beim Enzymdesign

Die computerbasierte Chemie und Biologie haben in den letzten Jahren durch die Entwicklung neuer Algorithmen und

Hardware – aber auch durch einen Anstieg und eine Verfestigung interdisziplinärer Kollaborationen – erhebliche Fortschritte erfahren. Forschern in diesen Gebieten ist es nun möglich, auf interessante Fragestellungen einzugehen, für die es noch vor einem halben Jahrzehnt nicht das nötige Werkzeug gab. Die routinemäßige Anwendung von QM, MM, MD und verschiedenen Kombinationen davon, die ständig steigende Leistungsfähigkeit von Supercomputern und von verteiltem oder dezentralisiertem Rechnen wie Folding@home und Rosetta@home, hochspezialisierte Computer wie Anton, verbesserte Strukturvorhersagealgorithmen und letztlich Crowd-Sourcing wurden dazu verwendet, aufregende neue Wege im Bereich der Proteinfaltung, Strukturvorhersage und des Protein-Engineerings zu beschreiten. Erhebliche Fortschritte in experimentellen Methoden und die einhergehenden reduzierten Kosten und Zeitinvestitionen haben es möglich gemacht, dass Vorhersagen weitaus routinemäßiger getestet werden können. Bisher hat uns die Arbeit auf diesen Gebieten bloß flüchtige Einblicke auf zukünftige Entwicklungen gegeben. Und während wir uns darum bemühen, die Methoden und Protokolle in eine routinemäßig einsetzbare Technologie zu entwickeln, so befindet sich das computerbasierte Enzymdesign in einer frühen Entwicklungsphase, und erhebliche Herausforderungen – sowohl rechnerische als auch experimentelle – liegen noch vor uns.

Die Werte in Tabelle 2 zeigen das Ausmaß der Überlegenheit natürlicher Enzyme gegenüber künstlichen Kataly-

5.1. Rechnerische Herausforderungen

Die kinetischen Parameter in Tabelle 2 zeigen, dass artifizielle Katalysatoren, die von computerbasiertem Design und immunologischer Produktion stammen, natürlichen Enzymen vor allem beim Bindes des Übergangszustands nachstehen. Dies kann verschiedenen Gründen zugeschrieben werden.

Eine Auslegung des Sachverhalts ist, dass die Ausrichtung der katalytischen Aminosäuren – sowohl in den De-novo-Designs als auch in den Antikörpern – bei weitem nicht die Präzision erreicht, für welche die aktiven Zentren natürlicher Enzyme bekannt sind.^[168, 199, 200] Abbildung 3 betont dies anhand vierer Kemp-Eliminasen und deutet darüber hinaus auf ein gewisses Maß der Ungenauigkeit im gegenwärtigen Designprotokoll hin, da die QM-Theozymgeometrien nicht originalgetreu in den fertigen Designs wiederzufinden sind.

Die enormen Geschwindigkeitserhöhungen natürlicher Enzyme sind weiterhin einer Minimierung von Konformationsbewegungen zugeschrieben worden.^[201, 202] Ein Vergleich mit computerbasierten Designs zeigt, dass sogar die aktivsten De-novo-Enzyme ein hohes Maß an Konformationsmobilität aufweisen, was am Beispiel von Abbildung 7 a,b ersichtlich ist. Gerichtete Evolution kann diesen Bestand zu einem begrenzten Ausmaß korrigieren und die Präorganisation des aktiven Zentrums erhöhen (Abbildung 11 c,d), aber auch dies ist auf bereits funktionelle Proteine beschränkt. Diejenigen,

denen es solch anfänglicher Aktivität fehlt, bevölkern unproduktive Konfigurationen des aktiven Zentrums (Abbildung 7 c und Abbildung 18).^[169] Bemühungen, die k_{cat} -Werte von De-novo-Enzymen zu verbessern, konzentrieren sich daher auf rigorosere Implementierungen von QM-Theozymgeometrien in die Vorlagenproteine, aber auch auf eine routinemäßige Physik-basierte Begutachtung der aktiven Zentren und ihren dynamischen Eigenschaften.

Natürlich vorkommende Enzyme sind darüber hinaus auf eine feinabgestimmte Regulierung von Wassermolekülen im aktiven Zentrum angewiesen. Sie erstellen fein abgestimmte Mikroumgebungen, in denen die $\text{p}K_s$ -Werte katalytischer Aminosäuren stark verschoben werden und aktivierte Wassermoleküle die existierende

katalytische Maschinerie erweitern können. Ein solches Ausmaß an Kontrolle ist zwar erstrebenswert, aber bislang außer Reichweite. Gerichtete Evolution kann auch hier behilflich sein – wie am Beispiel der KE59-Varianten ersichtlich ist (Abbildung 12) –, die resultierenden Mikroumgebungen sind jedoch weit davon entfernt, als „fein abgestimmt“ bezeichnet werden zu können. Wassermoleküle sind allerdings auch dazu in der Lage, suboptimale aktive Zentren zu

Tabelle 2: Kinetische Parameter computerbasierter Designs im Vergleich zu katalytischen Antikörpern und natürlichen Enzymen.

	$k_{\text{cat}}^{[a]}$	$K_M^{[b]}$	$k_{\text{cat}}/K_M^{[c]}$	$k_{\text{cat}}/k_{\text{uncat}}$	$[k_{\text{cat}}/K_M]/k_{\text{uncat}}$
nat. Enzyme	$\emptyset 10^5$	$\emptyset 10^{-4}$	10^6-10^9	10^6-10^{17}	10^8-10^{29}
kat. Antikörper	$10^{-2}-1$	$\emptyset 10^{-4}$	10^2-10^4	10^3-10^6	10^5-10^9
Kemp Elim.					
kat. Antikörper	$10^{-1}-1$	$10^{-3}-10^{-4}$	10^2-10^3	10^3-10^6	10^7-10^9
comp. Designs	$10^{-2}-1$	$\emptyset 10^{-3}$	$10-10^2$	10^3-10^6	10^7-10^9
evol. Designs	1–20	$10^{-3}-10^{-5}$	10^3-10^6	10^6-10^7	10^7-10^{11}
Retro-Aldol					
kat. Antikörper	$10^{-3}-10^{-1}$	$10^{-4}-10^{-5}$	$10-10^3$	10^5-10^6	10^7-10^9
comp. Designs	$10^{-2}-10^{-1}$	$\emptyset 10^{-4}$	$10^{-2}-10^{-1}$	10^3-10^4	10^6-10^7
Diels-Alder^[d]					
kat. Antikörper	$\emptyset 10^{-5}$	$\emptyset 10^{-3}$	$\emptyset 10$	$\emptyset 10^3$	10^9
comp. Designs	$10^{-5}-10^{-4}$	$10^{-1}-10^{-4}$	$1-10^2$	10^3-10^4	10^7-10^{11}

[a] In der Einheit s^{-1} . [b] In der Einheit M . [c] In der Einheit $\text{M}^{-1}\text{s}^{-1}$. [d] $k_{\text{cat}}/(K_{\text{M-diene}} \times K_{\text{M-dienophile}})$ anstelle von k_{cat}/K_M in der Einheit $\text{s}^{-1}\text{M}^{-1}\text{M}^{-1}$ und $(k_{\text{cat}}/(K_{\text{M-diene}} \times K_{\text{M-dienophile}}))/k_{\text{uncat}}$ anstelle von $(k_{\text{cat}}/K_M)/k_{\text{uncat}}$ in der Einheit $\text{M}^{-1}\text{M}^{-1}$.

satoren, was sowohl das Binden von Substraten (K_M) als auch die Geschwindigkeitserhöhung (k_{cat}) betrifft. Anhand von gerichteter Evolution können die Michaelis-Konstanten von De-novo-Designs soweit erhöht werden, dass deren K_M -Werte mit denen natürlicher Enzyme vergleichbar sind. Dennoch können sogar die aktivsten der gerichteten Evolutionsvarianten in Bezug auf k_{cat} nur mit den langsamsten natürlich vorkommenden Enzymen konkurrieren.

„retten“. Herschlag und Mitarbeiter zeigten dies für eine Ketosteroid-Isomerase (KSI), bei der die Mutation des katalytischen Tyr16 zu einem Phe den k_{cat} 20000-fach vermindert, während eine Tyr16Ser-Mutation den k_{cat} nur etwa 200-fach herabsetzt.^[199] Kristallographische und ¹⁹F-NMR-spektroskopische Untersuchungen zeigten, dass Wassermoleküle in der Lage sind, die katalytische Funktion zum Teil wiederherzustellen, indem sie die fehlende Tyr-Seitenkette ersetzen. In einer umgekehrten, aber ähnlichen Rolle ermöglichen drei Wassermoleküle die Funktion des aktiven Zentrums im Antikörpers 13G5, wobei ein vierter den k_{cat} -Wert um einen weiteren Faktor 10 anhebt.^[203] Die verbesserten k_{cat} -Werte der KE07-Varianten sind in ähnlicher Weise auf eine erhöhte Präsenz von Wassermolekülen in direktem Kontakt mit katalytischen Aminosäuren zurückzuführen (Abbildung 10).^[169,185]

5.2. Experimentelle Herausforderungen

Das „Inside-Out“-Design von Enzymen ist in experimenteller Hinsicht mit ähnlichen Herausforderungen konfrontiert wie auch gerichtete Evolutionsversuche. Vor allem die Wahl einer effizienten Detektionsmethode kann dabei zu einem ausschlaggebenden Faktor werden, da anfänglich aktive De-novo-Enzyme die gegebene Reaktion meist nur geringfügig beschleunigen. Eine robuste, aber auch empfindliche Methode muss daher in der Lage sein, einstellige Signalansteige im Vergleich zum Hintergrund zu erfassen.

Darüber hinaus kann es häufig dazu kommen, dass eine neue Proteinsequenz niedrige Expressionszahlen aufweist oder unlöslich ist. Dies hat damit zu tun, dass die Sequenzen von De-novo-Enzymen sich oft von denen der Vorlagenproteine um über fünfzehn Mutationen unterscheiden. Auch wenn die Computermodelle dahingehend optimiert werden, dass negative Effekte möglichst gering gehalten werden, kann davon ausgegangen werden, dass die meisten dieser Variationen die Stabilität des Proteins beeinträchtigen. Ähnlich zu den Erfahrungen in der gerichteten Evolution kann die Summe der Mutationen rasch zu einem Problem für die strukturelle Integrität des Proteins werden. Stabile, thermophile Proteine sind daher gute Ausgangspunkte für schwere Mutationslasten. Als potentieller Lösungsansatz kann die Thermostabilität und somit die Halbwertszeit computerbasierter Proteine durch funktionsneutrale Mutationen erhöht werden.^[204] Einige wenige Variationen (drei in diesem Fall) können zu einer Erhöhung der Schmelztemperatur um 10 °C führen, ohne dabei die katalytische Effizienz zu vermindern. Die Wahl der stabilisierenden Aminosäuren kann dabei von eng verwandten Proteinsequenzen beeinflusst werden (Konsensusmutation).^[205]

Stark gefallene DNA-Synthesekosten (von rund \$50/Aminosäure im Jahr 2000 auf \$1/Aminosäure in 2012) haben es erheblich leichter gemacht, neue Proteinsequenzen zu testen.^[206,207] Die DNA-Sequenz für ein typisches Enzym kann dann für einige hundert Dollar bestellt werden, sodass ein gängiges Biochemielabor eine Anzahl verschiedener neuer Proteinsequenzen mit einem relativ niedrigen Budget untersuchen kann. „Gene-Assembly“-Ansätze und „On-

Chip“-DNA-Synthese^[208] ermöglichen darüber hinaus den Aufbau von Fokusbibliotheken.

Die Toxizitätsprofile und die synthetische Verfügbarkeit von Substraten können zu zusätzliche Herausforderungen führen. Zwei jüngste Berichte, die sich die Sequestrierung von Organophosphat-Nervengiften zum Ziel setzten, untermauern diesen Punkt. Khare et al. verwendeten das „Inside-Out“-Protokoll, um eine Deaminase für die Hydrolyse eines Nervengift-Analogons umzubauen.^[209] Hemmert et al. verwendeten einen strukturbasierten Ansatz zur Herstellung von Varianten der Carboxylesterase 1 mit erhöhten Aktivitäten für Sarin, Soman und Cyclosarin, wobei die Aktivitätsmessung in einem Hochsicherheitslabor des US-Militärs durchgeführt werden musste.^[210]

6. Zusammenfassung und Ausblick

Computerbasierte Methoden haben einen Punkt erreicht, an dem es nun möglich ist, funktionelle Proteine für die Katalyse nichtnatürlicher Reaktionen herzustellen, für die es in der Natur kein Gegenstück gibt. Obwohl die Beschleunigungsrate solcher „Inside-Out“-Enzyme bislang relativ bescheiden sind, so zeigen sie bereits Anzeichen katalytischer Eigenschaften, die typisch für natürlicher Enzyme sind. Dazu gehören verschobene pK_s -Werte, Substratspezifität und Stereoselektivität. Diese Funktionen können beträchtlich erhöht werden, wenn computerbasiertes Design mit gerichteter Evolution verbunden wird. Letztere hat in jüngster Zeit im Rahmen industrieller Anwendungen einiges an Schwung gewonnen, wie am Beispiel der Entwicklung einer Transaminase für die Synthese des Diabetes-Medikaments Sitagliptin ersichtlich wird.^[106] Gerichtete Evolutionsversuche benötigen jedoch in der Regel Proteine, welche die gewünschte Funktionalität bereits zum Teil besitzen. Wenn solche Ansatzzpunkte in der Natur nicht vorkommen, können diese durch computerbasiertes Design generiert werden.

Bisher wurden mithilfe des „Inside-Out“-Protokolls Biokatalysatoren für relativ einfache uni- und bimolekulare Einzelschrittreaktionen hergestellt, während die Konstruktion aufwendigerer aktiver Zentren für mehrstufige Reaktionen in Arbeit sind.^[211] Der „Inside-Out“-Ansatz eignet sich jedoch auch zur Abwandlung bereits existierender Funktionen, wie Gordon et al. anhand der modifizierten Substratspezifität einer α -Gliadin-Endopeptidase zeigten.^[212] Mit der voranschreitenden Leistungsfähigkeit rechnerischer Methoden und der zugrundeliegenden Hardware werden gegenwärtig neue Algorithmen entwickelt, die es ermöglichen werden, sowohl komplexere aktive Zentren in Proteinvorlagen einzubauen als auch deren katalytische Kompetenz durch bedienungsfreundliche und akkurate Physik-basierte Simulationen zu testen und zu reparieren, bevor diese experimentell aufgearbeitet werden.

Die physikalischen Gesetze, die beim Design von Enzymen zum Einsatz kommen, sind wohlbekannt und innerhalb der einzelnen Designkomponenten mit hoher Genauigkeit beschrieben. Ihre Kombination führt jedoch nicht nur zu einer enormen Vielfalt möglicher Lösungen, sondern auch zu Inkonsistenzen zwischen einzelnen Komponenten. Die

Hauptherausforderung beim Design eines hochkompetenten Enzyms besteht dann in der exakten und schnellen Berechnung aller Lösungen. Dieser Befund hat eine gewisse Ähnlichkeit mit einem berühmten Zitat von Dirac: „*Die grundlegenden physikalischen Gesetze für den Großteil der Physik und die gesamte Chemie sind vollständig bekannt, und die Schwierigkeit liegt lediglich darin, dass die exakte Anwendung dieser Gesetze zu Gleichungen führt, die bei weitem zu komplex sind, als dass sie gelöst werden können.*“^[213]

Fast ein Jahrhundert später sind wir mehr denn je davon überzeugt, dass wir diese physikalischen Gesetze verstehen, einschließlich derjenigen, die enzymkatalytische Vorgänge beschreiben. Darüber hinaus ist es mittlerweile möglich, die Dirac-Gleichung mithilfe von Rechnern zu einer beliebigen Genauigkeit zu lösen – zumindest für kleine Systeme. Enzyme sind jedoch nicht nur groß, von Wasser umgeben und zur Adaption verschiedener Konformationen in der Lage, sondern stellen ihren jeweiligen Substraten eine präzise abgestimmte Mikroumgebung zur Verfügung, in denen Reaktionen – oft bimolekular, über mehrere Schritte hinweg und/oder Cofaktor-abhängig – effizient und selektiv beschleunigt werden können. Diracs Pessimismus stammt teilweise daher, dass Computer zu seiner Zeit noch unbekannt waren. Das heutige rasante Tempo der Hardware- und Algorithmenentwicklung gibt uns jedoch Grund zum Optimismus, was die Zukunft des computerbasierten Enzymdesigns betrifft. Der Idealfall ist dann einer, in welchem alle Merkmale effizienter Enzyme routinemäßig, hochgenau und schnell berechnet werden können. Dies ist zwar zurzeit noch nicht möglich, die gleichzeitige Umsetzung aller Aspekte ist jedoch absehbar, und es ist durchaus denkbar, dass dies in der nicht allzu fernen Zukunft zu einer alltäglichen Übung wird.

Wir danken der Defense Advanced Research Projects Agency (DARPA), dem National Institute of General Medical Sciences (NIGMS), den National Institutes of Health (NIH), dem Howard Hughes Medical Institute (HHMI) und dem Lawrence Scholars Programm für finanzielle Unterstützung.

Eingegangen am 25. Mai 2012
Online veröffentlicht am 25. März 2013

- [1] W. Kühne, *FEBS Lett.* **1976**, *62*, E4–E7.
- [2] L. Pauling, *Chem. Eng. News* **1946**, *24*(10), 1375–1377.
- [3] L. Pauling, *General Chemistry: an Introduction to Descriptive Chemistry and Modern Chemical Theory*, Freeman, San Francisco, **1947**.
- [4] C. C. Blake, D. F. Koenig, G. A. Mair, A. C. North, D. C. Phillips, V. R. Sarma, *Nature* **1965**, *206*, 757–761.
- [5] A. Warshel, *J. Biol. Chem.* **1998**, *273*, 27035–27038.
- [6] A. Radzicka, R. Wolfenden, *Science* **1995**, *267*, 90–93.
- [7] X. Zhang, K. N. Houk, *Acc. Chem. Res.* **2005**, *38*, 379–385.
- [8] D. R. Edwards, D. C. Lohman, R. Wolfenden, *J. Am. Chem. Soc.* **2012**, *134*, 525–531.
- [9] A. Radzicka, R. Wolfenden, *J. Am. Chem. Soc.* **1996**, *118*, 6105–6109.
- [10] R. Wolfenden, X. Lu, G. Young, *J. Am. Chem. Soc.* **1998**, *120*, 6814–6815.
- [11] E. A. Taylor, D. R. Palmer, J. A. Gerlt, *J. Am. Chem. Soc.* **2001**, *123*, 5824–5825.
- [12] R. A. R. Bryant, D. E. Hansen, *J. Am. Chem. Soc.* **1996**, *118*, 5498–5499.
- [13] K. Houk, A. G. Leach, S. P. Kim, X. Zhang, *Angew. Chem.* **2003**, *115*, 5020–5046; *Angew. Chem. Int. Ed.* **2003**, *42*, 4872–4897.
- [14] J. D. Stewart, *Adv. Appl. Microbiol.* **2006**, *59*, 31–52.
- [15] T. Furuya, K. Kino, *Appl. Microbiol. Biotechnol.* **2010**, *86*, 991–1002.
- [16] J. A. Gerlt, P. C. Babbitt, *Curr. Opin. Chem. Biol.* **2009**, *13*, 10–18.
- [17] *Directed Enzyme Evolution: Screening and Selection Methods* (Hrsg.: F. H. Arnold, G. Georgiou), Humana, New York, **2003**.
- [18] F. H. Arnold, G. Georgiou, P. C. Cirino, K. M. Mayer, D. Umeno, *Directed Evolution Library Creation*, Humana, New York, **2003**.
- [19] J. D. Bloom, M. M. Meyer, P. Meinhold, C. R. Otey, D. MacMillan, F. H. Arnold, *Curr. Opin. Struct. Biol.* **2005**, *15*, 447–452.
- [20] C. Jäckel, P. Kast, D. Hilvert, *Annu. Rev. Biophys.* **2008**, *37*, 153–173.
- [21] S. Bershtain, D. S. Tawfik, *Curr. Opin. Chem. Biol.* **2008**, *12*, 151–158.
- [22] D. Hilvert, *Annu. Rev. Biochem.* **2000**, *69*, 751–793.
- [23] P. Wentworth, Jr., *Science* **2002**, *296*, 2247–2249.
- [24] Y. Xu, N. Yamamoto, K. D. Janda, *Bioorg. Med. Chem.* **2004**, *12*, 5247–5268.
- [25] S. C. Pan, B. List, *Ernst Schering Found Symp. Proc.* **2007**, *1*–43.
- [26] T. Drepper, T. Eggert, W. Hummel, C. Leggewie, M. Pohl, F. Rosenau, S. Wilhelm, K.-E. Jaeger, *Biotechnol. J.* **2006**, *1*, 777–786.
- [27] V. Nanda, R. L. Koder, *Nat. Chem.* **2010**, *2*, 15–24.
- [28] A. Tramontano, K. D. Janda, R. A. Lerner, *Science* **1986**, *234*, 1566–1570.
- [29] S. J. Pollack, J. W. Jacobs, P. G. Schultz, *Science* **1986**, *234*, 1570–1573.
- [30] P. Wentworth, Jr., K. D. Janda, *Curr. Opin. Chem. Biol.* **1998**, *2*, 138–144.
- [31] L. Pauling, *Nature* **1948**, *161*, 707–709.
- [32] W. P. Jencks, *Catalysis in Chemistry and Enzymology*, McGraw-Hill, New York, **1969**.
- [33] M. Arnó, L. R. Domingo, *Int. J. Quantum Chem.* **2001**, *83*, 338–347.
- [34] B. G. Miller, R. Wolfenden, *Annu. Rev. Biochem.* **2002**, *71*, 847–885.
- [35] J. D. Stewart, S. J. Benkovic, *Chem. Soc. Rev.* **1993**, *22*, 213.
- [36] 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.
- [37] A. Tramontano, B. Ivanov, G. Gololobov, S. Paul, *Appl. Biochem. Biotechnol.* **2000**, *83*, 233–242; discussion A. Tramontano, B. Ivanov, G. Gololobov, S. Paul, *Appl. Biochem. Biotechnol.* **2000**, *83*, 242–243.
- [38] X. Wu, T. Zhou, J. Zhu, B. Zhang, I. Georgiev, C. Wang, X. Chen, N. S. Longo, M. Louder, K. McKee, S. O'Dell, S. Perfetto, S. D. Schmidt, W. Shi, L. Wu, Y. Yang, *Science* **2011**, *333*, 1593–1602.
- [39] J. S. McLellan, M. Pancera, C. Carrico, J. Gorman, J.-P. Julien, R. Khayat, R. Louder, R. Pejchal, M. Sastry, K. Dai, et al., *Nature* **2011**, *480*, 336–343.
- [40] T. Wojcik, K. Kiec-Kononowicz, *Curr. Med. Chem.* **2008**, *15*, 1606–1615.
- [41] D. Hilvert, K. W. Hill, K. D. Nared, M. T. M. Auditor, *J. Am. Chem. Soc.* **1989**, *111*, 9261–9262.
- [42] J. Xu, Q. Deng, J. Chen, K. N. Houk, J. Bartek, D. Hilvert, I. A. Wilson, *Science* **1999**, *286*, 2345–2348.

- [43] J. Chen, Q. Deng, R. Wang, K. Houk, D. Hilvert, *ChemBioChem* **2000**, *1*, 255–261.
- [44] V. E. Gouverneur, K. N. Houk, B. de Pascual-Teresa, B. Beno, K. D. Janda, R. A. Lerner, *Science* **1993**, *262*, 204–208.
- [45] A. Heine, E. A. Stura, J. T. Yli-Kauhaluoma, C. Gao, Q. Deng, B. R. Beno, K. N. Houk, K. D. Janda, I. A. Wilson, *Science* **1998**, *279*, 1934–1940.
- [46] C. E. Cannizzaro, J. A. Ashley, K. D. Janda, K. N. Houk, *J. Am. Chem. Soc.* **2003**, *125*, 2489–2506.
- [47] N. Bensel, N. Bahr, M. Reymond, C. Schenkels, J.-L. Reymond, *Helv. Chim. Acta* **1999**, *82*, 44–52.
- [48] M. Hugot, N. Bensel, M. Vogel, M. T. Reymond, B. Stadler, J.-L. Reymond, U. Baumann, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9674–9678.
- [49] A. G. Leach, K. N. Houk, J.-L. Reymond, *J. Org. Chem.* **2004**, *69*, 3683–3692.
- [50] S. P. Kim, A. G. Leach, K. N. Houk, *J. Org. Chem.* **2002**, *67*, 4250–4260.
- [51] S. N. Thorn, R. G. Daniels, M. T. Auditor, D. Hilvert, *Nature* **1995**, *373*, 228–230.
- [52] J. Na, K. N. Houk, D. Hilvert, *J. Am. Chem. Soc.* **1996**, *118*, 6462–6471.
- [53] Y. Hu, K. N. Houk, K. Kikuchi, K. Hotta, D. Hilvert, *J. Am. Chem. Soc.* **2004**, *126*, 8197–8205.
- [54] M. L. Casey, D. S. Kemp, K. G. Paul, D. D. Cox, *J. Org. Chem.* **1973**, *38*, 2294–2301.
- [55] D. S. Kemp, M. L. Casey, *J. Am. Chem. Soc.* **1973**, *95*, 6670–6680.
- [56] Y. Li, M. N. Paddon-Row, K. N. Houk, *J. Am. Chem. Soc.* **1988**, *110*, 3684–3686.
- [57] J. Wagner, R. A. Lerner, C. F. Barbas, *Science* **1995**, *270*, 1797–1800.
- [58] C. F. Barbas III, A. Heine, G. Zhong, T. Hoffmann, S. Gramatikova, R. Björnstedt, B. List, J. Anderson, E. A. Stura, I. A. Wilson, R. A. Lerner, *Science* **1997**, *278*, 2085–2092.
- [59] T. Hoffmann, G. Zhong, B. List, D. Shabat, J. Anderson, S. Gramatikova, R. A. Lerner, C. F. Barbas, *J. Am. Chem. Soc.* **1998**, *120*, 2768–2779.
- [60] B. List, C. F. Barbas, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 15351–15355.
- [61] G. Zhong, R. A. Lerner, C. F. Barbas III, *Angew. Chem.* **1999**, *111*, 3957–3960; *Angew. Chem. Int. Ed.* **1999**, *38*, 3738–3741.
- [62] M. Arnó, L. R. Domingo, *Org. Biomol. Chem.* **2003**, *1*, 637–643.
- [63] D. S. Kemp, K. G. Paul, *J. Am. Chem. Soc.* **1975**, *97*, 7305–7312.
- [64] D. S. Kemp, D. D. Cox, K. G. Paul, *J. Am. Chem. Soc.* **1975**, *97*, 7312–7318.
- [65] H. Zipse, G. Apaydin, K. N. Houk, *J. Am. Chem. Soc.* **1995**, *117*, 8608–8617.
- [66] J. Gao, *J. Am. Chem. Soc.* **1995**, *117*, 8600–8607.
- [67] C. Lewis, T. Kramer, S. Robinson, D. Hilvert, *Science* **1991**, *253*, 1019–1022.
- [68] K. Hotta, H. Lange, D. J. Tantillo, K. N. Houk, D. Hilvert, I. A. Wilson, *J. Mol. Biol.* **2000**, *302*, 1213–1225.
- [69] G. Ujaque, D. J. Tantillo, Y. Hu, K. Houk, K. Hotta, D. Hilvert, *J. Comput. Chem.* **2003**, *24*, 98–110.
- [70] K. D. Janda, C. G. Shevlin, R. A. Lerner, *Science* **1993**, *259*, 490–493.
- [71] K. Gruber, B. Zhou, K. N. Houk, R. A. Lerner, C. G. Shevlin, I. A. Wilson, *Biochemistry* **1999**, *38*, 7062–7074.
- [72] J. Na, K. N. Houk, C. G. Shevlin, K. D. Janda, R. A. Lerner, *J. Am. Chem. Soc.* **1993**, *115*, 8453–8454.
- [73] J. Na, K. N. Houk, *J. Am. Chem. Soc.* **1996**, *118*, 9204–9205.
- [74] J. M. Coxon, A. J. Thorpe, *J. Am. Chem. Soc.* **1999**, *121*, 10955–10957.
- [75] K. D. Janda, C. G. Shevlin, R. A. Lerner, *J. Am. Chem. Soc.* **1995**, *117*, 2659–2660.
- [76] K. D. Janda, D. Schloeder, S. J. Benkovic, R. A. Lerner, *Science* **1988**, *241*, 1188–1191.
- [77] J. D. Stewart, J. F. Krebs, G. Siuzdak, A. J. Berdis, D. B. Smi-thrud, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7404–7409.
- [78] C. Gao, B. J. Lavey, C.-H. L. Lo, A. Datta, P. Wentworth, K. D. Janda, *J. Am. Chem. Soc.* **1998**, *120*, 2211–2217.
- [79] V. A. Roberts, J. Stewart, S. J. Benkovic, E. D. Getzoff, *J. Mol. Biol.* **1994**, *235*, 1098–1116.
- [80] M. M. Thayer, E. H. Olander, A. S. Arvai, C. K. Koike, I. L. Canestrelli, J. D. Stewart, S. J. Benkovic, E. D. Getzoff, V. A. Roberts, *J. Mol. Biol.* **1999**, *291*, 329–345.
- [81] L. T. Chong, P. Bandyopadhyay, T. S. Scanlan, I. D. Kuntz, P. A. Kollman, *J. Comput. Chem.* **2003**, *24*, 1371–1377.
- [82] Y. M. Chook, H. Ke, W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8600–8603.
- [83] A. Y. Lee, P. A. Karplus, B. Ganem, J. Clardy, *J. Am. Chem. Soc.* **1995**, *117*, 3627–3628.
- [84] M. R. Haynes, E. A. Stura, D. Hilvert, I. A. Wilson, *Science* **1994**, *263*, 646–652.
- [85] P. Schultz, *Science* **1988**, *240*, 426–433.
- [86] O. Wiest, K. N. Houk, *J. Am. Chem. Soc.* **1995**, *117*, 11628–11639.
- [87] U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185–194.
- [88] M. T. Reetz, P. Soni, J. P. Acevedo, J. Sanchis, *Angew. Chem.* **2009**, *121*, 8418–8422; *Angew. Chem. Int. Ed.* **2009**, *48*, 8268–8272.
- [89] M. Zumárraga, T. Bulter, S. Shleev, J. Polaina, A. Martínez-Arias, F. J. Plou, A. Ballesteros, M. Alcalde, *Chem. Biol.* **2007**, *14*, 1052–1064.
- [90] J. Karanicolas, J. E. Corn, I. Chen, L. A. Joachimiak, O. Dym, S. H. Peck, S. Albeck, T. Unger, W. Hu, G. Liu, S. Delbecq, G. Montelione, C. Spiegel, D. R. Liu, D. Baker, *Mol. Cell* **2011**, *42*, 250–260.
- [91] G. J. Williams, C. Zhang, J. S. Thorson, *Nat. Chem. Biol.* **2007**, *3*, 657–662.
- [92] S. E. Deacon, M. J. McPherson, *ChemBioChem* **2011**, *12*, 593–601.
- [93] R. D. Gupta, M. Goldsmith, Y. Ashani, Y. Simo, G. Mullock-andov, H. Bar, M. Ben-David, H. Leader, R. Margalit, I. Silman, J. L. Sussman, D. S. Tawfik, *Nat. Chem. Biol.* **2011**, *7*, 120–125.
- [94] S. Bartsch, R. Kourist, U. T. Bornscheuer, *Angew. Chem.* **2008**, *120*, 1531–1534; *Angew. Chem. Int. Ed.* **2008**, *47*, 1508–1511.
- [95] M. T. Reetz, *Angew. Chem.* **2011**, *123*, 144–182; *Angew. Chem. Int. Ed.* **2011**, *50*, 138–174.
- [96] M. T. Reetz, B. Brunner, T. Schneider, F. Schulz, C. M. Cloutier, M. M. Kayser, *Angew. Chem.* **2004**, *116*, 4167–4170; *Angew. Chem. Int. Ed.* **2004**, *43*, 4075–4078.
- [97] S. Park, K. L. Morley, G. P. Horsman, M. Holmquist, K. Hult, R. J. Kazlauskas, *Chem. Biol.* **2005**, *12*, 45–54.
- [98] R. J. Fox, S. C. Davis, E. C. Mundorff, L. M. Newman, V. Gavrilovic, S. K. Ma, L. M. Chung, C. Ching, S. Tam, S. Muley, J. Grate, J. Gruber, J. C. Whitman, R. A. Sheldon, G. W. Huisman, *Nat. Biotechnol.* **2007**, *25*, 338–344.
- [99] P. L. Bergquist, R. A. Reeves, M. D. Gibbs, *Biomol. Eng.* **2005**, *22*, 63–72.
- [100] G. Amitai, R. D. Gupta, D. S. Tawfik, *HFSP J.* **2007**, *1*, 67–78.
- [101] J. D. Bloom, P. A. Romero, Z. Lu, F. H. Arnold, *Biol. Direct* **2007**, *2*, 17.
- [102] B. Seelig, J. W. Szostak, *Nature* **2007**, *448*, 828–831.
- [103] J.-C. Baret, O. J. Miller, V. Taly, M. Ryckelynck, A. El-Harrak, L. Frenz, C. Rick, M. L. Samuels, J. B. Hutchison, J. J. Agresti, D. R. Link, D. A. Weitz, A. D. Griffiths, *Lab Chip* **2009**, *9*, 1850–1858.

- [104] J.-P. Goddard, J.-L. Reymond, *Curr. Opin. Biotechnol.* **2004**, *15*, 314–322.
- [105] E. M. Brustad, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2011**, *15*, 201–210.
- [106] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, *Science* **2010**, *329*, 305–309.
- [107] T. T. Li, J. J. Liang, A. A. Ambrogelly, T. T. Brennan, G. G. Gloor, G. G. Huisman, J. J. Lalonde, A. A. Lekhal, B. B. Mijts, S. S. Muley, L. Newman, M. Tobin, G. Wong, A. Zaks, X. Zhang, *J. Am. Chem. Soc.* **2012**, *134*, 6467–6472.
- [108] M. Bocola, N. Otte, K.-E. Jaeger, M. T. Reetz, W. Thiel, *ChemBioChem* **2004**, *5*, 214–223.
- [109] M. T. Reetz, M. Puls, J. D. Carballeira, A. Vogel, K.-E. Jaeger, T. Eggert, W. Thiel, M. Bocola, N. Otte, *ChemBioChem* **2007**, *8*, 106–112.
- [110] G. A. Behrens, A. Hummel, S. K. Padhi, S. Schaetzle, U. T. Bornscheuer, *Adv. Synth. Catal.* **2011**, *353*, 2191–2215.
- [111] G. A. Strohmeier, H. Pichler, O. May, M. Gruber-Khadjawi, *Chem. Rev.* **2011**, *111*, 4141–4164.
- [112] A. S. Bommarius, J. K. Blum, M. J. Abrahamson, *Curr. Opin. Chem. Biol.* **2011**, *15*, 194–200.
- [113] G.-W. Zheng, J.-H. Xu, *Curr. Opin. Biotechnol.* **2011**, *22*, 784–792.
- [114] O. Khersonsky, D. S. Tawfik, *Annu. Rev. Biochem.* **2010**, *79*, 471–505.
- [115] A. C. Joerger, S. Mayer, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5694–5699.
- [116] G. J. Poelarends, J. J. Almrud, H. Serrano, J. E. Darty, W. H. Johnson, M. L. Hackert, C. P. Whitman, *Biochemistry* **2006**, *45*, 7700–7708.
- [117] W. S. Yew, J. Akana, E. L. Wise, I. Rayment, J. A. Gerlt, *Biochemistry* **2005**, *44*, 1807–1815.
- [118] E. L. Wise, W. S. Yew, J. Akana, J. A. Gerlt, I. Rayment, *Biochemistry* **2005**, *44*, 1816–1823.
- [119] D. M. Z. Schmidt, E. C. Mundorff, M. Dojka, E. Bermudez, J. E. Ness, S. Govindarajan, P. C. Babbitt, J. Minshull, J. A. Gerlt, *Biochemistry* **2003**, *42*, 8387–8393.
- [120] Y. Terao, K. Miyamoto, H. Ohta, *Chem. Commun.* **2006**, 3600–3602.
- [121] Y. Ijima, K. Matoishi, Y. Terao, N. Doi, H. Yanagawa, H. Ohta, *Chem. Commun.* **2005**, 877–879.
- [122] Y. Terao, K. Miyamoto, H. Ohta, *Appl. Microbiol. Biotechnol.* **2006**, *73*, 647–653.
- [123] H. Xiang, L. Luo, K. L. Taylor, D. Dunaway-Mariano, *Biochemistry* **1999**, *38*, 7638–7652.
- [124] R. B. Hamed, E. T. Batchelor, I. J. Clifton, C. J. Schofield, *Cell. Mol. Life Sci.* **2008**, *65*, 2507–2527.
- [125] R. B. Hill, D. P. Raleigh, A. Lombardi, W. F. DeGrado, *Acc. Chem. Res.* **2000**, *33*, 745–754.
- [126] B. R. Gibney, F. Rabanal, J. J. Skalicky, A. J. Wand, P. L. Dutton, *J. Am. Chem. Soc.* **1997**, *119*, 2323–2324.
- [127] B. R. Gibney, F. Rabanal, J. J. Skalicky, A. J. Wand, P. L. Dutton, *J. Am. Chem. Soc.* **1999**, *121*, 4952–4960.
- [128] K. T. Simons, C. Kooperberg, E. Huang, D. Baker, *J. Mol. Biol.* **1997**, *268*, 209–225.
- [129] B. I. Dahiyat, S. L. Mayo, *Science* **1997**, *278*, 82–87.
- [130] G. G. Dantas, B. B. Kuhlman, D. D. Callender, M. M. Wong, D. D. Baker, *J. Mol. Biol.* **2003**, *332*, 449–460.
- [131] B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, D. Baker, *Science* **2003**, *302*, 1364–1368.
- [132] P.-S. Huang, J. J. Love, S. L. Mayo, *Protein Sci.* **2007**, *16*, 2770–2774.
- [133] S. J. Fleishman, T. A. Whitehead, D. C. Ekiert, C. Dreyfus, J. E. Corn, E.-M. Strauch, I. A. Wilson, D. Baker, *Science* **2011**, *332*, 816–821.
- [134] H. Yin, J. S. Slusky, B. W. Berger, R. S. Walters, G. Vilaira, R. I. Litvinov, J. D. Lear, G. A. Caputo, J. S. Bennett, W. F. DeGrado, *Science* **2007**, *315*, 1817–1822.
- [135] S. J. Shandler, I. V. Korendovych, D. T. Moore, K. B. Smith-Dupont, C. N. Streu, R. I. Litvinov, P. C. Billings, F. Gai, J. S. Bennett, W. F. DeGrado, *J. Am. Chem. Soc.* **2011**, *133*, 12378–12381.
- [136] G. Grigoryan, Y. H. Kim, R. Acharya, K. Axelrod, R. M. Jain, L. Willis, M. Drndic, J. M. Kikkawa, W. F. DeGrado, *Science* **2011**, *332*, 1071–1076.
- [137] B. S. Chevalier, T. Kortemme, M. S. Chadsey, D. Baker, R. J. Monnat, B. L. Stoddard, *Mol. Cell* **2002**, *10*, 895–905.
- [138] J. Ashworth, J. J. Havranek, C. M. Duarte, D. Sussman, R. J. Monnat, B. L. Stoddard, D. Baker, *Nature* **2006**, *441*, 656–659.
- [139] U. Y. Ulge, D. A. Baker, R. J. Monnat, *Nucleic Acids Res.* **2011**, *39*, 4330–4339.
- [140] J. Ashworth, G. K. Taylor, J. J. Havranek, S. A. Quadri, B. L. Stoddard, D. Baker, *Nucleic Acids Res.* **2010**, *38*, 5601–5608.
- [141] S. B. Thyme, J. Jarjour, R. Takeuchi, J. J. Havranek, J. Ashworth, A. M. Scharenberg, B. L. Stoddard, D. Baker, *Nature* **2009**, *461*, 1300–1304.
- [142] J. S. Marvin, H. W. Hellinga, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4955–4960.
- [143] W. Yang, L. M. Jones, L. Isley, Y. Ye, H.-W. Lee, A. Wilkins, Z.-R. Liu, H. W. Hellinga, R. Malchow, M. Ghazi, J. J. Yang, *J. Am. Chem. Soc.* **2003**, *125*, 6165–6171.
- [144] L. L. Looger, M. A. Dwyer, J. J. Smith, H. W. Hellinga, *Nature* **2003**, *423*, 185–190.
- [145] M. Allert, S. S. Rizk, L. L. Looger, H. W. Hellinga, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7907–7912.
- [146] B. Schreier, C. Stumpf, S. Wiesner, B. Hocker, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18491–18496.
- [147] F. E. Boas, P. B. Harbury, *J. Mol. Biol.* **2008**, *380*, 415–424.
- [148] H. C. Fry, A. Lehmann, J. G. Saven, W. F. DeGrado, M. J. Therien, *J. Am. Chem. Soc.* **2010**, *132*, 3997–4005.
- [149] A. Morin, K. W. Kaufmann, C. Fortenberry, J. M. Harp, L. S. Mizoue, J. Meiler, *Protein Eng. Des. Sel.* **2011**, *24*, 503–516.
- [150] K. Johnsson, R. K. Allemann, H. Widmer, S. A. Benner, *Nature* **1993**, *365*, 530–532.
- [151] R. A. Lerner, S. J. Benkovic, P. G. Schultz, *Science* **1991**, *252*, 659–667.
- [152] T. Sasaki, E. T. Kaiser, *J. Am. Chem. Soc.* **1989**, *111*, 380–381.
- [153] K. S. Broo, L. Brive, P. Ahlberg, L. Baltzer, *J. Am. Chem. Soc.* **1997**, *119*, 11362–11372.
- [154] P. Rossi, P. Tecilla, L. Baltzer, P. Scrimin, *Chem. Eur. J.* **2004**, *10*, 4163–4170.
- [155] B. Lovejoy, S. Choe, D. Cascio, D. K. McRorie, W. F. DeGrado, D. Eisenberg, *Science* **1993**, *259*, 1288–1293.
- [156] C. Tommos, J. J. Skalicky, D. L. Pilloud, A. J. Wand, P. L. Dutton, *Biochemistry* **1999**, *38*, 9495–9507.
- [157] S. Chakraborty, J. Y. Kravitz, P. W. Thulstrup, L. Hemmingsen, W. F. DeGrado, V. L. Pecoraro, *Angew. Chem.* **2011**, *123*, 2097–2101; *Angew. Chem. Int. Ed.* **2011**, *50*, 2049–2053.
- [158] C. M. Summa, M. M. Rosenblatt, J.-K. Hong, J. D. Lear, W. F. DeGrado, *J. Mol. Biol.* **2002**, *321*, 923–938.
- [159] D. N. Bolon, S. L. Mayo, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14274–14279.
- [160] C. Pabo, *Nature* **1983**, *301*, 200–200.
- [161] J. W. Ponder, F. M. Richards, *J. Mol. Biol.* **1987**, *193*, 775–791.
- [162] J. Bowie, R. Luthy, D. Eisenberg, *Science* **1991**, *253*, 164–170.
- [163] K. Yue, K. A. Dill, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 4163–4167.
- [164] H. M. Berman, *Nucleic Acids Res.* **2000**, *28*, 235–242.
- [165] D. Tantillo, *Stimulating Concepts in Chemistry*, Wiley-VCH, Weinheim, **2000**.

- [166] A. Zanghellini, L. Jiang, A. M. Wollacott, G. Cheng, J. Meiler, E. A. Althoff, D. Röthlisberger, D. Baker, *Protein Sci.* **2006**, *15*, 2785–2794.
- [167] F. Richter, A. Leaver-Fay, S. D. Khare, S. Bjelic, D. Baker, *PLoS ONE* **2011**, *6*, e19230.
- [168] A. J. T. Smith, R. Müller, M. D. Toscano, P. Kast, H. W. Hellinga, D. Hilvert, K. N. Houk, *J. Am. Chem. Soc.* **2008**, *130*, 15361–15373.
- [169] G. Kiss, D. Röthlisberger, D. Baker, K. N. Houk, *Protein Sci.* **2010**, *19*, 1760–1773.
- [170] L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Röthlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas, D. Hilvert, K. N. Houk, B. L. Stoddard, D. Baker, *Science* **2008**, *319*, 1387–1391.
- [171] D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik, D. Baker, *Nature* **2008**, *453*, 190–195.
- [172] J. B. Siegel, A. Zanghellini, H. M. Lovick, G. Kiss, A. R. Lambert, J. L. St. Clair, J. L. Gallaher, D. Hilvert, M. H. Gelb, B. L. Stoddard, K. N. Houk, F. E. Michael, D. Baker, *Science* **2010**, *329*, 309–313.
- [173] F. W. Studier, *Protein Expression Purif.* **2005**, *41*, 207–234.
- [174] E. Sulkowski, *Trends Biotechnol.* **1985**, *3*, 1–7.
- [175] R. Janknecht, G. de Martynoff, J. Lou, R. A. Hipskind, A. Nordheim, H. G. Stunnenberg, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8972–8976.
- [176] *Guide to Protein Purification* (Hrsg.: R. R. Burgess, M. P. Deutscher), Academic Press, **2009**.
- [177] T. Gefflaut, C. Blonski, J. Perie, M. Willson, *Prog. Biophys. Mol. Biol.* **1995**, *63*, 301–340.
- [178] A. Heine, *Science* **2001**, *294*, 369–374.
- [179] J. Z. Ruscio, J. E. Kohn, K. A. Ball, T. Head-Gordon, *J. Am. Chem. Soc.* **2009**, *131*, 14111–14115.
- [180] J. K. Lassila, D. Baker, D. Herschlag, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4937–4942.
- [181] F. Hollfelder, A. J. Kirby, D. S. Tawfik, *J. Am. Chem. Soc.* **1997**, *119*, 9578–9579.
- [182] I. V. Korendovych, D. W. Kulp, Y. Wu, H. Cheng, H. Roder, W. F. DeGrado, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6823–6827.
- [183] A. J. Kirby, *Acc. Chem. Res.* **1997**, *30*, 290–296.
- [184] D. S. Kemp, *Nature* **1995**, *373*, 196–197.
- [185] O. Khersonsky, D. Röthlisberger, O. Dym, S. Albeck, C. J. Jackson, D. Baker, D. S. Tawfik, *J. Mol. Biol.* **2010**, *396*, 1025–1042.
- [186] O. Khersonsky, D. Röthlisberger, A. M. Wollacott, P. Murphy, O. Dym, S. Albeck, G. Kiss, K. N. Houk, D. Baker, D. S. Tawfik, *J. Mol. Biol.* **2011**, *407*, 391–412.
- [187] O. Khersonsky, G. Kiss, D. Röthlisberger, O. Dym, S. Albeck, K. N. Houk, D. Baker, D. S. Tawfik, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10358–10363.
- [188] A. N. Alexandrova, D. Röthlisberger, D. Baker, W. L. Jorgenson, *J. Am. Chem. Soc.* **2008**, *130*, 15907–15915.
- [189] M. P. Frushicheva, J. Cao, Z. T. Chu, A. Warshel, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16869–16874.
- [190] M. P. Frushicheva, J. Cao, A. Warshel, *Biochemistry* **2011**, *50*, 3849–3858.
- [191] H. K. Privett, G. Kiss, T. M. Lee, R. Blomberg, R. A. Chica, L. M. Thomas, D. Hilvert, K. N. Houk, S. L. Mayo, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 3790–3795.
- [192] T. Steiner, *Angew. Chem.* **2002**, *114*, 50–80; *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76.
- [193] L. Lo Leggio, S. Kalogiannis, K. Eckert, S. C. Teixeira, M. K. Bhat, C. Andrei, R. W. Pickersgill, S. Larsen, *FEBS Lett.* **2001**, *509*, 303–308.
- [194] B. I. Dahiyat, D. B. Gordon, S. L. Mayo, *Protein Sci.* **1997**, *6*, 1333–1337.
- [195] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenken, A. Leaver-Fay, D. Baker, Z. Popović, F. Players, *Nature* **2010**, *466*, 756–760.
- [196] F. Khatib, S. Cooper, M. D. Tyka, K. Xu, I. Makedon, Z. Popović, D. Baker, F. Players, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18949–18953.
- [197] F. Khatib, F. DiMaio, Foldit Contenders Group, Foldit Void Crushers Group, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zábranská, I. Pichová, J. Thompson, Z. Popović, M. Jaskolski, D. Baker, *Nat. Struct. Mol. Biol.* **2011**, *18*, 1175–1177.
- [198] C. B. Eiben, J. B. Siegel, J. B. Bale, S. Cooper, F. Khatib, B. W. Shen, F. Players, B. L. Stoddard, Z. Popović, D. Baker, *Nat. Biotechnol.* **2012**, *30*, 190–192.
- [199] D. A. Kraut, P. A. Sigala, T. D. Fenn, D. Herschlag, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1960–1965.
- [200] J. Villà, A. Warshel, *J. Phys. Chem. B* **2001**, *105*, 7887–7907.
- [201] W. R. Cannon, S. F. Singleton, S. J. Benkovic, *Nat. Struct. Mol. Biol.* **1996**, *3*, 821–833.
- [202] H. Liu, A. Warshel, *Biochemistry* **2007**, *46*, 6011–6025.
- [203] E. W. Debler, R. Müller, D. Hilvert, I. A. Wilson, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18539–18544.
- [204] A. Korkeian, M. E. Black, D. Baker, B. L. Stoddard, *Science* **2005**, *308*, 857–860.
- [205] M. Lehmann, L. Pasamontes, S. F. Lassen, M. Wyss, *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* **2000**, *1543*, 408–415.
- [206] R. Carlson, *Nat. Biotechnol.* **2009**, *27*, 1091–1094.
- [207] M. Baker, *Nature* **2011**, *473*, 403–408.
- [208] J. Tian, K. Ma, I. Saaem, *Mol. BioSyst.* **2009**, *5*, 714–722.
- [209] S. D. Khare, Y. Kipnis, P. J. Greisen, R. Takeuchi, Y. Ashani, M. Goldsmith, Y. Song, J. L. Gallaher, I. Silman, H. Leader, J. L. Sussman, *Nat. Chem. Biol.* **2012**, *8*, 294–300.
- [210] A. C. Hemmert, T. C. Otto, R. A. Chica, M. Wierdl, J. S. Edwards, S. M. Lewis, S. L. Lewis, C. C. Edwards, L. Tsurkan, C. L. Cadieux, S. A. Kasten, J. R. Cashman, S. L. Mayo, P. M. Potter, D. M. Cerasoli, M. R. Redinbo, *PLoS ONE* **2011**, *6*, e17441.
- [211] F. Richter, R. Blomberg, S. D. Khare, G. Kiss, A. P. Kuzin, A. J. T. Smith, J. Gallaher, Z. Pianowski, R. C. Helgeson, A. Grjasnow, R. Xiao, J. Seetharaman, M. Su, S. Vorobiev, S. Lew, F. Forouhar, G. J. Kornhaber, J. F. Hunt, G. T. Montelione, L. Tong, K. N. Houk, D. Hilvert, D. Baker, *J. Am. Chem. Soc.* **2012**, *134*, 16197–16206.
- [212] S. R. Gordon, E. J. Stanley, S. Wolf, A. Toland, S. J. Wu, D. Hadidi, J. H. Mills, D. Baker, I. S. Pultz, J. B. Siegel, *J. Am. Chem. Soc.* **2012**, DOI: 10.1021/ja3094795.
- [213] P. A. M. Dirac, *Proc. R. Soc. London Ser. A* **1929**, *123*, 714–733.