

Systematische Analyse von Epoxidhydrolasen:

**Erstellung einer familienspezifischen Datenbank,
Entwicklung einer strukturbezogenen Klassifizierung
und Amplifikation unbekannter Epoxidhydrolasen.**

Von der Fakultät Geo- und Biowissenschaften der Universität Stuttgart
zur Erlangung der Würde eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

vorgelegt von
Sandra Barth
aus Neuenbürg (Württ.)

Hauptberichter:	Prof. Dr. Rolf D. Schmid
Mitberichter:	Priv. Doz. Dr. Jürgen Pleiss
Prüfungsvorsitzender:	Prof. Dr. Hans-Dieter Görtz
Tag der mündlichen Prüfung:	18.12.2003

INSTITUT FÜR TECHNISCHE BIOCHEMIE

Es ist nicht genug, zu wissen, man muss es auch anwenden; es ist nicht genug, zu wollen, man muss es auch tun.

Johann Wolfgang von Goethe

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich für die Erstellung der vorliegenden Arbeit keine weiteren als die angegebenen Quellen und Hilfsmittel verwendet habe.

Danksagung

Im Laufe meiner Promotion habe ich von vielen Seiten Unterstützung erfahren, ohne die die Fertigstellung dieser Arbeit nicht möglich gewesen wäre. Ich möchte nun an dieser Stelle denen danken, die mir in den letzten Jahren bei dieser Arbeit geholfen haben.

Herrn **Prof. Rolf D. Schmid** für die Überlassung und Betreuung des Themas, seine unermüdliche Unterstützung und die hervorragenden Arbeitsbedingungen am ITB.

Herrn **Priv. Doz. Dr. Jürgen Pleiss** für die ausgezeichnete Betreuung dieser Arbeit während der vergangenen Jahre.

Der **BASF AG, Ludwigshafen**, im Besonderen Herrn **Dr. Bernhard Hauer** und Frau **Dr. Maria Keßeler**, für die finanzielle, sowie Herrn **Dr. Thomas Friedrich** und Herrn **Andreas Schädler** für die große Unterstützung bei der Reinigung der Epoxidhydrolase aus *Streptomyces antibioticus* Tü4.

Frau **Dr. Jutta Schmitt**, Herrn **Dr. Stefan Lange** und Herrn **Dr. Eckart Bonacker** für die Betreuung meiner Laborarbeit und für die Hilfe bei der Bewältigung der ersten Hürden.

Herrn **Kai Doderer** für seine unendliche Geduld bei chemischen und „gaschromatographischen“ Fragen und den vielen erhellenden Diskussionen.

Herrn **Markus Fischer** für seine unermüdliche Unterstützung in Fragen der LED, vielen Computerproblemen und für so manchen unterhaltsamen Abend.

Frau **Monika Rusnak**, Frau **Nicole Beuttenmüller** und Frau **Sabine Eiben** für viel Spaß, Sport und leckeres Essen auch abseits des Institutsalltags.

Meinem Lebensgefährten **Florian Barth** für schnelle Hilfe bei Cluster- und Linuxfragen, seine ständige Diskussionsbereitschaft und die vielen Ausflüge ins kulinarische Neuland.

Meinen Eltern **Karl Heinz und Roswitha Barth** für die Unterstützung während meines Studiums und meiner Promotion, die diese Arbeit erst ermöglichte.

Abkürzungsverzeichnis

% (M/V)	Massenprozent
% (V/V)	Volumenprozent
°C	Grad Celsius
μ	Mikro
μl	Mikroliter
Amp	Ampicillin
APS	Ammoniumpersulfat
AS	Aminosäure
ATCC	American Type Culture Collection
BL	<i>Bacillus subtilis</i> Lipase
bp	Basenpaare
<i>Burk. L</i>	<i>Burkolderia</i> Lipase
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
CL	<i>Candida antarctica</i> Lipase
d	demineralisiert
dATP	Desoxyadenosid-5'-triphosphat
DBMS	Datenbank-Management-System
DC	Dünnschichtchromatographie
dCTP	Desoxycytidin-5'-triphosphat
dd	doppelt-demineralisiert
DDBJ	DNA Data Bank of Japan
dest	destilliert
dGTP	Desoxyguanosin-5'-triphosphat
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
dNTP	Desoxyribonukleosid-5'-triphosphat
dNTP	Didesoxyribonukleosid-5'-triphosphat
DMF	Dimethylformamid
DSMZ	Deutsche Stammsammlung für Mikroorganismen und Zellkulturen
dTTP	Desoxythymidin-5'-triphosphat
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	N-N-N'-N'-Ethyldiamintetraacetat
ee	Enantiomerenüberschuss
EH	Epoxidhydrolase
EMBL	European Molecular Biology Laboratory
FC	<i>Fusarium</i> Cutinase
FID	Flammenionisations Detektor

g	Gramm
GC	Gaschromatographie
GL	Gastrische Lipase
gr.	griechisch
h	Stunde
HD	Haloalkan Dehalogenase
JH	Juvenilhormon
kb	Kilobasen
kDa	Kilodalton
kPa	Kilopascal
l	Liter
lacZ	verkürztes Galactosidase Gen (<i>E. coli</i>)
LB	Luria-Bertani (Medium)
LMW	Low Molecular Weight Standard
Lsg	Lösung
m	Milli
M	Mol pro Liter
min	Minute
ML	Moraxella Lipase
NCBI	National Center for Biotechnology Information
NP	Non-heme Peroxidase
mS	Millisiemens
OD	Optische Dichte
p.a.	pro analysi
PAGE	Polyacrylamid Gelelektrophorese
PEG	Polyethylenglykol
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PIR	Protein Information Resource
PL	Pankreatische Lipase
RDBMS	relationales Datenbank-Management-System
RL	<i>Rhizomucor miehei</i> Lipase
RNA	Ribonukleinsäure
RT	Raumtemperatur
s	Sekunde
SDS	Natriumdodecylsulfat
SET	Natriumchlorid/EDTA/Tris
sog.	sogenannt
TAE	Tris/Acetat/EDTA
<i>Taq</i>	<i>Thermus aquaticus</i>
TE	Tris-EDTA

TEMED	N,N,N',N'-Tetramethylethyldiamin
Tris	Tris-(hydroxymethyl)-aminomethan
TFA	Tri-Fluor-Essigsäure
TSB	Tryptic Soy Broth (Medium)
TSS	Transformation Storage Solution
U	Unit (Einheit der Enzymaktivität)
ÜK	Übernachtskultur
ÜN	Übernachtskultur
Upm	Umdrehungen pro Minute
V	Volt
Vol.	Volumen
WT	Wildtyp
z.B.	zum Beispiel

Des Weiteren wurden die üblichen Codes für Aminosäuren und Nucleotide, sowie die gängigen Abkürzungen für Restriktionsenzyme, Polymerasen usw. verwendet.

Inhaltsverzeichnis

<i>Eidesstattliche Erklärung</i>	3
<i>Danksagung</i>	4
<i>Abkürzungsverzeichnis</i>	5
<i>Inhaltsverzeichnis</i>	8
1 Kurzzusammenfassung	14
2 Abstract	15
3 Einleitung	21
3.1 Epoxidhydrolasen	21
3.1.1 Allgemeines	21
3.1.2 Säugetier EHs	21
3.1.3 Pflanzen EHs	22
3.1.4 Insekten EHs	22
3.1.5 EHs aus Hefen und Pilzen	23
3.1.6 Bakterielle EHs	23
3.1.7 Bekannte Strukturen	24
3.1.8 Epoxide	24
3.1.9 Reaktionsmechanismus	25
3.2 α/β Hydrolase Fold	27
3.2.1 Haloalkan Dehalogenasen	28
3.3 Streptomyceten	29
3.4 Phylogenie	31
3.4.1 Erstellung von Stammbäumen	34
3.4.2 Die molekulare Uhr	35
3.4.3 Mathematische Algorithmen zur Erstellung phylogenetischer Bäume	36
3.4.3.1 Datengrundlage	36
3.4.3.2 Distanzbasierte Methoden	37
3.4.3.2.1 UPGMA	37
3.4.3.2.2 Neighbor-Joining	37
3.4.3.3 Merkmalsbasierte Methoden	37

3.4.3.3.1	Maximum-Parsimony	38
3.4.3.3.2	Maximum-Likelihood	38
3.5	Datenbanken	39
3.5.1	Datenbanktypen	39
3.5.1.1	Flatfile-Datenbanken	39
3.5.1.2	Relationale Datenbanken	39
3.5.1.3	Objektorientierte Datenbanken	40
4	Ergebnisse	41
4.1	Zielsetzung	41
4.2	Teil 1: Systematische Analyse	42
4.2.1	Das Multisequenz-Alignment	42
4.2.2	Superfamilien	44
4.2.3	Einteilung in homologe Familien	46
4.2.4	Konservierte AS	49
4.2.5	Die β 6-Position	49
4.2.6	Remodellierung der Kristallstruktur der <i>Agrobacterium radiobacter</i> EH	51
4.2.7	Modularer Aufbau von EHs	52
4.2.7.1	Loop-Längen	55
4.2.7.2	Der NC-Loop	57
4.2.8	Der Cap-Loop	58
4.2.9	Homologiemodellierung	59
4.3	Familien-spezifische, degenerierte Primer	64
4.3.1	Amplifikation der EH aus <i>Streptomyces antibioticus</i> Tü4	65
4.3.1.1	Fermentation und Reinigung	67
4.3.1.2	N-terminale Ansequenzierung	68
4.3.1.3	Amplifikation von N-Terminus und CODEHOP-Fragment	69
4.3.1.4	Aufbau des C-Terminus	69
4.3.1.5	Trypsin-Verdau	71
4.3.2	Amplifikation des EH Fragments aus <i>Rhodococcus ruber</i>	72
4.4	Teil 2: Erstellung und Auswertung der EH/HD Datenbank	73
4.4.1	Die EH/HD Datenbank	73
4.4.2	Superfamilien	74

Inhaltsverzeichnis	10
<hr/>	
4.4.2.1	Cytosolische Hydrolasen Superfamilie (cHS) _____ 74
4.4.2.2	Mikrosomale Hydrolasen Superfamilie (mHS) _____ 74
4.4.3	Multisequenz-Alignments _____ 75
4.4.4	Phylogenetische Analyse _____ 75
4.4.5	Strukturelle Untersuchungen an α/β Hydrolasen _____ 75
4.4.6	Topologie des Übergangszustandes _____ 79
4.4.7	Anker Konzept _____ 80
4.4.8	Anker-Struktur Beziehung _____ 83
4.4.9	Das GXGXS-Motiv _____ 84
5	<i>Diskussion</i> _____ 86
5.1	Systematische Analyse von EHs _____ 86
5.1.1	Rolle der modularen Architektur _____ 86
5.1.2	Klassifizierung und Substratspektrum _____ 87
5.1.3	Funktionalität _____ 89
5.1.4	Homologiemodellierung _____ 89
5.2	Familienspezifische degenerierte Primer _____ 90
5.2.1	Screening nach unbekanntem EHs _____ 90
5.2.1.1	CODEHOP-Primer _____ 91
5.2.1.2	Amplifikation der CODEHOP-Fragmente _____ 92
5.2.1.3	Vervollständigen der Sequenz <i>Streptomyces antibioticus</i> Tü4 EH _____ 92
5.3	Datenbanken _____ 92
5.3.1	Die EH/HD Datenbank _____ 93
5.3.2	Reduktiver Fold und Anker-Prinzip _____ 94
5.3.3	Das GXGXS-Motiv _____ 96
6	<i>Zusammenfassung</i> _____ 98
7	<i>Material und Methoden</i> _____ 102
7.1	Theoretische Arbeiten _____ 102
7.1.1	Software und Hardware _____ 102
7.1.1.1	Software _____ 102
7.1.1.2	Datenbanken _____ 102
7.1.1.3	Hardware _____ 103
7.1.2	Bioinformatische Methoden _____ 103

7.1.2.1	Sequenz-Alignments	103
7.1.2.1.1	Globale Alignments	103
7.1.2.1.2	Lokale Alignments	104
7.1.2.2	Substitutionsmatrizen	105
7.1.2.3	Homologie Suche	106
7.1.2.4	Score und E-Wert	106
7.1.2.4.1	EH Sequenzen für die systematische Analyse	107
7.1.2.4.2	BLAST-Suchen kurzer AS-Sequenzen	107
7.1.2.4.3	EH/HD Datenbank	107
7.1.2.5	Phylogenetische Analyse	108
7.1.2.5.1	Neighbor-Joining	108
7.1.2.5.2	Maximum-Likelihood	109
7.1.2.6	Erstellung degenerierter Primer mit CODEHOP	111
7.1.2.7	Homologiemodellierung	112
7.1.2.7.1	Remodellierung	112
7.1.2.7.2	Homologiemodelle	113
7.1.2.8	Molekulardynamische Simulation	113
7.1.2.8.1	Ladungsberechnung	113
7.1.2.8.2	Minimierung	114
7.1.2.8.3	Solvatisierung und Neutralisierung des Proteins	115
7.1.2.8.4	Aufheizen	115
7.2	Experimentelle Arbeiten	117
7.2.1	Material	117
7.2.1.1	Chemikalien und Enzyme	117
7.2.1.2	Geräte	118
7.2.1.3	Plasmide	119
7.2.1.4	Mikroorganismen	119
7.2.1.5	Medien, Puffer und Lösungen	120
7.2.1.6	Puffer für die Agarose-Gel-Elektrophorese	121
7.2.1.7	Puffer und Lösungen für die SDS-PAGE	121
7.2.1.8	Lösungen für die Coomassie-Färbung von SDS-Gelen	122
7.2.1.9	Puffer für die Schnellisolierung von Plasmid-DNA	123
7.2.1.10	dNTP-Mix für die DNA-Sequenzierung	123
7.2.2	Molekularbiologische Methoden	124

7.2.2.1	Stammhaltung und Kultivierung von <i>E. coli</i> -Stämmen	124
7.2.2.1.1	Stammhaltung auf festen Nährböden	124
7.2.2.1.2	Anzucht in Flüssigmedien	124
7.2.2.2	Stammhaltung und Kultivierung von <i>Streptomyces</i> -Stämmen	124
7.2.2.2.1	Anzucht in Flüssigmedien	124
7.2.2.2.2	Fermentation	125
7.2.2.3	Zellaufschluss von <i>E. coli</i> -Zellen	125
7.2.2.4	Zellaufschluss von <i>Streptomyces</i> -Zellen	126
7.2.2.4.1	Zellaufschluss mit Ultraschall	126
7.2.2.4.2	Zellaufschluss für die Proteinreinigung	126
7.2.2.5	Zellaufschluss von <i>Rhodococcus</i> -Zellen	126
7.2.2.6	DNA-Isolierung	127
7.2.2.6.1	Gesamt-DNA-Isolierung aus Actinomyceten	127
7.2.2.6.2	Gesamt-DNA-Isolierung mit Anionenaustauschersäulen	127
7.2.2.7	Plasmidisolierung	127
7.2.2.7.1	Isolierung von Plasmid-DNA aus <i>E. coli</i>	127
7.2.2.7.2	Isolierung reiner Plasmid-DNA aus <i>E. coli</i>	128
7.2.2.7.3	Isolierung größerer Mengen an Plasmid-DNA aus <i>E. coli</i>	128
7.2.3	Enzymatische Methoden	129
7.2.3.1	Verdau von DNA mit Restriktionsendonukleasen	129
7.2.3.2	Ligation mit T4-Ligase	129
7.2.3.3	T/A-Klonierung	129
7.2.3.4	Herstellung kompetenter Zellen	130
7.2.3.5	Transformation nach der Heatshock-Methode	130
7.2.3.5.1	<i>E. coli</i> DH5 α -Zellen	130
7.2.3.5.2	One Shot [®] -Zellen (Invitrogen)	130
7.2.3.6	Polymerase-Kettenreaktion	131
7.2.3.7	Sequenzierung	132
7.2.3.7.1	DNA-Sequenzierung mit dem 377A-DNA-Sequencer	133
7.2.4	Analytische Methoden	135
7.2.4.1	Bestimmung der Epoxidhydrolase-Aktivität	135
7.2.4.1.1	Chirale Kapillar-Gaschromatographie	135
7.2.4.1.2	In der Mikrotiterplatte	136
7.2.4.2	Proteinreinigung	137

Inhaltsverzeichnis	13
<hr/>	
7.2.4.2.1 Ionenaustausch-Chromatographie	137
7.2.4.2.2 Hydrophobe Interaktionschromatographie	137
7.2.4.2.3 Molekularsieb-Chromatographie	137
7.2.4.3 In-Gel Verdau der Proteinbande mit Trypsin	139
7.2.4.4 N-terminale Sequenzanalyse: Edman-Abbau	140
8 Anhang	141
8.1 Accession Codes der verwendeten Sequenzen	141
8.2 Sequenzen	146
8.2.1 Primer	146
8.2.2 Sequenzen der CODEHOP-Fragmente	147
8.2.2.1 <i>Streptomyces antibioticus</i> Tü4	147
8.2.2.1.1 Sequenz des 250 bp-Fragments (236 bp)	147
8.2.2.1.2 DNA-Sequenz des CODEHOP-Fragments (771 bp)	147
8.2.2.1.3 AS-Sequenz des CODEHOP-Fragments (284 AS)	147
8.2.2.2 <i>Rhodococcus ruber</i> (LU760)	148
8.2.2.2.1 DNA-Sequenz des CODEHOP-Fragment (703 bp)	148
8.2.2.2.2 AS-Sequenz des CODEHOP-Fragments (215 AS)	148
9 Literatur	149
CURRICULUM VITAE	157

1 Kurzzusammenfassung

Im Laufe dieser Arbeit wurde eine systematische Analyse an 93 Epoxidhydrolase (EH) Sequenzen durchgeführt. Ein Vergleich der drei bekannten EH Strukturen aus *Agrobacterium radiobacter*, *Mus musculus* und *Aspergillus niger* zeigte die hohe Konservierung des α/β Hydrolase Folds und die Existenz zweier variabler Loop-Regionen, dem NC- und dem Cap-Loop (13). Aus dem Multisequenz-Alignment der 93 EHs wurden die Loop-Längen aller EHs bestimmt und in einem Diagramm gegeneinander aufgetragen. Hierbei bildeten sich drei nach Superfamilien und Organismengruppen getrennte Cluster. Jedes Cluster enthält eine der drei bekannten Strukturen. Unter der Annahme, dass ähnliche Loop-Längen eine ähnliche Struktur bedingen, konnten Homologiemodelle von fünf EHs erstellt werden. Vergleiche der Substratspektren mit der Cluster Zugehörigkeit führten zu der Annahme, dass Loop-Länge und Substratspezifität korrelieren, da z.B. nur für EHs mit langen Cap-Loops der Umsatz epoxidierter Fettsäuren beschrieben wurde.

Mit dem Programm CODEHOP wurden degenerierte familienspezifische Primer erstellt und an zwei bakteriellen Organismen getestet. Aus *Streptomyces antibioticus* Tü4, welcher eine EH Aktivität gegen Styroloxid aufweist (150) und *Rhodococcus ruber* LU760 konnten jeweils Fragmente mit signifikanter Homologie zu EHs amplifiziert werden.

Der zweite Teil dieser Arbeit befaßt sich mit der Erstellung und Auswertung einer familienspezifischen Datenbank auf der Grundlage der Lipase Engineering Database (LED) (45). Die Epoxidhydrolase/Haloalkandehalogenase (EH/HD) Datenbank enthält 397 Sequenzeinträge, 305 Proteineinträge und für 19 dieser Proteineinträge 51 Strukturdatensätze. Neben EHs und HDs enthält die Datenbank weitere Enzymfamilien, die in 14 homologe Familien und zwei Superfamilien eingeteilt wurden.

Untersuchungen der Strukturen aus EH/HD Datenbank und LED zeigten, dass trotz massiver struktureller Unterschiede, zwischen den Strukturen dieser beiden Datenbanken, das katalytische Zentrum strukturell in allen α/β Hydrolasen erhalten ist. Auch das für Lipasen ermittelte Anker-Konzept (111) für die Stabilisierung des Oxyanionholes greift bei den Proteinen der EH/HD Datenbank. Für das hochkonservierte GXGXS-Motiv (92) konnte eine mögliche strukturelevante Funktion entwickelt werden, die seine konservierte Struktur und Lage erklärt.

2 Abstract

Epoxide hydrolases (EHs) from eukaryotic and prokaryotic sources were found to be applicable for the kinetic resolution of racemic mixtures of epoxides, and high enantiomeric accesses have been reported (117). Enantiopure epoxides and vicinal diols are extensively employed as useful chiral building blocks for synthesis of various bioactive products in the pharmaceutical and agrochemical industry (51).

Aim of this work was to identify similarities and differences in this highly divers enzyme family by systematic analysis of available EH sequences and structures.

93 different EHs were identified by keyword and BLAST search. A phylogenetic analysis showed the clear separation into the two separate superfamilies cytosolic (cEHs) and microsomal (mEHs) EHs and several homologous families within each superfamily. In contrast to eukaryotic EHs, bacterial EHs show a high variability and barley form any well defined homologous families.

EHs belong to the family of α/β hydrolases. Analysis of the three known structures from *Agrobacterium radiobacter*, *Mus musculus* und *Aspergillus niger* lead to division of the modular structures into N-terminal domain, N-terminal catalytic domain, NC-loop, cap-domain, cap-loop and C-terminal catalytic domain (Figure 1).

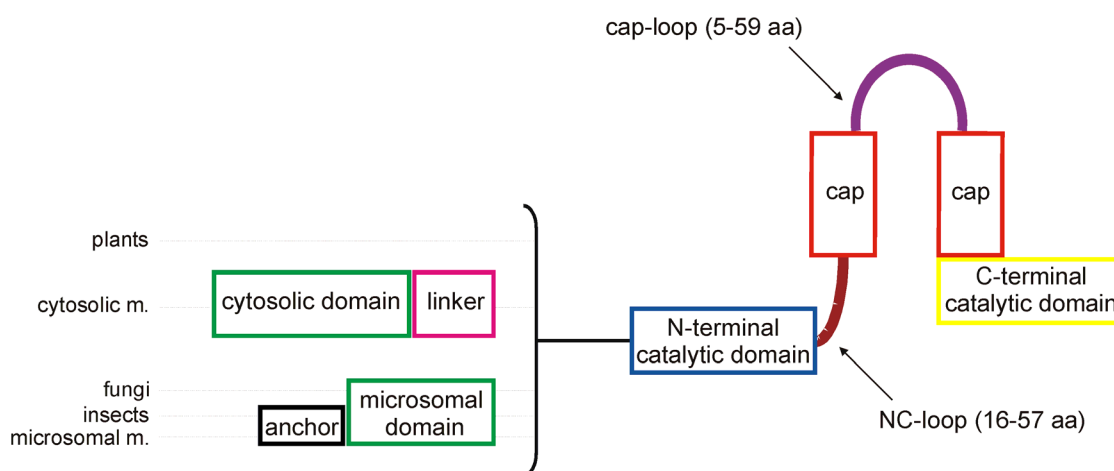


Figure 1

Schematic illustration of the modular architecture of EHs. The variable N-terminal domain (green, left half of the picture) occurring in cytosolic mammalian EHs (connected by a linker, pink), fungi/yeasts, insects and microsomal mammalian EHs. Insect and microsomal mammalian EHs additionally contain a membrane anchor (black). In the right half of the picture the three conserved domains are given: the N-terminal catalytic domain (blue), the cap domain (red) and the C-terminal catalytic domain (yellow) and the two variable loops: NC-loop (brown) and cap-loop (violet).

Besides obvious differences in molecule size of the three structures, caused by the variable N-terminal domain, EHs contain a highly conserved α/β hydrolase fold only interrupted by two variable loops: NC- (connects N-terminal catalytic domain and cap-domain) and cap-loop (located within the cap-domain). The bacterial EH from *Agrobacterium radiobacter* contains short loops (NC-loop: 19 residues, cap-loop: 9 residues), the EH from *Mus musculus* contains a medium sized NC-loop (23 residues) and a long cap-loop (35 residues) and the microsomal EH from *Aspergillus niger* contains a long NC-loop (35 residues) and a short cap-loop (8 residues). Loop-lengths of all EHs were determined using the multisequence alignment of 93 EHs and were plotted against each other in a diagram (Figure 2).

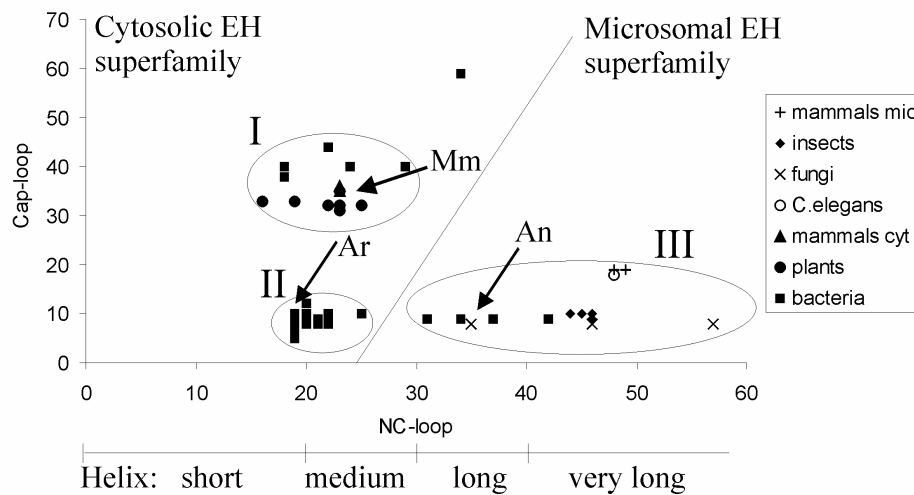


Figure 2

A plot of the lengths of NC- and cap-loop, three clusters are formed, separated into both superfamilies: cytosolic EHs, cluster I (mammalian EHs, plant EHs and bacterial EHs, related to eukaryotic EHs), cluster II (bacterial EHs), and microsomal EHs: cluster III. Positions of the structurally known EHs are marked with arrows: *Mus musculus* (Mm), *Agrobacterium radiobacter* (Ar) and *Aspergillus niger* (An).

Three clusters ordered by superfamily and organisms were formed: cluster I (cEHs from mammals, plants and to eukaryotic EHs related bacterial EHs) comprises the structurally known EH from *Mus musculus*, cluster II (cEHs from bacteria) contains the structurally known EH from *Agrobacterium radiobacter*, and cluster III (mEHs) comprises the structurally known EH from *Aspergillus niger*.

Although EHs contain a highly conserved structure, they show low sequences identities, what complicates reliable homology modelling. Assuming that EHs of one cluster, and thus with similar loop-lengths possess a similar structure, the loop-based classification into clusters and the multisequence alignment of 93 EHs was used to perform homology models of five

representative EHs of cluster I-III (from *Glycine max*, *Mycobacterium tuberculosis*, *Corynebacterium sp.*, *Streptomyces coelicolor* and *Mesorhizobium loti*). Of 78% of all 93 EHs the complete structure can be predicted using the insights resulting from the systematic analysis. In contrast, only 14% of the EHs can be predicted using the accepted method of pairwise alignments.

Another application for the loop-based assignment into clusters, is the prediction of substrate specificity of EHs. Looking at the substrate spectra of EHs, it is remarkable that EHs with long cap-loops (cyt. and mic. mammalian and plant EHs) are able to convert epoxy fatty acids (22, 48, 52), while EHs containing long NC-loops (mic. mammalian, insect, and fungal EHs) accept epoxy polycyclic aromatic hydrocarbons as substrates (18, 34, 48). For EHs with short cap-loops the conversion of fatty acids is not described. The long cap-loop forms a hydrophobic tunnel starting at the proteins surface down to the nucleophile of the catalytic triad (Figure 3), which is not observable for EHs with short cap-loops. This hydrophobic tunnel leads to the formation of a L-shaped binding pocket (8) and is similar to the fatty acid binding site of lipases (112) and would provide space for the hydrophobic alkane chain of a fatty acid. EHs with short cap-loops, without tunnel, and L-shaped binding pocket should not be able to convert epoxy fatty acids, while the bacterial EHs of cluster I, related to eukaryotic EHs and containing long cap-loops, should show activity.

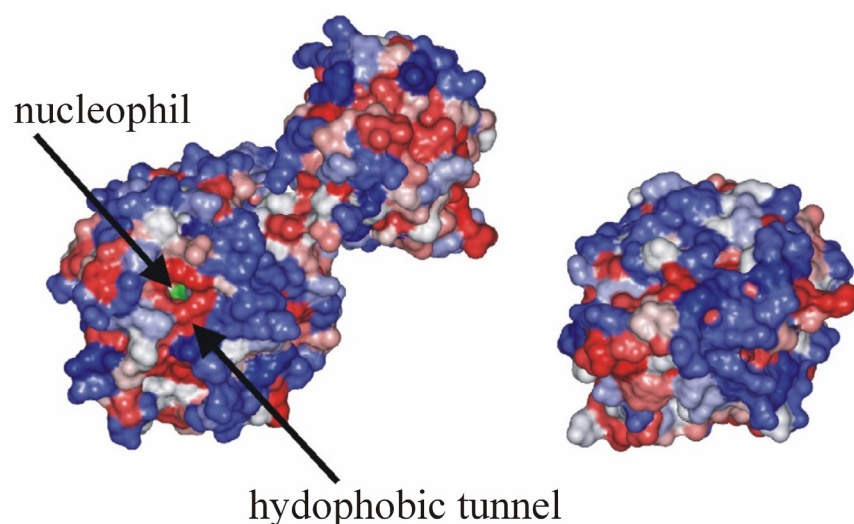


Figure 3

Structures of the EHs from *Mus musculus* (left) and *Agrobacterium radiobacter* (right) in surface design in topview on the cap domain. The structures are colored for hydrophobicity of their amino acids, hydrophilic residues are marked blue, hydrophobic residues are red. Hydrophobic tunnel and nucleophile (green) are labeled and marked with arrows.

The assignment to homologous families received from the phylogenetic analysis was employed for the design of family specific primers. Tests of these primers were performed with two bacterial organisms. *Streptomyces antibioticus* Tü4 shows EH activity towards styrene oxide (150). An already amplified 236 bp fragment of this species was located in the branch of the bacterial EHs of cluster I. With primers specific for this group, a 771 bp fragment with significant homology to EHs was amplified. After purification, N-terminal sequencing and tryptic digestion by the BASF AG, Ludwigshafen, it was possible to complement the fragment N-terminal. The C-terminus was constructed using PCR with the EH from *Streptomyces avermitilis* as template. Using the same set of primers, amplification of a 700 bp fragment from genomic DNA from *Rhodococcus ruber* (LU760), which belongs like the genus *Streptomyces* to the order of *Actinomycetales*, and identification as part of a putative EH was possible.

The second part of this work is concerned with building up and interpretation of a family specific database. The epoxide hydrolase/haloalkane dehalogenase database (EH/HD) is an extension to the 93 EH alignment with inclusion of further enzyme families and structure information. Building up the EH/HD database the automatic data extraction system of the Lipase Engineering Database (LED) (45) was used. A set of representative sequences, resulting from the 93 EHs of the systematic analysis and 18 HDs found by keyword search, was used for BLAST searches. The collected data was ordered hierarchically by the database system DWARF (46) in protein entry, homologous family and superfamily. The database consists of 397 sequence entries, 305 protein entries and contains for 19 protein entries 51 structure data sets. Besides EHs and HDs the database contains several other enzyme families like e.g. non-heme peroxidases, 2-hydroxy-muconic semialdehyde hydrolases, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid hydrolases, prolyl-amino-peptidases, esterases und luciferases, which were combined into 14 homologous families and two superfamilies.

To investigate similarities and differences of this huge amount of various α/β hydrolases of LED and EH/HD database, an analysis of available structural data was performed. While structures of the EH/HD database contain a canonical, a complete, α/β hydrolase fold, the hydrolases of the LED often show reductions. These are restricted to the β -strands $\beta 1$ and $\beta 2$, as well as the α -helices $\alpha 4$ and $\alpha 5$. $\beta 1$, $\beta 2$, $\alpha 4$ and $\alpha 5$ do not contain any catalytically relevant residues. Despite all significant structural differences the catalytic machinery is structurally conserved in all α/β hydrolases.

A so called anchor-concept (111) was developed for lipases, which is also valid for the structures of the EH/HD database. Despite the high conservation of the structures of the EH/HD database, stabilization of the oxyanion hole is realized differently, either with a hydrophobic pocket or a hydrophilic anchor. In case of prolyl-aminopeptidases no known anchor-type was detected.

The GXGXS-Motif (92) is highly conserved and forms a structurally also highly conserved loop. The GXGXS-Motif occurs only in the structures of the EH/HD database and is located beneath the first two β -strands of the α/β hydrolase fold and the cap-helices $\alpha 4$ and $\alpha 5$, which are connected by hydrogen-bonds (Figure 4).

The positions of the five residues involved in hydrogen bonding of the strands $\beta 1$ and $\beta 2$ and the GXGXS-motif are structurally conserved. Hydrogen bonding between GXGXS-Motif and cap-domain is more variable, but all residues involved are located at the loop between the cap-helices $\alpha 4$ and $\alpha 5$. The GXGXS-Motif stabilizes the strands $\beta 1$ and $\beta 2$ and is acting like a lock, connecting cap domain and N-terminal catalytic domain.

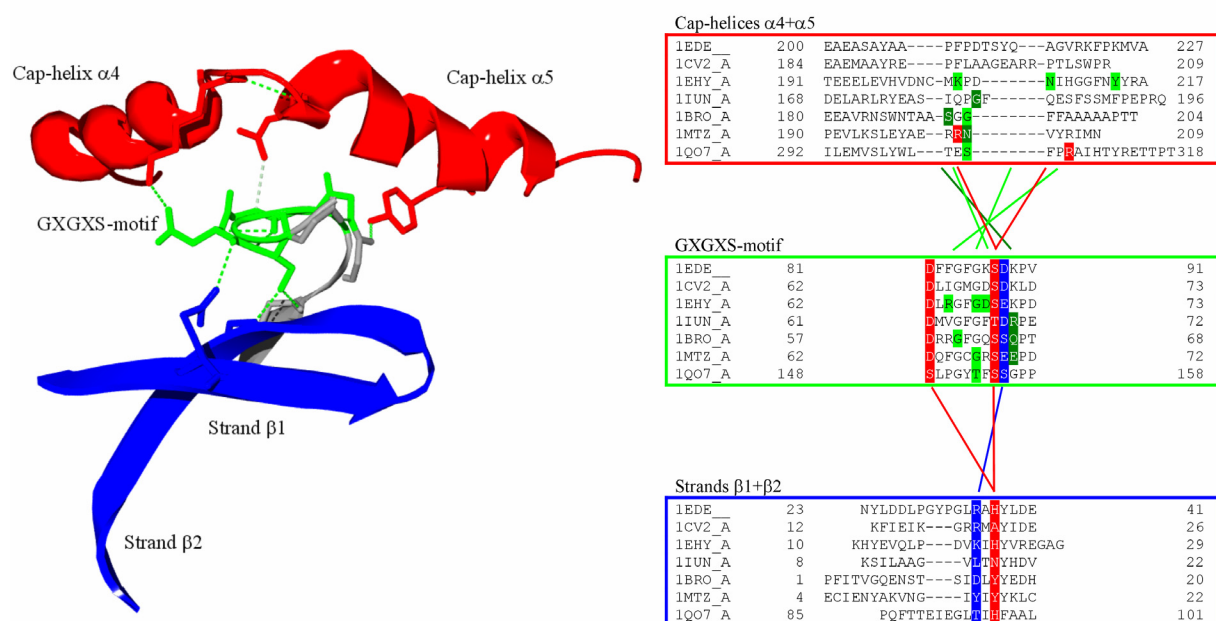


Figure 4

Left: The coin in the coin slot (epoxide hydrolase of *Agrobacterium radiobacter*, 1EHY): cap-helices $\alpha 4$ and $\alpha 5$ (red), GXGXS-motif (green), strands $\beta 1$ and $\beta 2$ (blue). Hydrogen bonds are shown as dashed lines. Right: Multisequence alignment of 1EDE (haloalkane dehalogenase), 1CV2_A (haloalkane dehalogenase), 1EHY_A (epoxide hydrolase), 1IUN_A (2-hydroxyruconic semialdehyde hydrolase), 1BRO_A (non-heme bromoperoxidase) 1MTZ (prolyl-aminopeptidases) and 1QO7_A (epoxide hydrolase). Frame colors are identical to coloring of the structure (left). In the alignment: residues connected by hydrogen bonds have the same color (hydrogen bonding indicated by lines).

By systematic analysis of sequence and structure of this large and diverse protein family, structurally and functionally relevant motifs and the modular architecture of α/β hydrolases were analysed, which provides a deeper insight into sequence-structure-function relationships. Understanding the modular structure allows reliable predictions of the biochemical properties for functional genomics and protein engineering.

3 Einleitung

3.1 Epoxidhydrolasen

3.1.1 Allgemeines

Epoxidhydrolasen (EHs, EC 3.3.2.3) gehören zu einer Gruppe hydrolytischer Enzyme, die Kofaktor unabhängig, enantio- und regiospezifisch die Addition von Wasser an Oxiran Derivate (Epoxide) unter der Bildung von vicinalen *trans*-Diolen katalysieren (10). Über eine S_N2 -spezifische Ringöffnung des Epoxids wird das *trans*-konfigurierte 1,2-Diol gebildet (Abbildung 1). EHs wurden in Organismen aller Arten gefunden, einschließlich Säugetieren (6), Insekten (34), Pflanzen (4), Pilzen und Bakterien (5). Aufgrund ihrer Primärsequenz wurden sie zur Struktur-Familie der α/β Hydrolasen gezählt (101). Die Enzymfamilie ist in zwei große Gruppen gegliedert: lösliche (cytosolische; cEHs) und membrangebundene EHs (mEHs) (48).

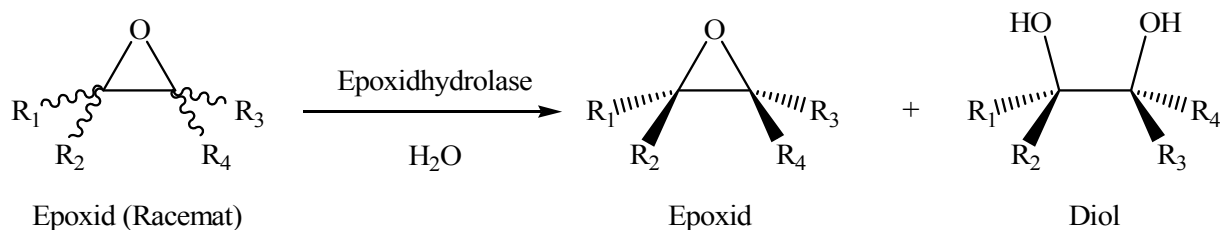


Abbildung 1

Hydrolyse von Epoxiden durch EHs

3.1.2 Säugetier EHs

Säugetiere enthalten sowohl cEHs als auch mEHs. Enzyme beider Gruppen wurden bezüglich ihrer biologischen Rolle detailliert untersucht (48). Weniger interessant ist, aufgrund ihrer schlechten Verfügbarkeit, ihr Einsatz in der organischen Synthese (7).

Die cEHs der Säugetiere sind Xenobiotica abbauende Enzyme, sie katalysieren die Hydrolyse *trans*-substituierter Epoxide, ebenso wie die vieler aliphatischer Epoxide des Fettsäure Metabolismus (48). Endogene Epoxide sind wichtige Effektormoleküle, die hormonähnliche Reaktionen in verschiedenen organischen Systemen hervorrufen können. Arachidonsäure z.B. wird zu einer Reihe von Monoepoxiden oxidiert, die dann zu den entsprechenden Dihydrodiolen hydrolysiert werden können, welche gefässerweiternde Eigenschaften haben

(78). cEHs sind hier also nicht ausschliesslich Entgiftungsenzyme, sondern spielen eine entscheidende Rolle bei z.B. der Regulation des Blutdrucks und entzündungshemmenden Prozessen.

Die mEHs der Säugetiere katalysieren die *trans*-Addition von Wasser an ein breites Spektrum von Epoxid-Substraten (100). Sie spielen eine wichtige Rolle bei der Entgiftung vieler reaktiver epoxidierte Intermediate des Stoffwechsels (10, 48), ebenfalls scheinen sie an der Aktivierung karzinogener polyzyklischer aromatischer Kohlenwasserstoffe beteiligt zu sein (48).

Mehrere Säugetier cEHs und mEHs wurden gereinigt und sequenziert: z.B. aus *Homo sapiens* (15, 125), *Mus musculus* (8), *Oryctolagus cuniculus* (60) und *Rattus norvegicus* (76).

3.1.3 Pflanzen EHs

Pflanzen enthalten cEHs (22) die epoxidierte Fettsäuren umsetzen. Die Expression findet häufig im Keimling statt (40, 42). Mehrere Funktionen wurden den Pflanzen EHs bisher zugeschrieben, wie eine Beteiligung an der Cutin-Biosynthese und der Abwehrreaktion gegen Feinde. Eine mögliche Funktion wäre eine Beteiligung an Stärkung und Reparatur von Schäden der Cuticula, verursacht durch Pflanzenpathogene (23). Ebenfalls vorgeschlagen wurde eine Beteiligung an der Biosynthese oxidierte Fettsäuren, wie z.B. dem Jasmonat. Jasmonat ist ein Signalmolekül, das Genexpressionen als Antwort auf Stress, im Besonderen auf mechanische Verwundung, induziert (83). Die Expression ist nicht in allen Pflanzen identisch: *Solanum tuberosum* accumuliert eine EH nach Verletzung oder Induktion durch Methyljasmonat(129). *Arabidopsis thaliana* exprimiert das Enzym als junge Pflanze nach Behandlung mit dem Wachstumshormon Auxin (74). *Nicotiana tabacum* exprimiert eine EH als Antwort auf eine Infektion mit dem Tabakmosaik Virus, zeigt jedoch keinerlei Reaktion auf eine Behandlung mit Auxin (55).

3.1.4 Insekten EHs

Insekten enthalten mEHs (139), die an der Regulation von Juvenil-Hormonen (JH) beteiligt sind. Die JH der Insekten sind epoxidierte Sesquiterpenester, die Wachstum und Entwicklung des Larvenstadiums und der sexuellen Reifung erwachsener Tiere kontrollieren (147). Die Kontrolle dieser JH wird durch eine JH Esterase und eine JH Epoxidhydrolase, moduliert durch ein JH Bindeprotein (113) ausgeführt. EHs wurden z.B. in den folgenden Insekten

beschrieben: *Ctenocephalides felis* (71), *Drosophila melanogaster* (72), *Galleria mellonella* (144), *Manduca sexta* (34) und *Trichoplusia ni* (139).

3.1.5 EHs aus Hefen und Pilzen

Pilze und Hefen enthalten mEHs (79, 151) die an Stoffwechselfunktionen beteiligt sind (18). Die erste EH aus einem Pilz wurde 1972 beschrieben (133), die EH aus *Helminthosporium sativum* katalysiert die enantioselektive Hydrolyse von 10,11-Epoxyfarnesol. Der Verursacher der Weißfäule *Pleurotus ostreatus* hat, wie viele andere holzzersetzende Pilze, die Fähigkeit polyzyklische aromatische Kohlenwasserstoffe wie Phenantren, Antrazen und Pyren abzubauen (18). Bei dem hierfür vorgeschlagenen Mechanismus z.B. für Phenantren, wird Phenantren durch ein Cytochrom P450 in Phenantren-9,10-Oxid umgesetzt, anschliessend wandelt eine Epoxidhydrolase dieses Epoxid in Phenantren *trans*-9,10-Dihydrodiol um. EH Aktivität wurde in vielen Pilzen und Hefen gefunden, z.B.: *Aspergillus niger* (151), *Beauveria sulfurens* (108), *Pleurotus ostreatus* (18), *Rhodotorula glutinis* (142), *Ulocladium atrum* (53) und *Xanthophyllomyces dendrorous* (141).

3.1.6 Bakterielle EHs

Bei einer Vielzahl von Bakterien wurde EH Aktivität nachgewiesen (7), häufig mit der Fähigkeit industriell wichtige Epoxide enantioselektiv umzusetzen. Viele enantiomerenreine Epoxide und vicinale Dirole sind Ausgangsstoffe für die pharmazeutische und agrochemische Industrie. Da bakterielle Organismen sich vergleichsweise einfach in großen Mengen kultivieren lassen, stellen sie eine interessante Quelle für den industriellen Einsatz dar. Glycidylphenylether ist ein nützliches Aryloxy Epoxid für die Synthese chiraler Aminoalkohole (67) und β -Blockern (75). Eine EH aus *Bacillus megaterium* ECU1001 setzt Glycidylphenylether mit hoher Enantioselektivität für (*R*)-Glycidylphenylether, durchführbar im präparativen Maßstab, um (51). Die Präferenz dieser EH ist komplementär zu bisher bekannten Enzymen (26, 87). (*S*)-Glycidylphenylether wird für die Synthese der medizinisch wichtigen β -Blocker verwendet (75). Bakterielle EHs wurde unter anderem in folgenden Organismen beschrieben: *Agrobacterium radiobacter* (116), *Corynebacterium* sp. (92), *Nocardia* sp. (80), *Pseudomonas* sp. (64), *Rhodococcus* sp. (130) und *Streptomyces antibioticus* (150).

3.1.7 Bekannte Strukturen

Bis heute wurden drei Epoxidhydrolase Strukturen aufgelöst:

- 1) Die Struktur einer bakteriellen cEH aus *Agrobacterium radiobacter* AD1 (pdb-Eintrag 1EHY) wurde 1999 kristallisiert (96). *Agrobacterium radiobacter* AD1 ist ein Gram-negatives Bakterium, das aufgrund seiner Fähigkeit Epichlorhydrin als einzige Kohlenstoff- und Energiequelle zu nutzen, isoliert wurde (138).
- 2) Als nächstes wurde die Struktur einer cEH aus der Hausmaus *Mus musculus* (pdb-Eintrag 1CQZ), ebenfalls 1999, veröffentlicht (8). Die Struktur besteht aus zwei großen Domänen, einer Epoxidhydrolase Domäne und einer N-terminalen Domäne, welche strukturelle Ähnlichkeiten zu Haloacetat Dehalogenasen aufweist. Kürzlich wurde für diese Domäne eine Phosphataseaktivität nachgewiesen, die unabhängig von der EH-Aktivität zu sein scheint (30, 98).
- 3) Die erste Struktur einer mEH aus dem Pilz *Aspergillus niger* (pdb-Eintrag 1QO7) wurde 2000 veröffentlicht (151). Die EH aus *Aspergillus niger* stellt einen Sonderfall eines mikrosomalen Proteins dar, sie ist ein „lösliches“ mikrosomales Protein, der bei eukaryotischen EHs übliche Membrananker fehlt ihr.

3.1.8 Epoxide

Epoxide (Oxirane) sind dreigliedrige, heterozyklische Verbindungen, die ein Sauerstoffatom enthalten. Diese Verbindungen lassen sich auch als organische Oxide oder zyklische Ether auffassen. Aufgrund der hohen Spannung im Oxiranring-System und der starken Polarisierung der C-O Bindungen sind Epoxide reaktive Elektrophile. Sie sind chirale Verbindungen und als optisch reine Epoxide wichtige Bausteine in der organischen Synthese. Für die Darstellung chiraler Epoxide existieren unterschiedliche chemische Verfahren wie z.B.: die Darstellung ausgehend von prochiralen Verbindungen (109), die Sharpless-Oxidation (69) und das Verfahren nach Jacobsen und Katsuki (63). Die Racematspaltung mit EHs zur Herstellung enantiomerenreiner Epoxide stellt eine interessante Alternative zu den aufwendigen chemischen Verfahren dar (7, 41, 103, 109, 130, 134).

3.1.9 Reaktionsmechanismus

EHS enthalten eine katalytische Triade, gebildet aus einem Nukleophil und einem His/Säure-Paar(62). Zusätzlich enthalten sie mindestens ein, für die Aktivität essentielles, Tyr (118), das über Wasserstoff-Brücken die Positionierung und Aktivierung des Substrats bewirkt. Im Falle der EHS wird das Nukleophil, welches sich in einem „nucleophilic elbow“ nach dem β -Strand β_5 befindet, von einem Asp gebildet, aber auch Ser oder Cys sind in anderen Enzymklassen möglich. Das Nukleophil führt den Angriff auf das primäre Kohlenstoffatom des Oxiranringes und somit die Ringöffnung durch (Abbildung 2). Das hierdurch gebildete, kovalent an das Nukleophil gebundene Esterintermediat wird durch das His/Säure-Paar hydrolysiert und das Produkt wird freigesetzt. Der Säurerest, ein Asp oder Glu, nach β_6 oder β_7 positioniert, assistiert dem His (nach β_8) bei der Aktivierung eines Wassermoleküls. Dieses hydrolysiert schliesslich das Esterintermediat an der Carbonylfunktion des Nukleophils. Während der Reaktion muß die negative Ladung, die sich an der nukleophilen Seitenkette bildet, stabilisiert werden. Dies wird durch das Oxyanionhole (137) gewährleistet. Das Oxyanionhole besteht aus zwei Backbone Amidinen, deren korrekte Ausrichtung für die Reaktion essentiell ist. Während eine der Oxyanionhole AS in der Sequenz direkt auf das Nukleophil folgt und somit stabil ist, liegt die andere AS im sog. HGX-Motiv (92) auf einem Loop. Die Stabilisierung dieser Loop-Region erfolgt über den sog. Anker (111), der mit der Seitenkette der Oxyanionhole AS wechselwirkt.

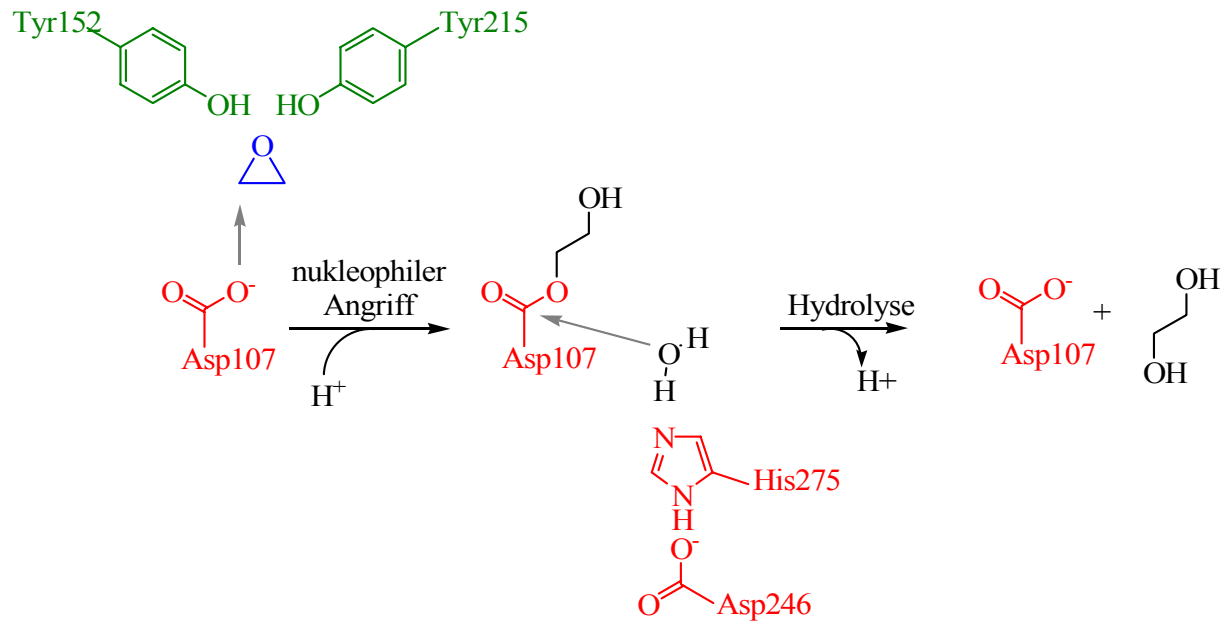


Abbildung 2

Reaktionsmechanismus von EHs am Beispiel der EH aus *Agrobacterium radiobacter*. Tyr152 und Tyr215 (grün) positionieren das Epoxid (blau) durch die Bildung von Wasserstoff Brücken zum Epoxid Sauerstoff und aktivieren das Epoxid für den nukleophilen Angriff durch Asp107 (rot). Der nukleophile Angriff von Asp107 erfolgt auf den Epoxid-Kohlenstoff und führt zu einem Enzym-Substrat Esterintermediat. Die Hydrolyse wird durch ein aktiviertes Wassermolekül durchgeführt. Die Aktivierung erfolgt über das Elektronen-Transfer-System des His275/Asp246 Paares (rot) und führt zur Freisetzung des Produkts und der Wiederherstellung des aktiven Enzyms.

3.2 α/β Hydrolase Fold

Anfang der 1990er Jahre wurden die 3D Strukturen von fünf scheinbar unterschiedlichen hydrolytischen Enzymen aufgedeckt. Die auffällige strukturelle Gemeinsamkeit dieser Strukturen wurde „ α/β Hydrolase Fold“ genannt (101). Die Familie der α/β Hydrolasen ist eine der größten Gruppen strukturell verwandter Enzyme mit unterschiedlichen katalytischen Funktionen. Das Modell der divergenten Evolution beschreibt die Entwicklung von Proteinen mit unterschiedlicher Funktion ausgehend von einem gemeinsamen Vorfahren. Im Laufe der Evolution veränderte sich ein Vorläuferprotein durch den Austausch von Aminosäuren (AS) und das Einfügen oder Entfernen von Loop-Regionen. Dadurch könnten sich Enzyme mit unterschiedlichen katalytischen Funktionen entwickelt haben, welche nur noch eine geringe Sequenzähnlichkeit aufweisen. Die Familie der α/β Hydrolasen repräsentiert ein solches Beispiel (62). Mitglieder dieser Familie sind beispielsweise: Acetylcholinesterasen, Lipasen, Thioesterasen, Haloalkan Dehalogenasen, Non-heme Peroxidasen, EHs und Serincarboxypeptidasen.

Der kanonische α/β Hydrolase Fold enthält ein zentrales β -Faltblatt, welches aus acht β -Strands, sieben parallelen und einem antiparallelen (β_2), und sechs α -Helices besteht. Die erste und die letzte der α -Helices (α_1 und α_6) befinden sich auf einer Seite des β -Faltblattes, die restlichen vier Helices auf der entgegengesetzten Seite (Abbildung 3). Viele Mitglieder dieser Enzymfamilie enthalten Loop-Insertionen, die sich zu Subdomänen falten. Diese Sekundärstrukturelemente bilden sogenannte „Caps“, „Lids“ oder „Flaps“, welche wichtige Rollen bei der Bildung von Form und Erreichbarkeit der Bindungstaschen der unterschiedlichen Enzyme spielen (101). Zusätzlich zur Hydrolase Domäne enthält z.B. die Epoxidhydrolase Struktur eine Cap Region. Das Cap, bestehend aus fünf α -Helices, ist in den α/β Hydrolase Fold nach dem β -Sheet β_6 insertiert und spielt eine wichtige Rolle bei Substratbindung und Regioselektivität (62).

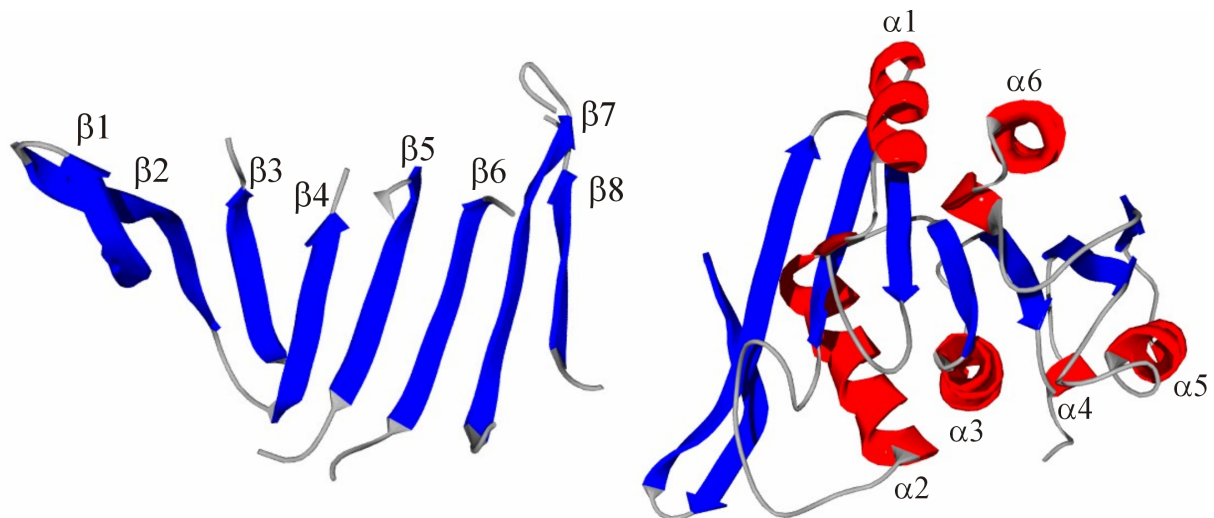


Abbildung 3

Darstellung des α/β Hydrolase Folds am Beispiel der Non-heme Peroxidase aus *Streptomyces aurefaciens* (1BRO). Der α/β Hydrolase Fold besteht aus einem zentralen β -Faltblatt (links, blau), bestehend aus einem antiparallelen und sieben parallelen β -Strands. Das zentrale β -Faltblatt wird von sechs α -Helices flankiert (rechts, rot).

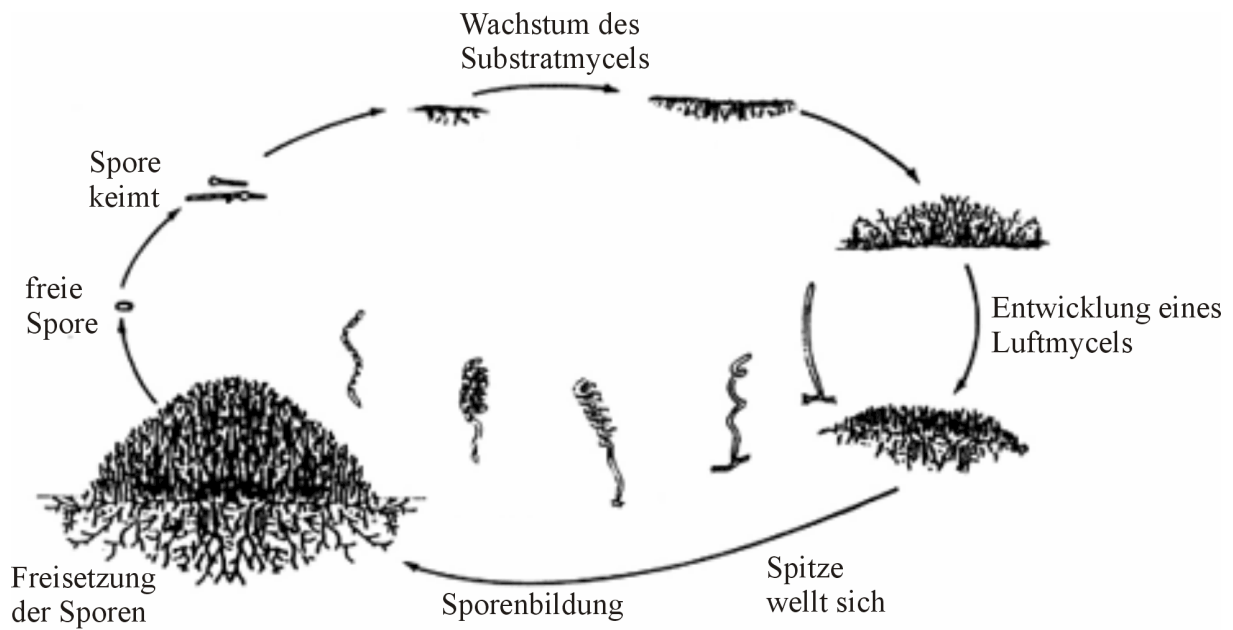
3.2.1 Haloalkan Dehalogenasen

EHS weisen neben der strukturellen Übereinstimmung, eine deutliche Ähnlichkeit auf Sequenzebene zu Haloalkan Dehalogenasen (HD, E.C. 3.8.1.5) auf (116). HDs sind bakterielle Enzyme, die Kohlenstoff-Wasserstoff Bindungen halogenerter aliphatischer Verbindungen durch Hydrolyse spalten (62). 1985 wurde die erste HD aus *Xanthobacter autrophicus* veröffentlicht (65). Die Enzyme setzen halogenierte Alkane in die entsprechenden Alkohole um. Die Substratspezifitäten von HDs sind im Regelfall breit, die Enzyme akzeptieren verschiedene Chloro- und Bromoalkane mit einer Präferenz für primäre Kohlenstoff-Halogen Bindungen. Die Enzyme wurden aufgrund von Substratspezifität und Sequenzhomologie in zwei Gruppen unterteilt (31, 95): HDs die C2-C4 Haloalkane umsetzen und HDs die Mono-, Di- und Mehrfach halogenierte C2-C4 Alkane und zyklische Substrate akzeptieren.

3.3 Streptomyceten

Die Gattung *Streptomyces* gehört zur Ordnung der *Actinomycetales*. Actinomyceten sind eine große Gruppe filamentöser, gram-positiver Bakterien, deren DNA einen GC-Gehalt von 63-78% aufweist. Diese faserartig wachsenden Bakterien bilden ein verzweigtes Netzwerk, Mycel genannt, welches Analogien zu den Mycelien filamentös wachsender Pilze aufweist. Die meisten Actinomyceten sind Sporenbildner und weisen einen komplexen Differenzierungszyklus auf (Abbildung 4). Die Hyphen dieses Mycels haben einen Durchmesser von 0,5-1,0 µm und sind von unbegrenzter Länge. In der vegetativen Phase fehlen den ihnen häufig die Querwände. Das Wachstum erfolgt an den Spitzen der Hyphen und ist von Teilungen begleitet, so dass sich ein dichtes, verwobenes Mycel bildet. Altert eine *Streptomyces*-Kultur, bilden sich Lufthyphen sogenannte Sporophoren, aus welchen sich Sporen, Conidia, abschnüren. *Streptomyces*-Sporen werden durch die Bildung von Querwänden innerhalb der mehrkernigen Sporophoren gebildet.

Die Gattung *Streptomyces* wird durch über 500 unterschiedliche Species vertreten, welche überwiegend Bodenorganismen sind. Ihr charakteristischer Erdgeruch wird durch mehrere Metabolite, Geosime genannt, verursacht. Diese Substanzen sind Sesquiterpenoide, ungesättigte Ringverbindungen von Kohlenstoff, Sauerstoff und Wasserstoff. Ein typisches Geosim ist *trans*-1,10-dimethyl-*trans*-9-decalol. Von besonderer Bedeutung ist die Fähigkeit von vieler der *Streptomyces* Species eine große Anzahl therapeutisch wichtiger Antibiotika zu bilden. Über 500 unterschiedliche, antibiotisch wirksame Substanzen wurden aus *Streptomyces* isoliert. Viele Species produzieren mehrere Antibiotika, die, obwohl in einem Organismus produziert, chemisch nicht zwingend verwandt sind. Beispiele für *Streptomyces* Antibiotika sind: Streptomycin (*Streptomyces griseus*), Chloromycetin (*Streptomyces venezuelae*) und verschiedene Tetracycline (*Streptomyces aureofaciens*) (88). Für *Streptomyces antibioticus* Tü4 wurde eine EH-Aktivität gegen Styroloxid nachgewiesen (150).

**Abbildung 4**

Lebenszyklus von *Streptomyces*. Nach dem Keimen, bildet *Streptomyces* ein Substratmycel und anschließend ein Luftmycel (Sporophoren), mit zunehmendem Alter beginnt sich die Spitze eines Sporophoren zu wellen und die Sporen werden freigesetzt.

3.4 Phylogenie

Phylogenie/Phylogenese:

[Griechisch: *phylos* – Stamm; Griechisch: *genesis* – Entstehung]: Stammesgeschichtliche Entwicklung von Organismen (Begriff 1866 von E. Haeckel geprägt) (143).

Bis in das 18. Jahrhundert hinein galt die „Konstanz der Arten“ als unumstößliche Lehrmeinung. Auch der schwedische Naturforscher Carl von Linné (1707-1778), der über ein umfassendes naturwissenschaftliches Wissen verfügte, glaubte an diese Theorie. Allerdings kamen ihm bei der Aufstellung seines Systems der Pflanzen und Tiere Zweifel, als er "fließende Übergänge" zwischen einzelnen Arten feststellte. Der französische Zoologe Georges de Cuvier (1769-1832) kannte zwar die Fossilien aus den Gesteinsschichten im Pariser Becken, hielt aber dennoch an der Konstanz der Arten fest. Er deutete die fossile Artenvielfalt als Abfolge von weltumspannenden Katastrophen (Sindfluten). An der Wende vom 18. zum 19. Jahrhundert wurde versucht, Faktoren für die offensichtliche Auseinanderentwicklung der Lebewesen zu finden. Jean-Baptiste de Lamarck (1744-1829) äußerte sich in seiner "Philosophie zoologique" so: "Alles, was ein Lebewesen durch den Einfluss anhaltender Daseinsbedingungen gewinnt oder verliert, wird durch Vererbung erhalten und geht auf die kommende Generation über." Lamarck war der Auffassung, dass im Laufe eines individuellen Lebens erworbene Eigenschaften an die Nachkommen vererbt werden können. Diese Eigenschaften können durch eine Anpassung, wie sie von einer sich ändernden Umwelt ausgelöst wird, erworben werden, aber auch durch ein inneres Bestreben, einer bestimmten Umwelt gerecht zu werden. Durch Gebrauch oder Nichtgebrauch eines Organs kann dieses also entweder leistungsfähiger werden oder verkümmern. So ließen sich rudimentäre Organe erklären, aber auch die Sprungkraft des Tigers und der lange Hals der Giraffe.

In der Mitte des 19. Jahrhunderts gelang englischen Naturforschern der entscheidende Durchbruch. Ihr bedeutendster Vertreter war Charles Darwin (1809-1882). Von ihm erschien 1859 das Buch "Origin of Species by Means of Natural Selection", in dem Darwin Fakten für eine Abstammung der Lebewesen von ursprünglicheren Formen vorlegte, und auch auf Faktoren des Evolutionsgeschehens hinwies. Seine wesentlichen Aussagen waren:

- 1) In der Natur findet eine Überproduktion von Nachkommen statt. Trotzdem bleibt die Größe einer Population innerhalb gewisser Grenzen konstant.
- 2) Die Mitglieder einer Population unterscheiden sich voneinander (Variabilität).
- 3) Das am besten an seine Umwelt angepasste Individuum hat beim Kampf ums Dasein

die besten Überlebenschancen.

Darwin verwendete anstelle von 'Evolution' den Begriff 'descent with modification', also in etwa 'Abstammung mit Veränderung'. Wenn diese zuträfe, sollte es möglich sein, Stammbäume aufzustellen. Ernst Haeckel (1834-1919) vertrat die Thesen Darwins in Deutschland und entwickelte erste konkrete Rekonstruktionen der Stammesgeschichte von Tieren und Pflanzen. Er prägte den Begriff der *Phylogenie* für die Abstammungsgeschichte. Die Phylogenetik widmet sich der Rekonstruktion der evolutionären Verwandtschaftsbeziehungen von Arten und Artengruppen.

Haeckel (1894)

„Natürlich bleiben aber auch diese Schemata (Stammbäume) ... immer nur Versuche, tiefer in die Geheimnisse der Stammesgeschichte einzudringen; sie sollen nur den Weg andeuten, auf welchem nach dem jetzigen beschränkten Zustand unserer Kenntnisse die weitere phylogenetische Forschung am besten vorzudringen hat. Ich brauche daher hier wohl nicht zu wiederholen, dass ich meinen Entwürfen von Stammbäumen und systematischen Tabellen keinen dogmatischen Wert beimesse.“

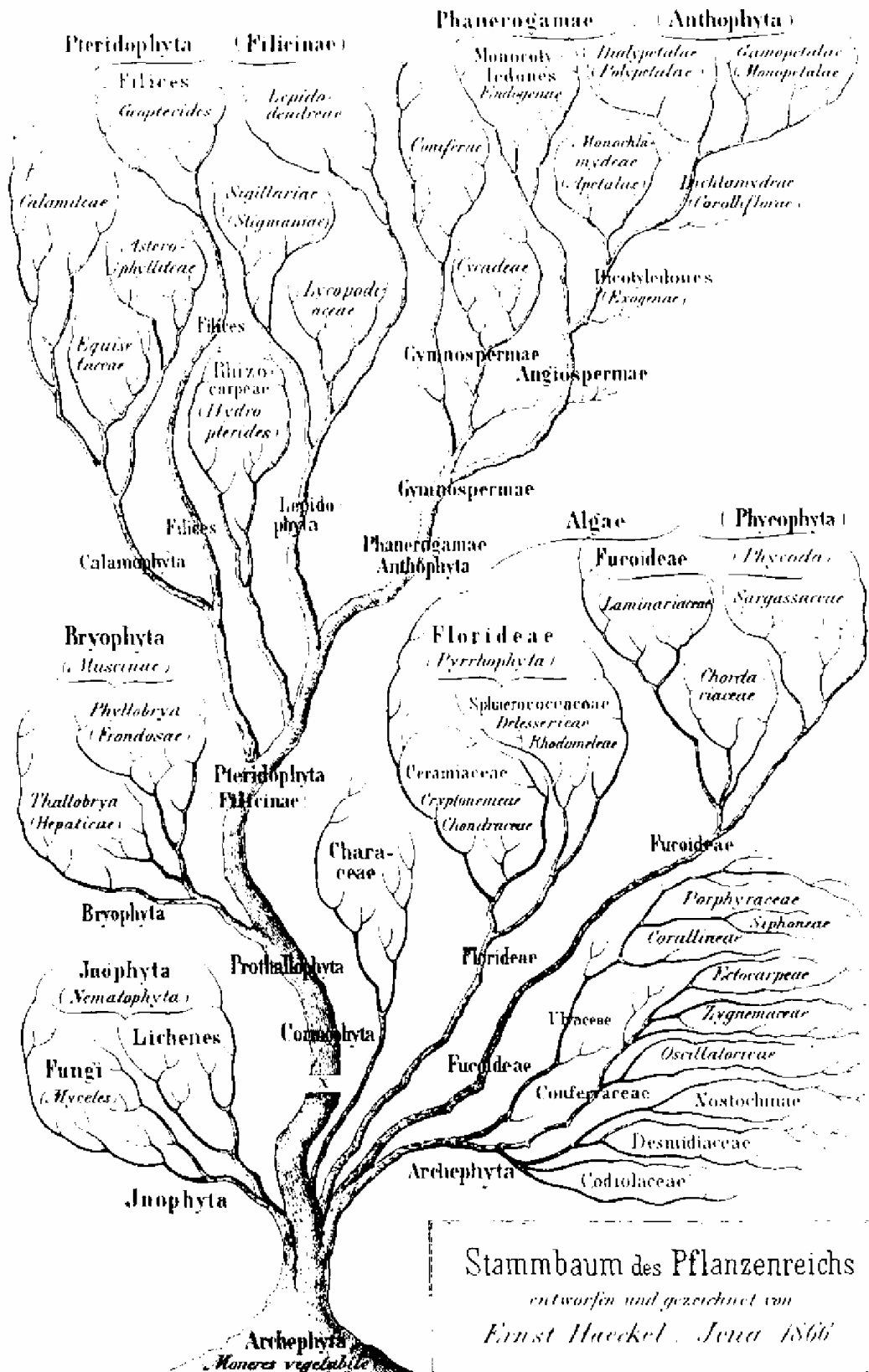


Abbildung 5

Stammbaum des Pflanzenreichs, Ernst Haeckel, 1866 aus seinem Buch „Generelle Morphologie der Organismen“.

3.4.1 Erstellung von Stammbäumen

Bis Mitte des 20. Jahrhunderts beruhten biologische Klassifizierungssysteme auf sichtbaren morphologischen Merkmalen. Ernst Haeckel teilte, ausgehend von Darwins Theorie, alle Lebewesen in die drei Reiche der Tiere, Pflanzen und Protisten auf. Diese fasste er unter Annahme der damals bekannten Gattungen und Arten von Pflanzen und Tieren unter dem Gesichtspunkt ihrer möglichen Entwicklung in einem phylogenetischen Baum zusammen (E. Haeckel, *Generelle Morphologie der Organismen*, 1866). Besonders schwierig war damals die Klassifizierung der Einzeller, da diese arm an auswertbaren sichtbaren Merkmalen sind. Daher ordnete Haeckel dem Reich der Protisten nicht nur einzellige Eukaryoten, sondern auch die als "Moneres" bezeichneten Organismen, die Prokaryoten zu.

Versuche einer natürlichen, d. h. in der Abstammung begründeten Klassifikation einzelliger Eukaryoten und Prokaryoten sind bis in die Gegenwart problematisch. Der Grund hierfür ist der gleiche, mit dem sich auch Haeckel konfrontiert sah, das Fehlen abgestufter Merkmalsreihen.

Ein möglicher Ausweg aus der Problematik der Klassifizierung der Einzeller wurde von Zuckerkandl und Pauling (152) geliefert. Sie schlugen vor, anstatt morphologischer Merkmale die Sequenzen von Proteinen dazu zu nutzen, vergleichende Studien zur Evolution durchzuführen. Als Beispiel sollte das Cytochrom c dienen, daher wurden zwischen 1970 und 1980 sehr viele Sequenzen für Cytochrome c aus Prokaryoten ermittelt. Als Problem stellte sich jedoch der laterale Gentransfer für diese Gene unterhalb verschiedener Bakteriengruppen heraus (37). Daher setzte sich als Standardmolekül für die Analyse phylogenetischer Beziehungen innerhalb der Gruppe der Einzeller die 16s RNA, eine kleine Untereinheit der ribosomalen RNA, durch.

Durch Untersuchungen mit 16s RNA wurde die Gruppe der Archaeobakterien entdeckt (146). Archaeobakterien umfassen alle Methanbakterien, sowie alle extrem halophilen und extrem thermoacidophilen Arten. Sie unterscheiden sich phylogenetisch und bezüglich vieler Besonderheiten wesentlich von den restlichen Prokaryoten, den Eubakterien. Archaeobakterien fehlt z.B. der Murein-Sacculus der Zellwand (145).

Der Vorschlag von Zuckerkandl und Paulus Proteinsequenzen für phylogenetische Untersuchungen einzusetzen (152) führte einen Paradigmenwechsel herbei. Anstatt des Phaenotyps wurde zukünftig der Informationsgehalt des Genotyps von Organismen zur Klassifizierung verwendet.

3.4.2 Die molekulare Uhr

Das Auftreten von Mutationen ähnelt dem Zerfall radioaktiver Elemente. Über kurze Zeiträume ist die Häufigkeit von Einzelereignissen nicht oder nur sehr ungenau vorhersehbar, über lange Zeiträume jedoch sind die Vorhersagen recht genau. Aus dieser gleichmäßigen Substitutionsrate leitet sich der Wunsch ab, bei phylogenetischen Analysen die Sequenzen als molekulare Uhren zu verwenden, um Aussagen über die Zeit zu machen, die zwischen den einzelnen Verzweigungen im phylogenetischen Baum vergangen ist.

Untersuchungen an z.B. an Vertebraten haben allerdings ergeben, dass sich die Mutationsraten verschiedener Organismen stark unterscheiden können. Dieses Phänomen könnte seinen Ursprung in den unterschiedlichen Generationszeiten mehrzelliger Organismen haben. Erklärt werden kann dies dadurch, dass für phylogenetische Analysen wichtige Mutationsereignisse in den Keimbahnen stattfinden. Haben nun zwei Organismen eine vergleichbare Anzahl an DNA-Replikationen in der Keimbahn, so ist die Mutationsrate in dem Organismus mit kürzerer Generationsdauer höher, da hier eine größere Anzahl an Zellteilungen pro Zeiteinheit stattfinden und somit mehr Mutationen entstehen können. Eine weitere Erklärung wäre, dass unterschiedliche Organismen unterschiedlich gut funktionierende Reparatursysteme enthalten, welche unterschiedlich hohe Mutationsraten ermöglichen (49, 77, 84). Obwohl also keine globale molekulare Uhr existiert, ist es möglich, lokal, d.h. bei eng verwandten Organismen, anhand molekularer Uhren Zeitabstände zu schätzen. So wird z.B. für die Analyse der zeitliche Entwicklung der Affen und Menschen eine molekulare Uhr verwendet, die mit dem Wissen um den Zeitpunkt der Divergenz einer Außengruppe aus fossilen Daten kalibriert wird (58, 84).

3.4.3 Mathematische Algorithmen zur Erstellung phylogenetischer Bäume

3.4.3.1 Datengrundlage

Um phylogenetische Bäume rekonstruieren zu können, werden Daten benötigt, die die stammesgeschichtliche Entwicklung widerspiegeln. Früher wurden zu diesem Zweck hauptsächlich morphologische Daten verwendet, heute werden hauptsächlich Sequenzdaten (AS-, DNA- oder RNA-Sequenzen) als Datenquelle genutzt.

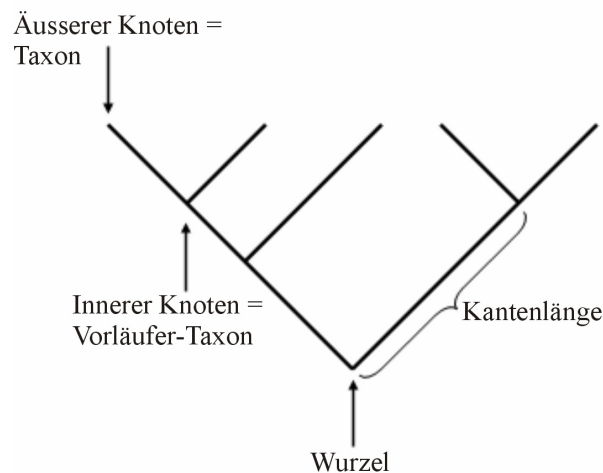


Abbildung 6

Schematische Darstellung eines binären phylogenetischen Baums.

Zur Visualisierung phylogenetischer Analysen werden binäre Bäume verwendet (Abbildung 6). Solche Bäume sind azyklisch verbundene, ungerichtete Graphen aus einer Menge von Knoten, die mittels Kanten miteinander verbunden sind. Jeder Knoten ist nie mit mehr als drei weiteren Knoten verbunden, d.h. er hat einen Grad von $d \leq 3$. Diese inneren Knoten repräsentieren ausgestorbene Vertreter, gemeinsame Vorfahren der rezenten Organismen, die die Blätter (Taxon) bilden ($d = 1$). Enthält der phylogenetische Baum eine Wurzel, so hat diese den Grad $d = 2$. Dieser Vorfahr beziehungsweise eine entsprechende DNA- oder Protein-Sequenz ist jedoch oft unbekannt. In diesen Fällen sind die Bäume wurzellos, die aufgestellten Baumdiagramme repräsentieren dann die Verwandtschaftsverhältnisse aller Taxa untereinander, nicht aber die evolutionären Wege ihrer Entstehung.

3.4.3.2 Distanzbasierte Methoden

Distanzbasierte Methoden verwandeln ein Sequenz-Alignment in eine paarweise Distanzmatrix und verwenden diese Werte zur Berechnung des phylogentischen Baums.

3.4.3.2.1 UPGMA

UPGMA (unweighted pair group method using arithmetic averages) (91) ist die einfachste Cluster-Methode. Zu Beginn bildet jede Sequenz ein eigenes Cluster und der zugehörige Zweig des Baumes besitzt die Länge 0. Im nächsten Schritt werden die beiden Cluster mit der geringsten evolutiven Distanz zu einem Cluster (Englisch: *Cluster-Anhäufung*) vereinigt. Ein Verzweigungspunkt, der die beiden Zweige verbindet wird definiert. Die Kantenlänge des Zweiges reflektiert die Distanz der beiden Taxa.

3.4.3.2.2 Neighbor-Joining

Die Kantenlängen der Äste eines phylogenetischen Baums sind additiv, d.h. die Kantenlängen des direkten Wegs zwischen den Taxa werden addiert, um den Abstand zweier Taxa von einander zu bestimmen. Der Neighbor-Joining Algorithmus (121) sucht nicht nur nach der minimalsten Distanz zweier Sequenzen, sondern subtrahiert die durchschnittlichen Distanzen der beiden Sequenzen zu den restlichen Sequenzen des phylogenetischen Baums. Hiermit wird verhindert, dass ein Paar von Nachbarn unterschiedlich lange Kantenlängen aufweist, und somit die Gefahr besteht, dass das Taxon mit der kürzeren Kantenlänge, aufgrund der Additivität, näher an einem anderen Taxa, als dem eigentlich nächsten Nachbarn liegt.

3.4.3.3 Merkmalsbasierte Methoden

Merkmalsbasierte Methoden verwenden direkt das Multisequenz-Alignment zu Erstellung des phylogenetischen Baums, indem sie die Variationsmöglichkeiten jeder Position des Alignments einbeziehen.

3.4.3.3.1 Maximum-Parsimony

Die Maximum-Parsimony-Methode (39) (Englisch: *parsimony*-Geiz) beruht auf dem Ökonomieprinzip, formuliert von dem mittelalterlichen Philosophen Ockham, das besagt, dass die einfachste Erklärung wahrscheinlich die richtige ist („It is futile to do with more what can be done with fewer“). Die Maximum-Parsimony-Methode sucht aus einer Menge möglicher Bäume den Baum heraus, der mit der geringsten Anzahl an Mutationen die beobachteten Unterschiede zwischen den Sequenzen erklären kann.

3.4.3.3.2 Maximum-Likelihood

Die Maximum-Likelihood-Methode (43) (Englisch: *Likelihood*-Wahrscheinlichkeit) berechnet ebenfalls jeden möglichen phylogenetischen Baum einer gegebenen Auswahl an Sequenzen. Die Maximum-Likelihood-Methode ist wahrscheinlichkeitstheoretisch, d.h. jeder möglichen ausgetauschten Position in den Sequenzen werden Wahrscheinlichkeiten zugeordnet. Der optimale phylogenetische Baum besitzt eine maximale Gesamtwahrscheinlichkeit. Maximum-Likelihood-Methoden verwenden hierfür Informationen über Substitutionsraten, analog zu den Substitutionsmatrizen, die in MSAs Verwendung finden.

3.5 Datenbanken

Seit dem Aufkommen des World Wide Web sind biologische Datenbanken zu einer wichtigen Ergänzung der Literaturrecherche geworden. Informationen in einer zentralen biologischen Datenbank zu finden und herunterzuladen gehört inzwischen genauso zu den Kenntnissen die man als Wissenschaftler besitzen muß, wie die traditionelle Literaturrecherche (50).

3.5.1 Datenbanktypen

Es gibt zwei Typen von Datenbank-Management-Systemen (DBMS): Systeme, die eine indizierte Datei benutzen, ein sogenanntes Flat-File, und relationale DBMS. Ein dritter Typ, das objektorientierte DBMS, wird immer populärer. Die Wahl zwischen einem Flatfile-System und einem relationalen Datenbank-System ist eine wichtige Entscheidung, die weitreichende Auswirkungen auf das Fassungsvermögen und die Benutzbarkeit einer Datenbank hat.

3.5.1.1 Flatfile-Datenbanken

Flatfile-Datenbanken sind die einfachsten Datenbanktypen, sie bestehen aus einer sortierten Ansammlung ähnlicher Dateien. Dieser Datenbanktyp wird häufig durch eine Indizierung mit einer Suchfunktion ausgestattet. Für einen Index werden bestimmte Merkmale aus einer Datei herausgeholt und mit der Datei und dem Speicherort verknüpft. Flatfile-Datenbanken können nur begrenzte Mengen an Daten verwalten, steigt die Datenmenge stetig an, wird dieser Datenbanktyp immer ineffizienter. Der eindimensionale Aufbau des Index erschwert die Handhabung großer Datenmengen. Viele biologische Datenbanken haben ihren Ursprung in einer Flatfile-Datenbank. Ein Beispiel ist die PDB-Datenbank (17), die bis 1990, ca. 20 Jahre nach ihrer Entstehung, auf nicht mehr als ein paar hundert Dateien angewachsen war. Als die PDB in den 90er Jahren an Größe zunahm mußten neue Lösungen gefunden werden. Heute ist die PDB eine objektorientierte Datenbank.

3.5.1.2 Relationale Datenbanken

In relationalen Datenbanken werden Informationen in einer Ansammlung von Tabellen abgelegt. In diesem Datenbanktyp werden die Informationen eines jeden Kapitels in getrennten Tabellen gespeichert, die miteinander verknüpft werden können. Es gibt z.B. Tabellen für Versuchsbedingungen, die Elemente der Sekundärstruktur, die Positionen der

Atome.... Ein relationales Datenbank-Management-System (RDBMS) erlaubt es z.B. alle Proteinstrukturdaten auf einmal zu lesen. Die Daten in den Tabellen relationaler Datenbanken stehen jeweils in Zeilen, wobei jede Zeile einem Datensatz entspricht. Jede Zeile kann unterschiedliche Felder mit Informationen enthalten. Jedes Feld innerhalb der Datenbank muß eine eindeutige Information enthalten. Es kann nicht aus einer Liste oder einem weiteren Satz an Information bestehen, der wiederum unterteilt werden kann. Die Tabellen einer relationalen Datenbank sind nicht nur verbesserte Flatfiles. Die Aufgabe des RDBMS besteht darin, Verbindungen zwischen verwandten Tabellen zu erstellen, indem diejenigen Elemente aufgespürt werden, die in anderen Tabellen vorhanden sind.

Das Netzwerk aus Tabellen und ihre Beziehungen zueinander, wird Datenbankschema genannt. Um den Nutzen einer Datenbank über eine lange Zeit hinweg zu gewährleisten, sollte das Schema einer Datenbank sorgfältig entwickelt werden.

3.5.1.3 Objektorientierte Datenbanken

Der Begriff „objektorientiert“ wird sowohl bei Programmiersprachen, als auch bei Datenbanken verwendet. Ein DBMS ist dann objektorientiert, wenn es die gleichen Grundprinzipien verwendet, wie die objektorientierte Programmierung. Der praktische Nutzen objektorientierter Datenbank-Systeme ist ihre Flexibilität, mehr als nur Tabellen zu speichern. Da Informationen nicht als Tabellen, sondern Objekte gespeichert werden, kann solch eine Datenbank den Zugriff auf alles Mögliche erlauben, von der einfachen Textdatei, bis hin zur Videosequenz.

Viele Datenbank-Systeme werden heute so programmiert, dass sie sowohl objektorientierte als auch relationale Ansätze verfolgen.

4 Ergebnisse

4.1 Zielsetzung

Ziel dieser Arbeit war die systematische Analyse von EHs. Die bei dieser Analyse gewonnenen Erkenntnisse sollten den Verwandtschaftsgrad zwischen den einzelnen EHs unterschiedlicher Herkunft aufzeigen und den Weg zur Erstellung familienspezifischer Primer ebnet, um das Gen zu identifizieren, das für die bekannte EH Aktivität gegen Styroloxid in *Streptomyces antibioticus* Tü4 (150) verantwortlich ist. Bekannt war ein 250 bp Fragment, das bei BLAST-Suchen signifikante Ähnlichkeit zu EH Sequenzen zeigte.

Um die Datenmenge zu erhöhen und eventuelle Verwandtschaften zu weiteren Enzymklassen zu ermitteln, wurde eine Datenbank basierend auf BLAST-Suchen mit repräsentativen Ausgangssequenzen erstellt. EHs zeigen auf Sequenz- und Strukturebene eine enge Verwandtschaft mit HDs (62). Daher wurden bei der Erstellung der Datenbank nicht nur EHs, sondern auch HD Sequenzen als Ausgangspunkte verwendet. Die EH/HD Datenbank wurde mit der schon bestehenden LED (Lipase Engineering Database) (45) verglichen und strukturelle Untersuchungen der Gesamtstrukturen der in beiden Datenbanken enthaltenen α/β Hydrolasen des GX-Typs (111) durchgeführt.

Der Ergebnisteil dieser Arbeit wurde in zwei Teile aufgeteilt. Der erste Teil befasst sich mit der systematischen Analyse von 93 EH Sequenzen, dem Design familienspezifischer Primer und der Amplifikation der EH aus *Streptomyces antibioticus* Tü4. Im zweiten Teil der Arbeit wird die Erstellung und Auswertung der EH/HD Datenbank beschrieben.

4.2 Teil 1: Systematische Analyse

Für die systematische Analyse von EHs wurden 93 unterschiedliche Sequenzen verwendet. 59% dieser durch Stichwort- und BLAST-Suche ermittelten Sequenzen sind putative Proteine. Sie stammen aus verschiedenen bakteriellen und eukaryotischen Organismen wie Pilzen bzw. Hefen, Pflanzen, Insekten oder Säugetieren. Eukaryotische EHs sind gut untersucht (86, 99). Diese Enzyme sind jedoch industriell unbedeutend, da zwar hohe Enantiomerenüberschüsse erzielt wurden, die Ausbeuten jedoch gering waren (25). Als Alternative besitzen viele bakterielle EHs die Fähigkeit industriell interessante Epoxide mit hoher Enantioselektivität umzusetzen (7, 70, 130). Bakterielle EHs haben den Vorteil, dass sie sich fermentativ in großen Mengen herstellen lassen.

4.2.1 Das Multisequenz-Alignment

Das Multisequenz-Alignment dieser 93 Sequenzen (Abbildung 7) zeigt deutlich die große, für diese Enzymfamilie typische Variabilität, besonders N-terminal zeigen die EHs auffällige Unterschiede. Die Differenzen in der Sequenzlänge können mehr als 200 Aminosäuren betragen. Diese N-terminalen Unterschiede korrelieren mit der Herkunft der verschiedenen EHs. Die längsten N-terminalen Sequenzen haben die cytosolischen Säugetier EHs, verursacht durch eine zusätzlich zur eigentlichen EH vorhandenen Domäne. Diese Domäne weist eine eigene Phosphataseaktivität auf (30, 98), welche unabhängig von der EH Aktivität zu sein scheint. Pflanzliche und cytosolische bakterielle EHs dagegen haben einen kurzen N-Terminus, der keine Ähnlichkeit zur zusätzlichen Domäne der cytosolischen Säugetier EHs aufweist und variable ist. Mikrosomale EHs aus Säugetieren, Insekten, Pilzen und einigen Bakterien enthalten N-terminal einen „Meander“ (151). Zusätzlich zu diesem weisen mikrosomale Säugetier und Insekten EHs einen Membrananker auf, welcher allerdings bei mikrobiellen mikrosomalen EHs nicht vorkommt.

Das Multisequenz-Alignment zeigt im weiteren Verlauf abwechselnd homologe und variable Bereiche. Die Aminosäuren der katalytischen Triade befinden sich in den homologen Bereichen. Das Nukleophil und die $\beta 6$ -Position der katalytischen Säuregruppe befinden sich in einem ca. 150 AS langen Bereich, der auf den N-Terminus folgt. Die $\beta 7$ -Position der katalytischen Säuregruppe und das katalytische His befinden sich am C-terminalen Ende der Sequenz in einem ca. 100 AS langen Bereich.

<i>Agrobacterium radiobacter</i>	-----	
<i>Mus musculus</i>	MALRVAAFDLGVLALPSIAGAFRRSEEALALPRDFLLGAYQTEFPEGPT	50
<i>Aspergillus niger</i> EH	-----	
<i>Agrobacterium radiobacter</i>	-----	
<i>Mus musculus</i>	EQLMKGKITFSQWVPLMDESYRKSSKACGANLPENFSISQIFSQAMAARS	100
<i>Aspergillus niger</i> EH	-----	
<i>Agrobacterium radiobacter</i>	-----	
<i>Mus musculus</i>	INRPMLOAAIALKKGFTTCIVTNNWLDGDKRDSLAQMMCELSQHFDL	150
<i>Aspergillus niger</i> EH	----- MSAPFAKE	8
<i>Agrobacterium radiobacter</i>	-----	
<i>Mus musculus</i>	IESCQVGMIKPEPQIYNFLDITLAKPNEVVFLDDFGSNLKPARDMGMVT	200
<i>Aspergillus niger</i> EH	PSSASISPFPVTSIPDEQLDDLKTLVRLSKIAPPPTYESLQADGRFGITS	58
<i>Agrobacterium radiobacter</i>	-----	
<i>Mus musculus</i>	ILVHNTASALR ELEKVTGTQFPEAPLPVPCNPNDVSHGVVTVKPGIRLHF	250
<i>Aspergillus niger</i> EH	EWLTTMREKWLSEFDWR ----- PPEARLN SFPQ-FTEIEGLTIHF	98
	HGX-Motif Oxyanionhole	
<i>Agrobacterium radiobacter</i>	VREGA----GPTLLLL HW PGFWWEWSKVIPLAEH-----YDVIVPD	62
<i>Mus musculus</i>	VEMGS---GPALCLC HW PESWFSWRYQIPALAQAG-----FRVLAID	290
<i>Aspergillus niger</i> EH	AALFSEREDAVPIALL HW PGSFVEFYPIQLFREYTPETLPFHLVVP	148
	GXGXS-Motif Nucleophilic ellbow Nukleophil Oxyanionhole	
<i>Agrobacterium radiobacter</i>	LR HW GD SEKPDNLNLSKYSLDKAADDQAALLDALGIEKAYVV HW - HDFAAI	111
<i>Mus musculus</i>	MK HW GD SSP--PEIEEYAMELLCKEMVTFLDKLGIPQAVFI HW - HDWAGV	337
<i>Aspergillus niger</i> EH	LP HW TF SSGP--PLDKDFGLMDNARVVDQLMKDLGFGSGYIIQ HW GD ISF	196
	Protonendonatoren	
<i>Agrobacterium radiobacter</i>	VLHKFIRKYSDRVIAKAAIFD---PIQPDFGPVYFGLGHVHESW HW SQFHQ	157
<i>Mus musculus</i>	MVWNMALFYPERVRAVASLNTPFMPPDPVSPMKVIRSIPVFN HW QLYFQE	387
<i>Aspergillus niger</i> EH	VGRLLGVGFDACKAVHLNLCAMRAPPEGPSIESLSAAEKEGIARMEKFMT	246
<i>Agrobacterium radiobacter</i>	LDMAVEVVGSSREVCKKYFKHFFDHSYR-----	186
<i>Mus musculus</i>	PGVAEAELEKNMSRTFKSFFRASDETGFIAVHKATEIGGILVNTPEDPNL	437
<i>Aspergillus niger</i> EH	DGLA HW AMEHSTRPSTIGHVLSSSPIALLAWIG-----EKYLQWV	285
<i>Agrobacterium radiobacter</i>	DELLTEEELEVHVDNCKMPDNIHGGFNY HW RANIRPDAALWTDLDHTMSDL	236
<i>Mus musculus</i>	SKITTEEEIEFYIQFCKTG-FRGPLNW HW RNTER-NWKWCKGLGRKILV	485
<i>Aspergillus niger</i> EH	DKPLPSETILEMVSLYWLTESFPRAIHT HW RETTPTASAPNGATMLQKELY	335
	kat. Säuregruppe kat. His	
<i>Agrobacterium radiobacter</i>	PVTMIWGLG HW DTCVFYAPLIEFVVPKYYSNYTMETIEDCG HW FLMVEKP----	282
<i>Mus musculus</i>	PALMVTAEK HW DIVLRPEMSKNMEKWI PFLKR-GHIEDCG HW TQIEKPTEVNQ	535
<i>Aspergillus niger</i> EH	IHKPFSGSFFPK HW DLCVPRSWIATTGNLVFFRDHAEGG HW FAALERPRELK	385
<i>Agrobacterium radiobacter</i>	EIAIDRIKTAFR-----	294
<i>Mus musculus</i>	ILIKWLQTEVQNPSVTSKI	554
<i>Aspergillus niger</i> EH	TDLTAFVEQVWQK-----	398

Abbildung 7

Variabilität des N-terminalen Bereichs von EH-Sequenzen (grün) am Bsp. der EHs aus *Agrobacterium radiobacter* (cytosolisch), *Mus musculus* (cytosolisch) und *Aspergillus niger* (mikrosomal).

4.2.2 Superfamilien

Der aus dem Multisequenz-Alignment resultierende phylogenetische Baum (Abbildung 8), zeigt eine deutliche Auftrennung in zwei Superfamilien: cytosolische und mikrosomale EHs. Superfamilien werden aus homologen Familien gebildet, die eine niedrige aber signifikante Sequenzähnlichkeit aufweisen. Als Definition einer Superfamilie wurde die Konservierung des Nukleophils, des katalytischen His und der $\beta 6$ -Position angewendet. Die $\beta 7$ -Position, die bei der überwiegenden Anzahl der EHs die katalytische Säuregruppe trägt, eignet sich nicht hierfür, da sie bei einem Teil der $\beta 6$ EHs nicht konserviert ist. Die $\beta 6$ -Position jedoch ist in allen EHs einer Superfamilie konserviert und trägt für die einzelnen homologen familienspezifische Reste.

Das Vorhandensein von EHs beider Superfamilien in einem Organismus ist nur bei Säugetieren zu beobachten. Insekten und Pilze/Hefen weisen nur mikrosomale EHs auf, während die pflanzlichen Enzyme zu den cytosolischen EHs gehören. Bakterielle EHs sind Bestandteil beider Superfamilien, der Großteil bakterieller Sequenzen ist jedoch cytosolisch.

Die sternförmige Anordnung des cytosolischen Teiles des phylogenetischen Baumes bedeutet, dass die durch das Multisequenz-Alignment gegebenen Informationen nicht ausreichen, um die Verwandtschaftsverhältnisse der einzelnen EHs darzustellen. Um eine bessere Auflösung zu erhalten, wurden separate Bäume für mikrosomale und cytosolische EH Sequenzen berechnet (Abbildung 9 und 10).

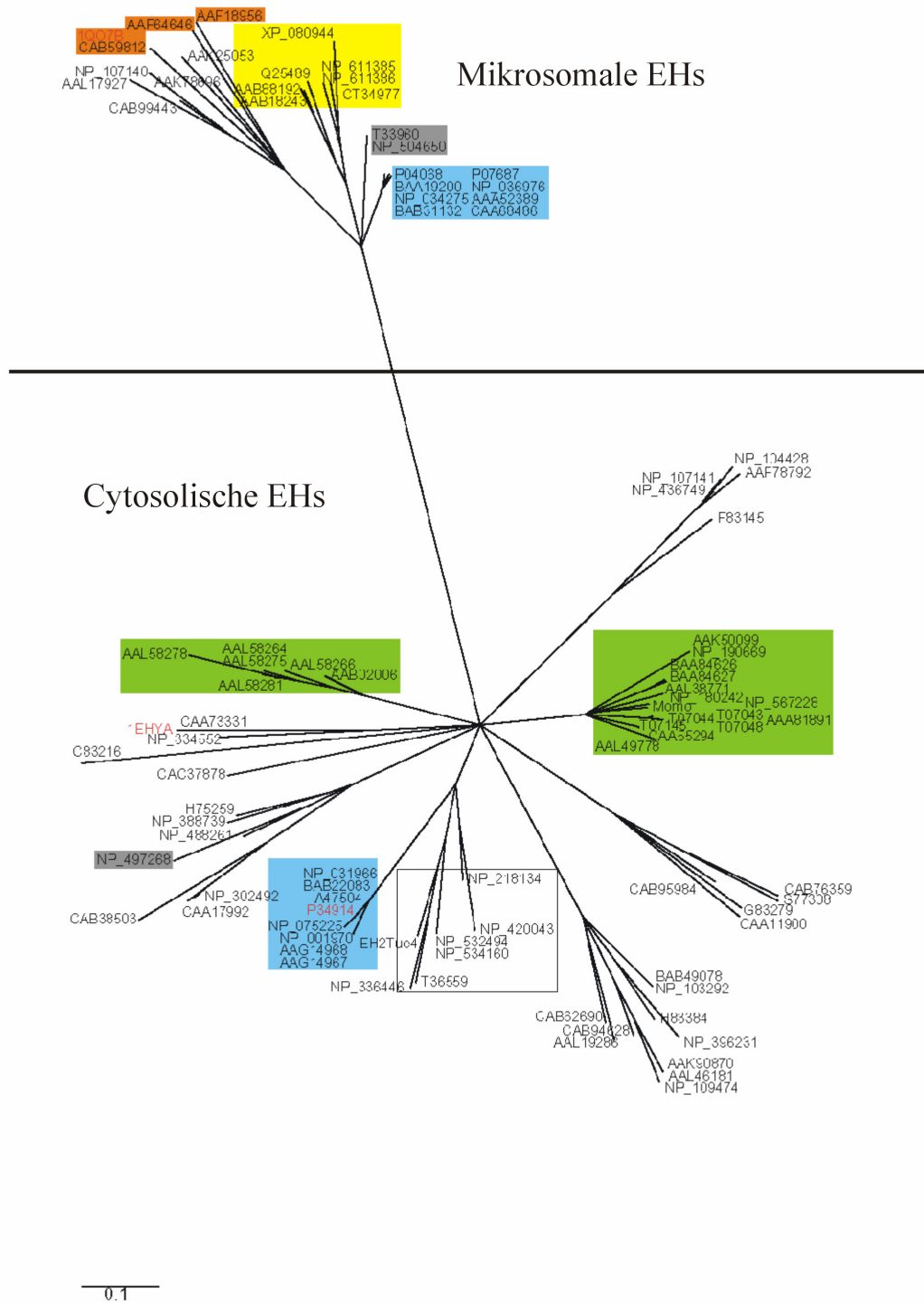


Abbildung 8

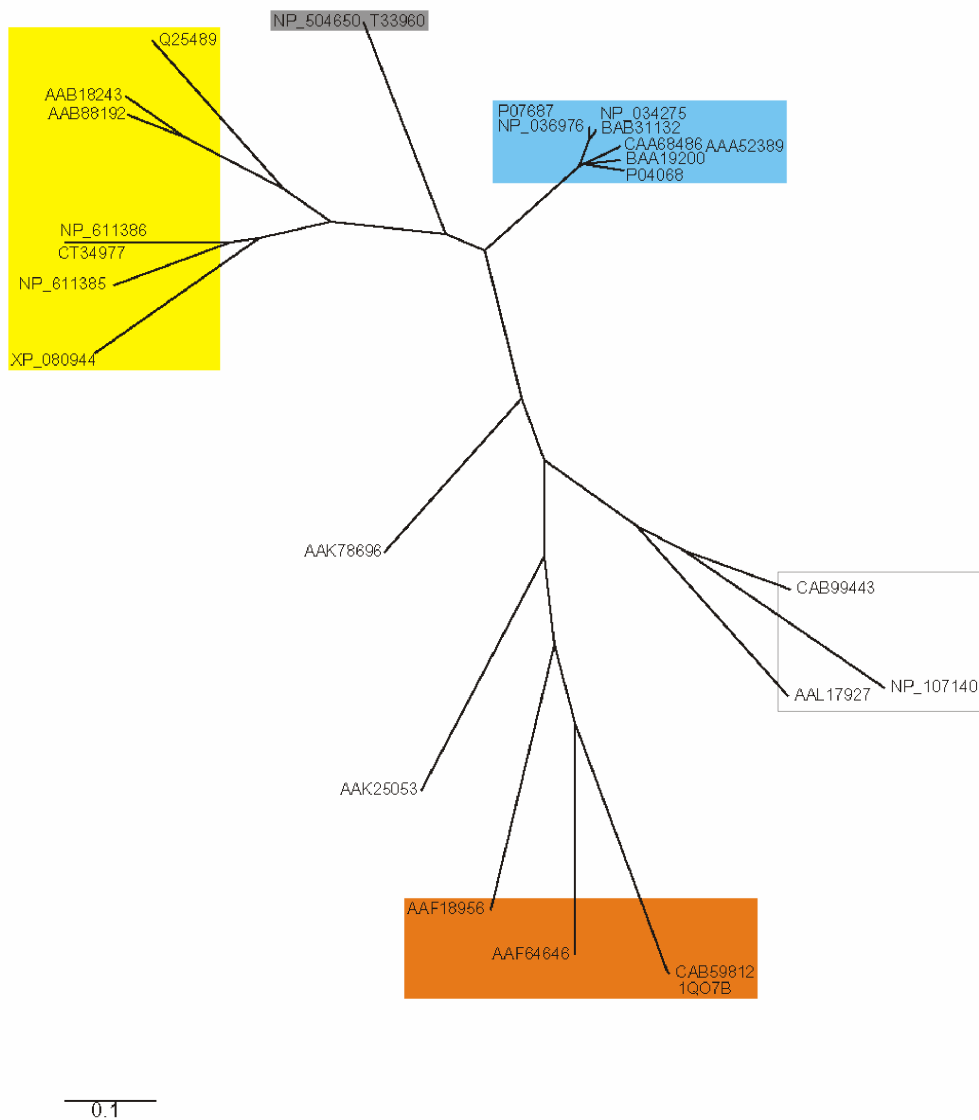
Phylogenetischer Baum aus 93 Ets. Der Baum ist unterteilt in mikrosomale (oben) und cytosolische Ets (unten). Farblich hervorgehoben wurden die homologen Familien der Säugetier Ets (blau), pflanzlichen Ets (grün), Insekten Ets (gelb) und Pilze/Hefen Ets (orange). Nicht gekennzeichnet wurden bakterielle Ets, mit Ausnahme einer isolierten Familie cytosolischer bakterieller Ets, verwandt mit eukarotischen Ets, welche mit einem schwarzen Rahmen gekennzeichnet wurde

4.2.3 Einteilung in homologe Familien

Die Mitglieder einer homologen Familie zeichnen sich durch eine hohe Sequenzähnlichkeit aus. Da phylogenetische Bäume die verwendeten Sequenzen nach ihrer Ähnlichkeit ordnen, beruht die Einteilung in homologe Familien in dieser Arbeit auf der phylogenetischen Analyse der 93 EHs. Wie in den Abbildungen 9 und 10 zu sehen, wurden 12 homologe Familien ermittelt:

Die Superfamilie der mikrosomalen EHs weist 4 homologe Familien, bestehend aus 1 Familie Säugetier EHs, 1 Familie Insekten EHs, 1 Familie Pilze/Hefen EHs und 1 Familie bakterieller EHs auf. Zusätzlich enthält die mikrosomale Superfamilie 4 isolierte EH Sequenzen: 2 EHs aus *Caenorhabditis elegans* und 2 bakterielle EHs. Die Superfamilie der cytosolischen EHs besteht aus 8 homologen Familien, welche sich auf 1 Familie Säugetier EHs, 2 Familien Pflanzen EHs und 5 Familien bakterieller EHs verteilen. Zusätzlich enthält die cytosolische Superfamilie 8 isolierte EH Sequenzen: 1 EH aus *Caenorhabditis elegans* und 7 bakterielle EHs, die sich keiner homologen Familie zuordnen ließen.

Unter den 5 homologen Familien cytosolischer bakterieller EHs befinden sich zwei Familien ($\beta 6$ EHs 1+2), deren Mitglieder zum Teil an der $\beta 6$ -Position oder an der $\beta 6$ - und der $\beta 7$ -Position eine Säuregruppe enthalten, die sich jeweils in Wasserstoffbrücken Distanz zum katalytischen His befinden. Eine weitere homologe Familie befindet sich isoliert von den restlichen bakteriellen EHs in der Nähe der eukaryotischen EHs. Während bakterielle EHs eine große Variabilität auf Sequenzebene aufweisen (Abbildung 9), sind die homologen Familien der eukaryotischen EHs aus Säugetieren, Insekten und Pflanzen gut zu erkennen

**Abbildung 10**

Phylogenetischer Baum der mikrosomalen EH Superfamilie. Farblich hervorgehoben wurden die homologen Familien der Säugetier EHs (blau), Insekten EHs (gelb) und Pilze/Hefen EHs (orange). Die homologe Familie der bakteriellen EHs wurde mit einem schwarzen Rahmen gekennzeichnet. Die EHs aus *C. elegans* wurden grau hinterlegt.

4.2.4 Konservierte AS

Fünf katalytisch relevante AS sind für alle EHs absolut konserviert:

- 1) Zwei AS der katalytischen Triade, das Nukleophil und das katalytische His, welche in direktem Kontakt mit dem Substrat sind. Die dritte AS der katalytischen Triade, die katalytische Säuregruppe, zeigt Plastizität: in den meisten EHs befindet sie sich nach dem Strand $\beta 7$, in manchen EHs aber nach $\beta 6$. In einigen bakteriellen EHs befinden sich an beiden Positionen Säurereste. Die Seitenketten beider AS sind auf das katalytische His ausgerichtet und sind in Wasserstoffbrücken Distanz, um die positive Ladung des Imidazolringes zu stabilisieren. Die katalytische Säuregruppe kann durch Asp oder Glu gebildet werden.
- 2) Das Oxyanionhole wird durch die Backbone Stickstoffe zweier AS gebildet. Die erste AS ist ein Teil des HGX-Motivs (92), die zweite ist der direkte Nachfolger des Nukleophils in der Sequenz.
- 3) Alle EHs enthalten ein absolut konserviertes Tyr in Helix $\alpha 5$ der Cap-Domäne. Die meisten EHs enthalten ein zweites Tyr in Helix $\alpha 1$ der Cap-Domäne, dessen Position nicht in allen EHs konserviert ist. Einigen EHs beider Superfamilien, bei $\beta 6$ EHs sogar allen, fehlt dieses Tyr.

4.2.5 Die $\beta 6$ -Position

Die $\beta 6$ -Position, bei einigen cytosolischen bakteriellen EHs von katalytischer Bedeutung, scheint mehr Information über Herkunft und Evolution einer EH zu enthalten als auf den ersten Blick ersichtlich. An dieser Position kommen hauptsächlich Asn, Ser, Thr, Gly, Asp und Glu vor, alle, mit Ausnahme von Gly, AS mit nukleophilem Potential. Ausnahmen bilden zwei putative cytosolische bakterielle EHs, die Ala bzw. Val an dieser Position tragen. Während innerhalb der Superfamilie der mikrosomalen EHs mit zwei Ausnahmen (Thr) ein Asn an der $\beta 6$ -Position konserviert ist, ist die Position bei den cytosolischen EHs weniger gut erhalten. Hier beschränkt sich die Konservierung auf die einzelnen homologen Familien. Die eukaryotischen homologen Familien, Säugetier EHs (Asn), Pflanzen1 (Gly), Pflanzen2 (Ser) weisen familienspezifische Reste auf, während bei cytosolischen bakteriellen EHs die Variabilität größer ist. Hier weisen vor allem die homologen Familien der zu eukaryotischen EHs verwandten bakteriellen EHs (BakterienII, Ser) und die Familie der EHs aus den gram-negativen Wurzelbakterien der Rhizobien (Ser) konservierte $\beta 6$ -Positionen auf. Beide Familien bestehen aus Bodenbakterien (Ausnahme: BakterienII, *Mycobacterium tuberculosis*

EH) und tragen wie die homologe Familie der Pflanzen2 ein Ser. Die Familie der $\beta 6$ EHs 1 enthalten ein Asp oder Ser, die $\beta 6$ EHs 2 ein Glu oder ein Gly. Die restlichen cytosolischen bakteriellen EHs variieren an der $\beta 6$ -Position zwischen Ser und Asn.

Wie anhand der $\beta 6$ EHs ersichtlich, befindet sich diese Position in der Nähe des katalytischen His. AS wie Asn, Ser, Thr, Asp oder Glu könnten die Funktion der katalytischen Säuregruppe unterstützen, während Gly, als kleinste aller AS keinen Einfluss hätte. AS wie z.B. das basische Lys dagegen, hätten einen eher gegenteiligen Effekt und AS mit sperrigen Ringsystemen scheitern an sterischer Hinderung. Auf DNA-Ebene fällt auf, dass sich Ser, Thr, Asp und Glu durch Punktmutationen aus Asn darstellen lassen. Gly könnte sich durch eine zweite Punktmutation aus Ser, Asp oder Glu entwickelt haben. Die $\beta 6$ -Position könnte also einen Hinweis auf die Entwicklungsgeschichte der einzelnen homologen EH Familien geben.

Tabelle 1: Aminosäuren an der $\beta 6$ -Position der einzelnen homologen Familien

Homologe Familie	AS der $\beta 6$ -Position
cytosolische EHs	
Säugetiere	Asn
Pflanzen1	Gly
Pflanzen2	Ser
BakterienI	Ser, Asn, Gly, Asp, Glu
BakterienII	Asn
mikrosomale EHs	
Säugetiere	Asn
Insekten	Asn
Pilze/Hefe	Asn, Thr
Bakterien	Asn, Thr

4.2.6 Remodellierung der Kristallstruktur der *Agrobacterium radiobacter* EH

Die Struktur der *Agrobacterium radiobacter* EH (1EHY) wurde als Tetramer kristallisiert (96). Jedes Monomer besteht aus 294 AS. Für die Kristallstruktur wurden einige ungewöhnliche Konformationen beschrieben. Intermolekulare Kontakte zwischen den einzelnen Monomeren legen den Schluss nahe, dass diese auf Kräfte, die während der Kristallisation auftraten, zurückzuführen sind. Das als katalytische Säuregruppe identifizierte Asp246 befindet sich nicht in ausreichender Nähe zum katalytischen His um seine katalytische Funktion erfüllen zu können. Stattdessen zeigt der Loop 244-257 in Richtung des Lösungsmittels, verursacht durch einen Kontakt mit dem Nachbar-Monomer. Die entstandene Lücke wurde durch Gln134 gefüllt, welches nun wie ein Inhibitor das aktive Zentrum blockiert. Der Loop 138-148 konnte aufgrund schlechter Werte für die Elektronendichte nicht aufgelöst werden (Abbildung 11).

Für die systematische Analyse der bekannten EH Strukturen wurde die Struktur der EH aus *Agrobacterium radiobacter* remodelliert. Als Template wurde die Struktur der Non-heme Peroxidase aus *Streptomyces aureofaciens* (1BRO) verwendet, welche strukturell sehr große Ähnlichkeit aufweist. Remodelliert wurden der Loop 244-257 und der Loop 132-148. Die AS 132-137 wurde ebenfalls modelliert, da dieser Teil des Loops Gln134 trägt, das das aktive Zentrum blockiert. Nach der Modellierung (Abbildung 11) befand sich das katalytische Asp246 in ausreichender Nähe zum His275, das Gln134 befand sich nicht mehr im aktiven Zentrum und der Loop 138-148 war modelliert. Die Stabilität des Modells wurde mit einer MD-Simulation erfolgreich getestet.

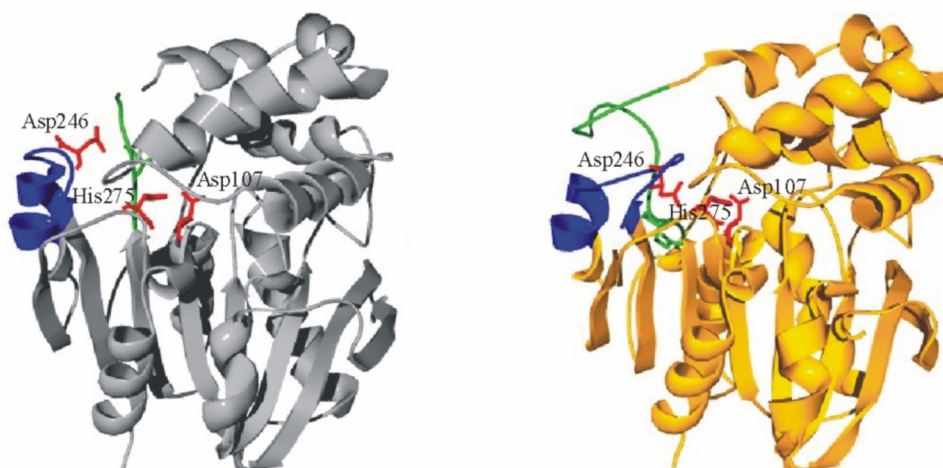


Abbildung 11

Kristallstruktur (grau) und remodellierte Struktur (gelb) der *Agrobacterium radiobacter* EH. Farbig markiert wurden die Loops 244-257 (blau) und 132-148 (grün), die katalytische Triade (rot) und das Gln134 (grün).

4.2.7 Modularer Aufbau von EHs

Bei den drei bekannten Strukturen handelt es sich um eine cytosolische Säugetier EH (*Mus musculus*, PDB Eintrag: 1CQZ), eine cytosolische bakterielle EH (*Agrobacterium radiobacter*, PDB Eintrag: 1EHY) und eine mikrosomale EH (*Aspergillus niger*, PDB Eintrag: 1QO7). Auf den ersten Blick erscheinen die drei Strukturen sehr unterschiedlich. Diese Unterschiede werden durch den variablen N-terminalen Bereich verursacht. Werden die Strukturen überlagert, wird eine erstaunliche strukturelle Übereinstimmung des α/β Hydrolase Folds und der Cap-Region sichtbar. Aufgrund der guten Konservierung wurden die Strukturen in Module eingeteilt (Abbildung 12):

N-terminale Domäne: Dieser Bereich weist die größte Variabilität innerhalb der EHs auf: cytosolische Säugetier EHs enthalten eine durch einen prolinreichen Linker verbundene Domäne, wohingegen cytosolische bakterielle und pflanzliche EHs keine N-terminale Domäne haben. Mikrosomale EHs dagegen enthalten einen „Meander“ aus 4 α -Helices. Die eukaryotischen mikrosomalen EHs von Säugetieren und Insekten haben zusätzlich einen Membrananker.

N-terminale katalytische Domäne: Enthält die erste Hälfte des α/β Hydrolase Folds, bis zum Strand β_6 . Sie besteht aus 6 β -Strands und 4 α -Helices. Die N-terminale katalytische Domäne enthält das Nukleophil und die β_6 -Position der katalytische Säuregruppe.

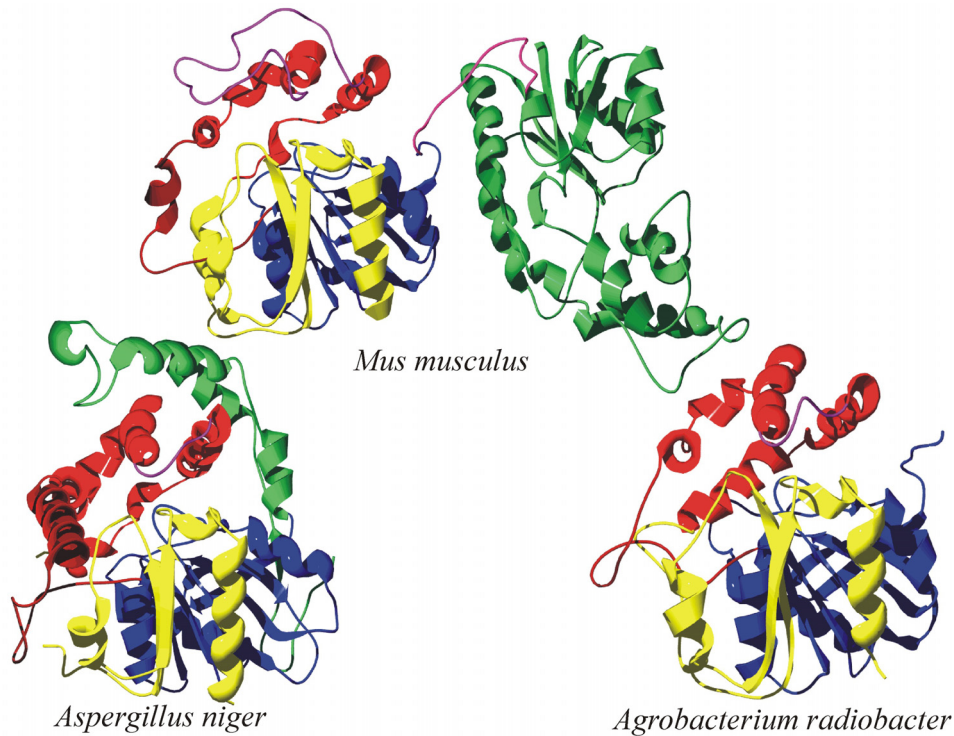
NC-Loop: Variabler Bereich, der N-terminale katalytische Domäne und Cap-Domäne verbindet. Der NC-Loop beginnt direkt nach der β_6 -Position und endet vor der ersten Cap-Helix. Er ist überwiegend hydrophob, seine Länge variiert für die 3 bekannten Strukturen zwischen 19 AS (*Agrobacterium radiobacter* EH), 23 AS (*Mus musculus* EH) und 35 AS (*Aspergillus niger* EH). Die beiden längeren NC-Loops enthalten eine α -Helix während der kurze Loop der *Agrobacterium radiobacter* EH keine Sekundärstrukturelemente aufweist.

Cap-Domäne: Diese, oberhalb des α/β Hydrolase Folds positionierte Domäne ist bei EHs nach Strand β_6 des α/β Hydrolase Folds inseriert. Sie besteht aus 5 α -Helices, die ähnlich dem Uteroglobin in zwei Lagen angeordnet sind (62). Die obere, trapezförmige Lage wird durch die Helices α_1 , α_2 und α_3 gebildet, die untere, einen Winkel von ca. 100° einschließend, von den Helices α_4 und α_5 . In den drei Strukturen haben die Helices α_1 , α_2 , α_3 und α_4 ähnliche Längen (Helix α_1 : 5-8, Helix α_2 : 7-10, Helix α_3 : 14 und Helix α_4 : 8-12 AS). Helix α_5 dagegen besteht aus 11 AS (*Agrobacterium radiobacter* und *Aspergillus niger* EH), aber nur aus 4 AS für die EH aus *Mus musculus*. Die Cap-Domäne enthält ein absolut konserviertes Tyr in Helix α_5 und die meisten EHs enthalten ein zweites Tyr in Helix α_1 .

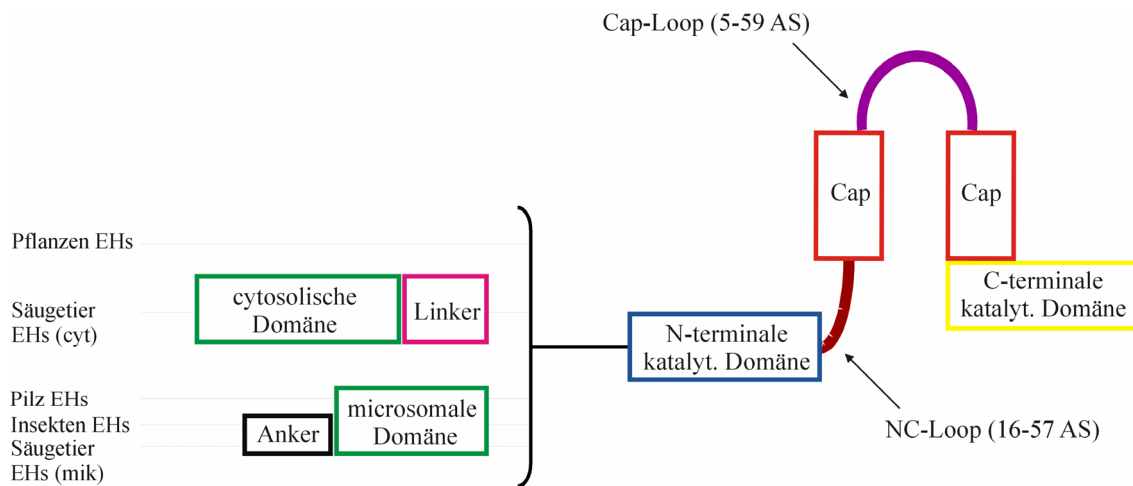
Cap-Loop: Variabler Loop innerhalb der Cap-Domäne. Der Cap-Loop verbindet die beiden Lagen der Cap-Domäne und befindet sich zwischen den Helices $\alpha 3$ und $\alpha 4$. Seine Länge variiert für die 3 bekannten Strukturen zwischen 8 AS (*Aspergillus niger* EH), 9 AS (*Agrobacterium radiobacter* EH) und 36 AS (*Mus musculus* EH). Bei den Cap-Loops der drei bekannten Strukturen steigert sich der Grad der Hydrophobizität mit wachsender Länge. Die *Aspergillus niger* EH enthält einen hydrophilen Cap-Loop (25% hydrophobe AS), die *Agrobacterium radiobacter* EH einem erhöhten hydrophoben Anteil (43% hydrophobe AS) und die *Mus musculus* EH einen überwiegend hydrophoben Cap-Loop (63% hydrophobe AS).

C-terminale katalytische Domäne: Enthält den zweiten Teil des α/β Hydrolase Folds ab Strand $\beta 7$. Sie besteht aus 2 β -Strands und 4 α -Helices. Die C-terminale katalytische Domäne enthält das katalytische His und die $\beta 7$ -Position der katalytischen Säuregruppe.

Diese Einteilung in „Module“ wurde auf das Multisequenz-Alignment übertragen. Die Sekundärstrukturen einzelner EH Sequenzen wurden berechnet, um die Übertragbarkeit der Module auf das Multisequenz-Alignment abzuschätzen. Wie in der Schema-Zeichnung in Abbildung 13 zu sehen, spiegelt das Alignment diesen modularen Aufbau wider, die variablen und konservierten Regionen sind gut erkennbar.

**Abbildung 12**

Bekannte Strukturen aus *Agrobacterium radiobacter* (unten rechts), *Mus musculus* (oben) und *Aspergillus niger* (unten links). EHs zeigen einen modularen Aufbau unterteilt in N-terminale Domäne (grün), N-terminale katalytische Domäne (blau), NC-Loop (braun), Cap-Domäne (rot), Cap-Loop (violett) und C-terminale katalytische Domäne (gelb).

**Abbildung 13**

Schematische Darstellung des modularen Aufbaus der EHs. Die variable N-terminale Domäne (grün, linke Bildhälfte), enthalten in cytosolischen Säugetier EHs (über einen Linker (pink) verbunden), Pilze/Hefen, Insekten und mikrosomalen Säugetier EHs. Insekten und mikrosomale Säugetier EHs enthalten zusätzlich einen Membrananker (schwarz). In der rechten Bildhälfte sind die drei konservierten Domänen aufgeführt: N-terminale katalytische Domäne (blau), Cap-Domäne (rot) und C-terminale katalytische Domäne (gelb) und die beiden variablen Loops: NC-Loop (braun) und Cap-Loop (violett).

4.2.7.1 Loop-Längen

Obwohl die Sequenz Identitäten unter den EHs mit bekannter Struktur nur ca. 30% betragen, sind α/β Hydrolase Fold und die Cap-Domäne strukturell sehr gut erhalten. Sehr unterschiedlich sind dagegen die beiden variablen Loops. Für die systematische Analyse der Loop-Längen aller 93 EHs wurden mit Hilfe des Multisequenz-Alignments die ersten und letzten AS der Loops ermittelt. Um sie zu charakterisieren wurden die Loop-Längen aller EHs in einem Diagramm gegeneinander aufgetragen. Hierbei bildeten sich drei nach Superfamilien und Organismengruppen geordnete Cluster (Abbildung 14). Cluster I+II enthalten cytosolische EHs, während in Cluster III nur mikrosomale EHs vorkommen:

- 1) Die EHs von Cluster I haben mittel lange NC- und lange Cap-Loops (16-40 bzw. 31-59 AS) und beinhalten Säugetier und Pflanzen EHs und eine Gruppe mit eukaryotischen EHs verwandten bakteriellen EHs. Cluster I enthält die EH von *Mus musculus* mit bekannter Röntgenstruktur, welche einen NC- und einen Cap-Loop von 23 bzw. 35 AS Länge besitzt. Diese Loops sind lang innerhalb des Clusters, so dass diese Struktur als Template für die Modellierung der meisten EHs von Cluster I verwendet werden kann. Säugetier EHs haben Loops ähnlicher Länge (NC-Loop: 23 und Cap-Loop: 35 oder 36 AS), Pflanzen EHs zeigen eine stärkere Variabilität (NC-Loop: 16-25 und Cap-Loop: 31-35 AS) und die bakteriellen EHs von Cluster I zeigen Variabilität in beiden Loops (NC-Loop: 18-40 und Cap-Loop: 38-59 AS).
- 2) Die EHs von Cluster II sind charakterisiert durch kurze NC- und Cap-Loops (18-25 bzw. 5-12 AS) und beinhalten die bakteriellen EHs der cytosolischen Superfamilie, ausgenommen die mit Säugetier EHs verwandten bakteriellen EHs. Cluster II enthält die EH von *Agrobacterium radiobacter* mit bekannter Röntgenstruktur (NC-Loop: 19 und Cap-Loop: 9 AS). Cluster II ist deutlich separiert von den beiden anderen Clustern cytosolischer EHs (Cluster I und Cluster III), wie auch im zugehörigen phylogenetischen Baum (Abbildung 9) zu sehen, in dem die Mehrheit der bakteriellen EHs ebenfalls klar getrennt von allen anderen cytosolischen EHs ist. Für Homologiemodelle bakterieller cytosolischer EHs mit NC-Loops bis zu einer Länge von 21 AS kann die *Agrobacterium radiobacter* EH als Template verwendet werden. Für die zwei homologen β_6 EH Familien kann diese EH ebenfalls als Template eingesetzt werden, da sie Asp Reste an der β_6 - wie auch an der β_7 -Position (nach β_6 : Asp131 und nach Strand β_7 : Asp246) besitzt.
- 3) Alle mikrosomalen EHs, einschließlich der bakteriellen EHs, bilden ein einziges Cluster III. Die Längen der NC-Loops variieren stark (21-57 AS), während die Cap-

Loops von Insekten, Pilzen/Hefen und Bakterien kurz sind (8-10 AS), haben die mikrosomalen Säugetier EHs längere Cap-Loops (19 AS). Cluster IV enthält die Röntgenstruktur der *Aspergillus niger* EH (NC-Loop: 35, Cap-Loop: 8 AS). Die *Aspergillus niger* EH kann als Template nur für eine begrenzte Anzahl mikrosomaler EHs verwendet werden. Pilze und Insekten EHs und eine bakterielle EH haben zu lange NC-Loops (42-57 AS), während die Cap-Loops eine vergleichbare Länge aufweisen (8-10 AS). Für die Säugetier und die *Caenorhabditis elegans* EHs sind beide Loops zu lang (NC-Loop: 48, 49 und Cap-Loop: 19 AS).

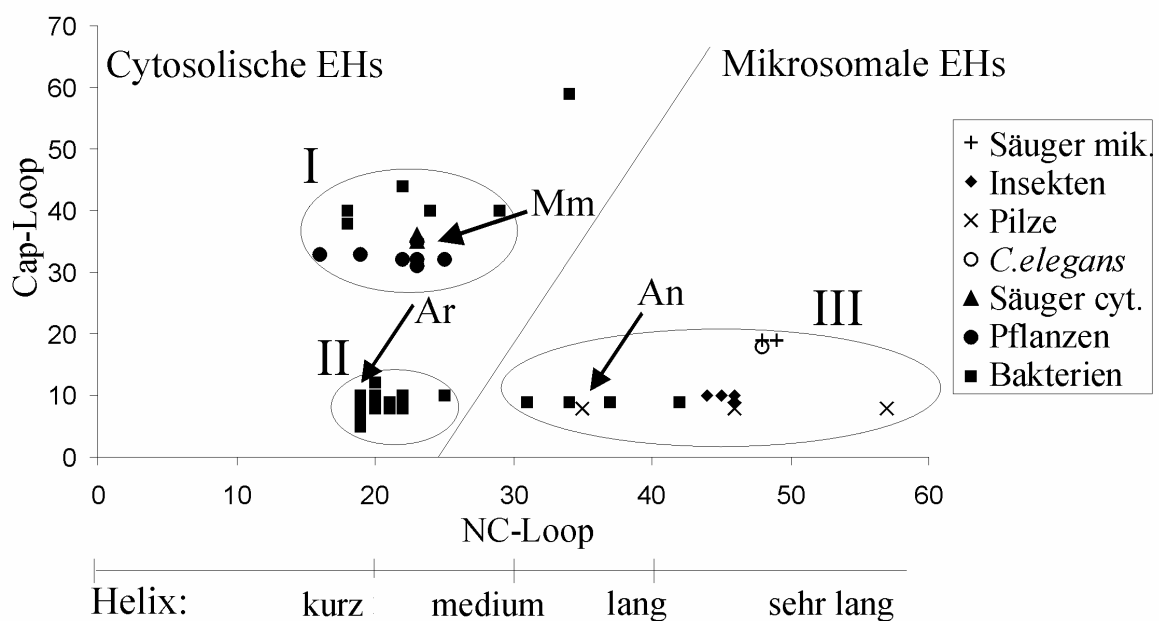


Abbildung 14

Aufgetragen nach der Anzahl an AS von NC- und Cap-Loop bilden sich drei Cluster, separiert in die beiden Superfamilien. Cytosolische EHs: Cluster I (Säugetier, Pflanzen EHs und zu eukaryotischen EHs verwandte bakterielle EHs), Cluster II (bakterielle EHs) und Cluster III (mikrosomale EHs). Die Positionen der strukturbekannteren EHs sind durch Pfeile gekennzeichnet: *Mus musculus* EH (Mm), *Agrobacterium radiobacter* EH (Ar) und *Aspergillus niger* EH (An).

4.2.7.2 Der NC-Loop

Die drei Cluster besitzen überlappende NC-Loop-Längen. Der NC-Loop der *Agrobacterium radiobacter* EH (19 AS) enthält keine helikalen Elemente, wohingegen die längeren NC-Loops der EHs von *Mus musculus* (23 AS) und *Aspergillus niger* (36 AS) eine Helix enthalten. Bei allen drei NC-Loops beträgt die Länge des nicht-helikalen Teils 16-19 AS. Für NC-Loops mit einer Länge größer 19 AS scheinen die zusätzlichen AS eine Helix zu bilden. Cluster I enthält NC-Loops von 16-25 AS Länge. Der Loop der *Oryza sativa* EH (16 AS) ist kürzer als der Loop der *Agrobacterium radiobacter* EH (19 AS) aus Cluster II, so dass für kurze NC-Loops im allgemeinen keine helikalen Elemente erwartet werden. Im Gegensatz dazu enthält die EH aus *Streptomyces coelicolor* (Cluster II) einen NC-Loop mit einer Länge von 25 AS, ausreichend lang um eine Helix zu bilden. Möglicherweise findet die Bindung von Helices in allen Clustern statt. Mit Hilfe dieser Erkenntnisse lässt sich Abbildung 14 in vier Bereiche unterteilen. Bis zu einer Länge von 19 AS findet keine Helix-Bildung statt (Bsp. 19 AS bei *Agrobacterium radiobacter* EH), von ca. 20 bis 32 AS ist die Helix kurz (Bsp. 23 AS bei *Mus musculus* EH), von 33 bis 40 AS ist die Helix lang (Bsp. 36 AS bei *Aspergillus niger* EH). Für NC-Loops länger als 40 AS kann keine Aussage getroffen werden, da keine Daten für solche lange Loops existieren. Der variable NC-Loop wird von zwei strukturell konservierten Regionen flankiert, der N-terminalen katalytischen Domäne und der Cap-Domäne. Aufgrund dessen scheint sich der NC-Loop unabhängig vom konservierten Rest des Proteins zu falten, so dass das Vorhandensein und die Länge helikaler Bereiche ausschließlich abhängig ist von der Loop-Länge.

Ein potentieller Substrat Tunnel befindet sich direkt neben dem flexiblen NC-Loop. In der Struktur der EH aus *Agrobacterium radiobacter* ist dieser NC-Loop aufgrund zu hoher B-Faktoren nicht aufgelöst (96). Auch aus den nah verwandten Haloalkan Dehalogenasen ist bekannt, dass dieser Loop eine große Flexibilität besitzt (124). Zusätzlich zeigt die dritte Cap-Helix in molekular dynamischen Simulationen der Haloalkan Dehalogenase *Sphingomonas paucimobilis* ebenfalls ein flexibles Verhalten (110). Da die dritte Cap-Helix oberhalb des NC-Loops liegt, könnte eine gemeinsame Bewegung beider Regionen einen Substrat Tunnel öffnen, der auch für größere Substrate ausreichend wäre.

4.2.8 Der Cap-Loop

Eine Substratgruppe der cytosolischen EHs sind epoxidierte Fettsäuren. Eine Fettsäure ist ein hydrophobes Molekül. Von Lipasen ist bekannt, dass sie in ihrer Bindungstasche eine Fettsäure- und eine Alkoholbindestelle enthalten (112). Die Beschaffenheit dieser Bindestellen ist direkt an die Substratspezifität gekoppelt. So setzen Lipasen mit großer Fettsäurebindestelle längerkettige Triacylglyceride um als Lipasen mit kleiner Fettsäurebindestelle (112). Die beiden bekannten EH Strukturen, deren Enzyme keine epoxidierten Fettsäuren umsetzen, weisen einen kurzen Cap-Loop auf, während die cytosolische EH einen sehr langen Cap-Loop enthält. Der Cap-Loop der *Mus musculus* EH wölbt sich oberhalb des Nukleophils über die Cap-Domäne. Hierdurch entsteht ein hydrophober Tunnel, der von der Enzymoberfläche direkt zum Nukleophil führt (Abbildung 15). Mit Hilfe des Cap-Loops wird bei eukaryotischen cytosolischen EHs eine Fettsäurebindestelle gebildet, die in den Strukturen der cytosolischen bakteriellen EH aus *Agrobacterium radiobacter* und der mikrosomalen EH aus *Aspergillus niger* nicht auftritt. Auch mikrosomale Säugetier EHs haben die Fähigkeit epoxidierte Fettsäuren umzusetzen. Wie in Abbildung 14 ersichtlich, besitzen diese EHs zwar einen kürzeren Cap-Loop als die cytosolischen Säugetier EHs, jedoch ist dieser noch deutlich länger als die Cap-Loops der restlichen mikrosomalen und cytosolischen bakteriellen EHs, welche keine epoxidierten Fettsäuren umsetzen. Eine Ausnahme bildet die cytosolische Familie der mit eukaryotischen Proteinen verwandten bakteriellen EHs, diese besitzen einen, für bakterielle EHs untypischen, langen Cap-Loop und müssten daher epoxidierte Fettsäuren als Substrat akzeptieren.

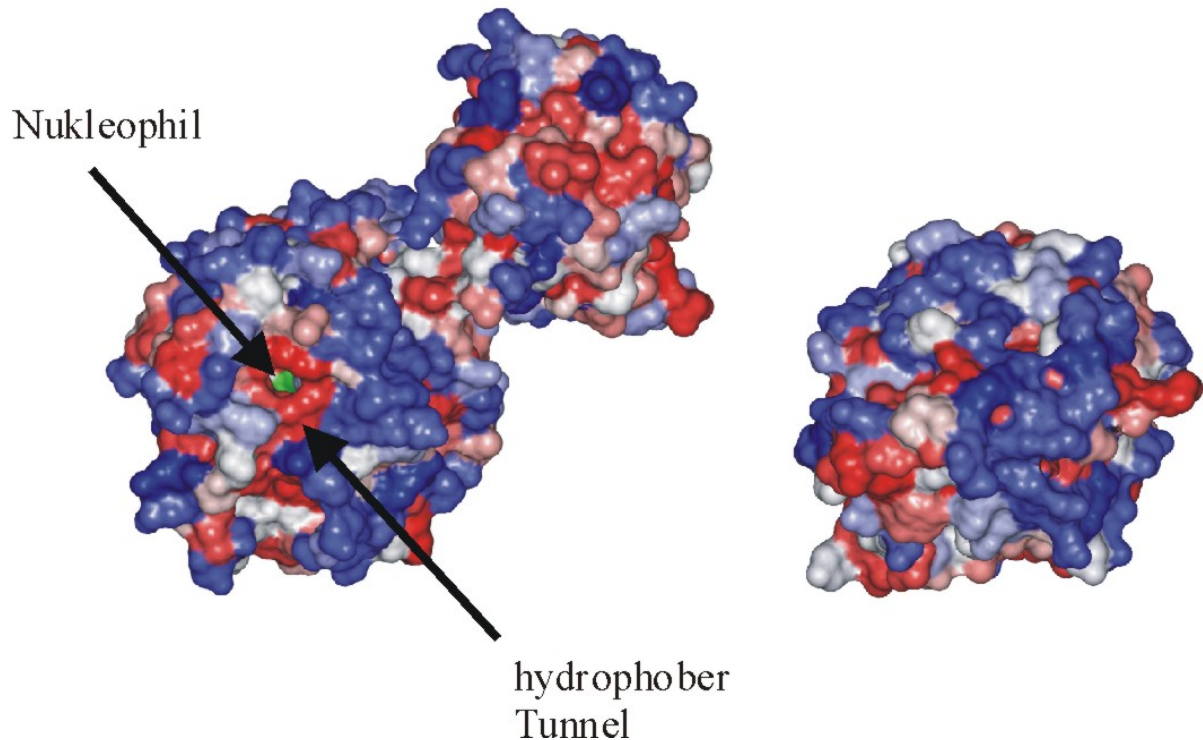


Abbildung 15

Strukturen der Ets aus *Mus musculus* (links) und *Agrobacterium radiobacter* (rechts) in der Oberflächendarstellung mit Ansicht auf die Cap-Domäne. Die Strukturen wurden nach Hydrophobizität ihrer AS eingefärbt, hydrophile AS sind blau und hydrophobe AS rot gekennzeichnet. Bei der *Mus musculus* Ets ist deutlich der durch den langen Cap-Loop gebildete hydrophobe Tunnel erkennbar, der sich direkt oberhalb des Nukleophils (grün) befindet. Die Struktur der *Agrobacterium radiobacter* Ets, welche einen kurzen Cap-Loop besitzt, weist keinen hydrophoben Tunnel oberhalb des Nukleophils auf.

4.2.9 Homologiemodellierung

Bei einem ersten Ansatz für Homologiemodellierung wurden Ziel- und Template-Sequenzen paarweise aligned. Aufgrund der niedrigen Sequenzähnlichkeiten innerhalb der Ets und der Existenz von nur drei bekannten Strukturen, könnten auf diese Weise maximal für 14 % der 93 untersuchten Ets Strukturen modelliert werden. Um möglichst alle Ets zu modellieren wurde das 93 Ets Alignment als Grundlage für die Homologiemodellierungen verwendet. Basierend auf dem 93 Ets Alignment liegen die Sequenzidentitäten für die meisten Ets zwischen 20 und 30%, dem Grenzwert für erfolgreiche Homologiemodellierungen. Um zu ermitteln, ob Homologiemodellierungen in diesem Bereich sinnvoll sind, wurden aus jedem Cluster repräsentative Ets ausgewählt: Ziel Ets haben 20-30% Sequenzidentität zur Template-Sequenz und enthalten Loops ähnlicher Länge.

- 1) Cluster I: Für die meisten EHs aus Cluster I kann die cytosolische EH aus *Mus musculus* als Template-Sequenz verwendet werden. Für einige EHs jedoch können die Loops auf Grundlage dieses Templates nicht modelliert werden: zwei Pflanzen EHs enthalten NC-Loop von 16 und 19 AS Länge, welche keine helikalen Elemente enthalten sollten. Während der NC-Loop der *Mus musculus* EH (23 AS) eine Helix enthält, weist der entsprechende Loop der *Agrobacterium radiobacter* EH (19 AS) eine geeignete Länge auf und kann als Template dienen. Zusätzlich haben mehrere bakterielle EHs dieses Clusters NC-Loops länger als 28 AS, für welche der NC-Loop der *Aspergillus niger* EH (35 AS) als Template verwendet werden kann. Manche bakterielle EHs jedoch enthalten Cap-Loops länger als 40 AS, für welche kein Template zur Verfügung steht. Somit können 94% der Strukturen aller Cluster I EHs komplett vorhergesagt werden. Als repräsentative Modelle wurden die Strukturen der EHs aus *Glycine max* (19% identische AS, 30% identische AS ohne die N-terminale Domäne) (Abbildung 16) und *Streptomyces coelicolor* (20% identische AS) (Abbildung 17) gewählt.

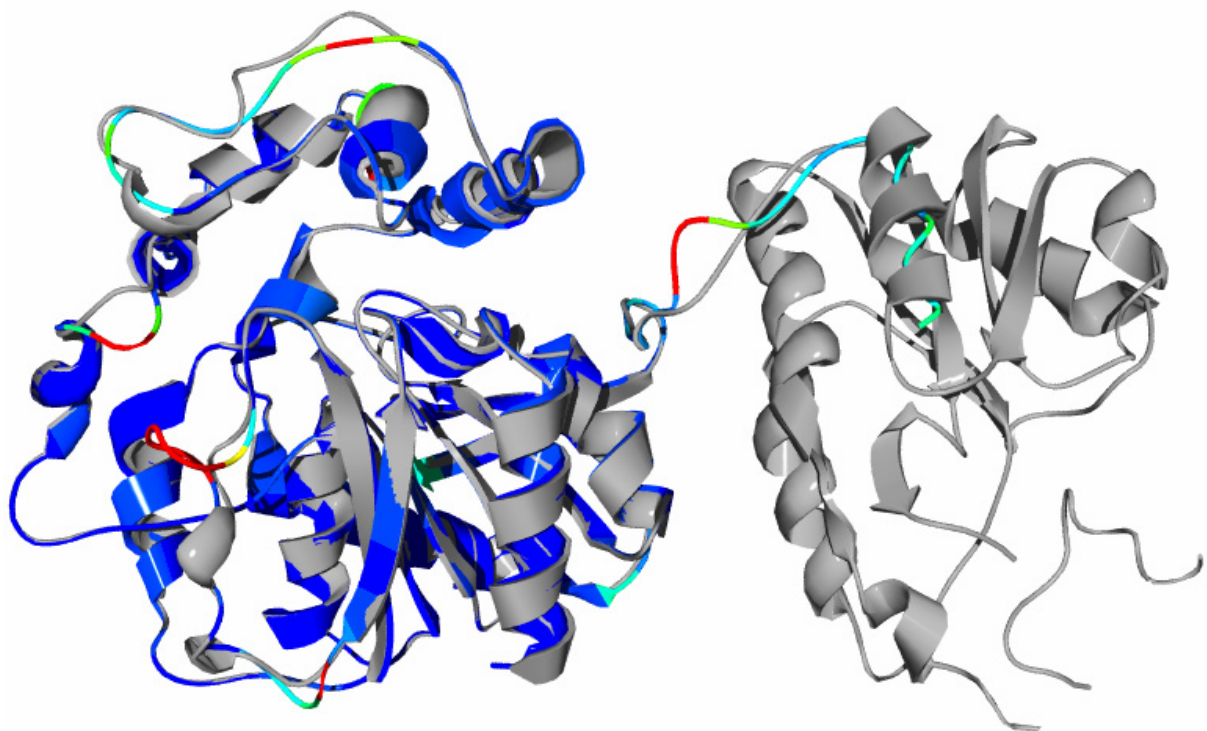


Abbildung 16

Überlagerung der Template-Struktur der *Mus musculus* EH (grau) mit der nach rmsd eingefärbten modellierten Struktur der *Glycine max* EH.

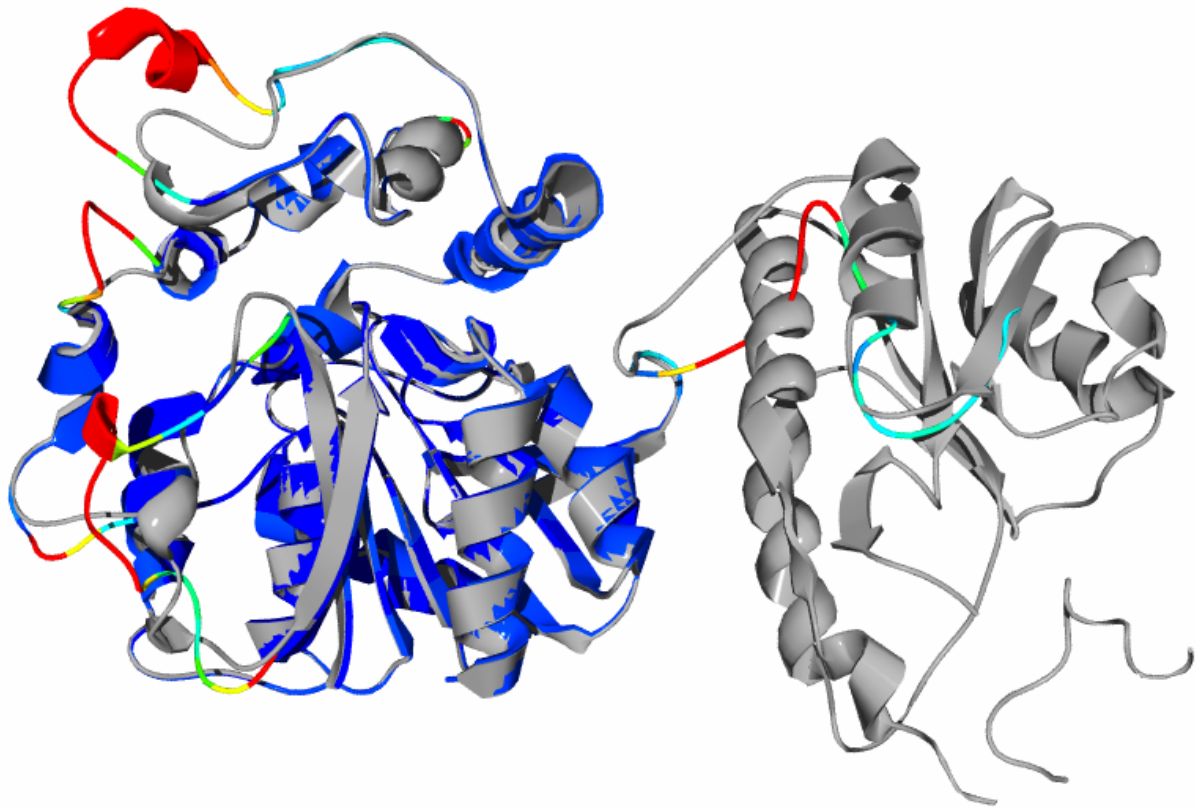


Abbildung 17

Überlagerung der Template-Struktur der *Mus musculus* EH (grau) mit der nach rmsd eingefärbten modellierten Struktur der *Streptomyces coelicolor* EH.

- 2) Cluster II: Da alle EHs von Cluster II ähnliche Loop-Längen enthalten, können die Strukturen aller Cluster II EHs vorhergesagt werden. Für eine EH mit einem NC-Loop länger als 20 AS kann der NC-Loop der *Mus musculus* EH als Template verwendet werden. Die meisten EHs dieses Clusters haben ihre katalytische Säuregruppe an der $\beta 7$ -Position, aber einige enthalten sie an der $\beta 6$ -Position. Für beide EH-Typen kann die EH aus *Agrobacterium radiobacter* als Template-Sequenz verwendet werden: sie hat geeignete NC- und Cap-Loop-Längen (19 bzw. 9 AS) und sie enthält je ein Asp an beiden Positionen, welche sich in Wasserstoffbrückenabstand zum katalytischen His befinden. Als repräsentative Modelle wurden die Strukturen der EHs aus *Mycobacterium tuberculosis* (26% identische AS, Asp in $\beta 7$ -Position) (Abbildung 18) und *Corynebacterium sp.* (25% identische AS, Asp in $\beta 6$ -Position) (Abbildung 19) gewählt.

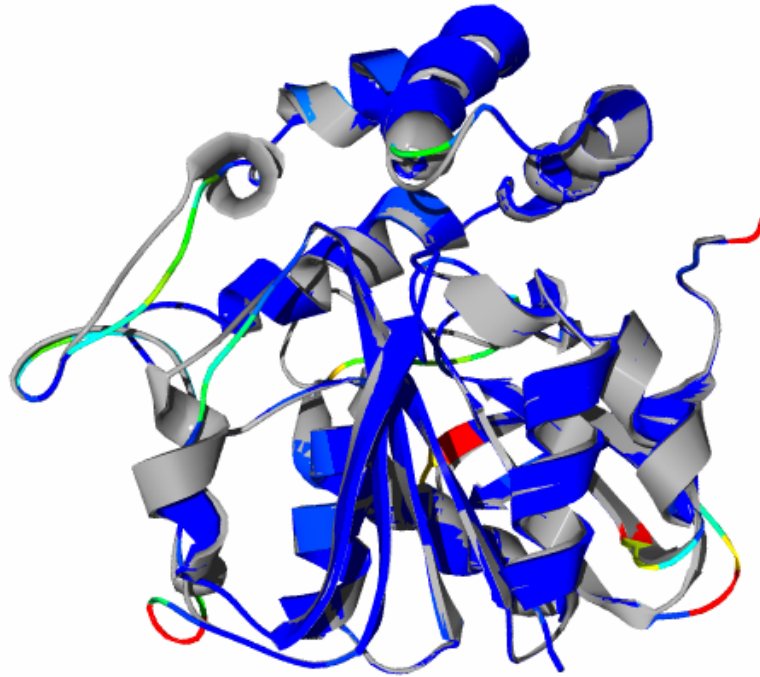


Abbildung 18

Überlagerung der Template-Struktur der *Agrobacterium radiobacter* EH (grau) mit der nach rmsd eingefärbten modellierten Struktur der *Mycobacterium tuberculosis* EH.



Abbildung 19

Überlagerung der Template-Struktur der *Agrobacterium radiobacter* EH (grau) mit der nach rmsd eingefärbten modellierten Struktur der *Corynebacterium* sp. EH.

- 3) Cluster III: Für 50% aller mikrosomale EHs kann der Cap-Loop auf Grundlage der *Aspergillus niger* EH modelliert werden, während nur für 19% aller mikrosomalen EHs die Struktur beider Loops vorhergesagt werden kann. Beide Loops der Säugetier EHs sind zu lang für die Modellierung, bei Insekten und Pilz EHs sind die NC-Loops viel länger als in der Template-Struktur. Als repräsentative Modelle wurde die Strukturen der EH aus *Mesorhizobium loti* (23% identische AS) (Abbildung 20).



Abbildung 20

Überlagerung der Template-Struktur der *Agrobacterium radiobacter* EH (grau) mit der nach rmsd eingefärbten modellierten Struktur der *Mesorhizobium loti* EH.

4.3 Familienspezifische, degenerierte Primer

Die bei der systematischen Analyse erfolgte Einteilung der EHs in Superfamilien und homologe Familien ermöglicht das Design familienspezifischer Primer für das Screening nach unbekanntem EHs (Abbildung 21). Die Primer wurden mit dem webbasierten Programm CODEHOP (120) erstellt.

Für die Erstellung der CODEHOP-Primer wurde die Einteilung der EHs in homologe Familien genutzt. Die mikrosomalen EHs wurden in eine Gruppe Säugetier EHs (mSäuger), eine Gruppe Insekten EHs (mInsekten), eine Gruppe Pilze/Hefen EHs (mPilze) und eine Gruppe bakterieller EHs (mBakterien) unterteilt. Zwei bakterielle Sequenzen aus *Clostridium acetobutylicum* (AAK78696) und *Caulobacter crescentus* (AAK25053) konnten aufgrund ihrer isolierten Lage im phylogenetischen Baum nicht beim Design der Primer berücksichtigt werden. Die cytosolischen EHs wurden in eine Gruppe Säugetier EHs (cSäuger), zwei Gruppen pflanzlicher EHs (cPflanzen1 und cPflanzen2) und 5 Gruppen bakterieller EHs (cBakterien1-5) unterteilt (Tabelle 2). Aufgrund der hohen Diversität der bakteriellen EHs der cytosolischen Superfamilie konnten einige Sequenzen nicht für das Design der Primer verwendet werden.

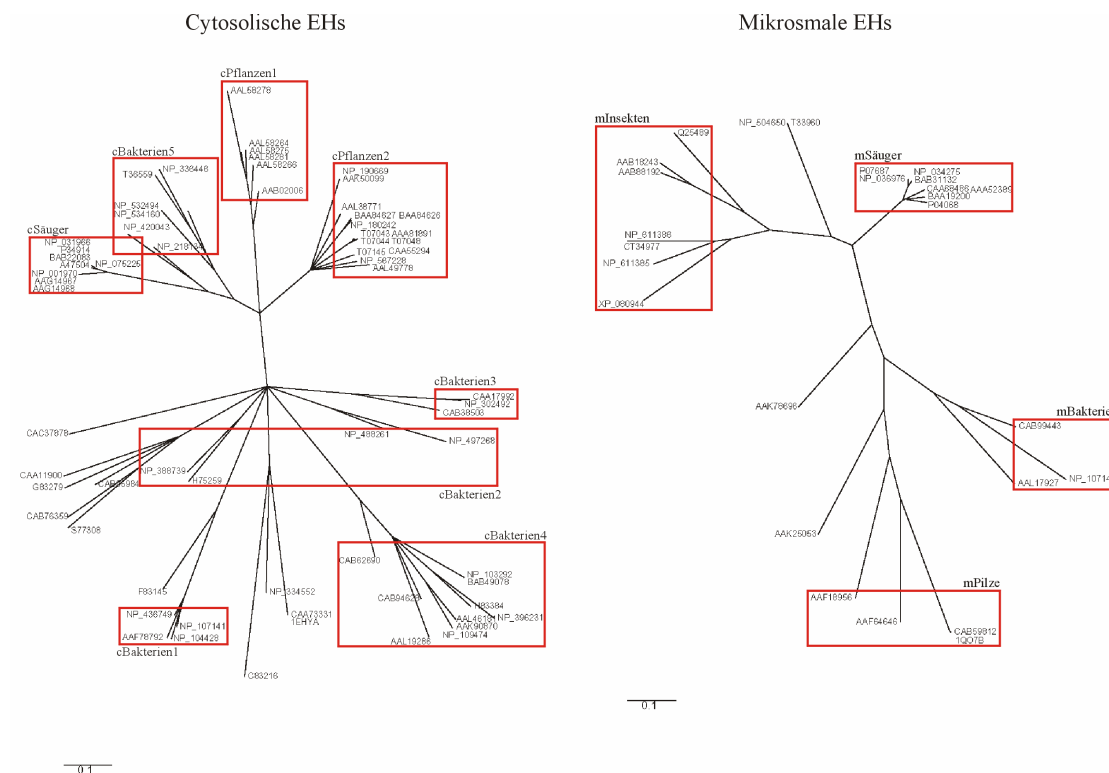


Abbildung 21

Für das Design von CODEHOP-Primern eingesetzte homologe Familien.

Über Multisequenz-Alignments der jeweiligen Sequenzen einer homologen Familie, wurden die zu erwartenden Fragmentgrößen der Amplifikate abgeschätzt:

Tabelle 2: Erwartete Fragmentgröße bei Amplifikation mit CODEHOP-Primern.

Homologe Familie	erwartete Fragmentgröße
mSäuger	~750 bp
mInsekten	~1150 bp
mPilze	~1000 bp
mBakterien	~300 bp
cSäuger	~1350 bp
cPflanzen1	~750 bp
cPflanzen2	~600 bp
cBakterien1	~750 bp
cBakterien2	~650 bp
cBakterien3	~750 bp
cBakterien4	~300 bp
cBakterien5	~750 bp

Die praktische Anwendung der Primer erfolgte stichprobenartig mit den Primern der Gruppe cBakterien5.

4.3.1 Amplifikation der EH aus *Streptomyces antibioticus* Tü4

Wie in der Einleitung beschrieben, wurde bei dem Stamm *Streptomyces antibioticus* Tü4 eine Aktivität gegen Styroloxid nachgewiesen (150).

Mit herkömmlichen degenerierten Primern, erhalten aus einem Multisequenz-Alignment mehrerer EHs, konnte aus genomischer DNA von *Streptomyces antibioticus* ein ca. 250 bp Fragment amplifiziert werden. Dieses Fragment wurde ohne zu klonieren sequenziert und eine ca. 100 bp lange Sequenz mit signifikanter Ähnlichkeit zu Epoxidhydrolasen erhalten (Persönliche Mitteilung Dr. Eckart Bonacker). Das Fragment war mit der Schnittstelle *Bam*H I versehen und wurde in den Vektor pUC18 ligiert. Eine Sequenzierung dieses Konstruktes mit M13-Primern ergab eine 236 bp lange Sequenz, die bei einer BLAST-Suche eine hohe Ähnlichkeit zu der putativen EH aus *Streptomyces coelicolor* A2(3) (T36559) zeigte. Mit den Primern, die zur Amplifikation des 236 bp Fragmentes verwendet wurden, ließ sich aus genomischer DNA von *Streptomyces coelicolor* A2(3) ein Fragment gleicher Größe

4.3.1.1 Fermentation und Reinigung

Eine Möglichkeit, die N-terminale Sequenz der EH aus *Streptomyces antibioticus* Tü4 zu ermitteln, ist die N-terminale Ansequenzierung mit dem Edman-Abbau. Hierfür wird gereinigtes Enzym benötigt. Für die Reinigung wurde eine größere Menge Zellmasse fermentativ hergestellt. Die Reinigung der EH aus *Streptomyces antibioticus* Tü4 wurde im Labor von Herrn Dr. Friedrich bei der BASF durchgeführt. Die EH aus *Streptomyces antibioticus* Tü4 wurde in einem vierstufigen Prozess aufgereinigt. Der durch den Zellaufschluss erhaltene Rohextrakt wurde durch eine Anionenaustausch Chromatographie auf einer Sephadex FF-Säule als ersten Reinigungsschritt vorgereinigt. Anschließend wurden die aktiven Fraktionen gepoolt und durch eine Ammoniumsulfatfällung das Volumen verringert. Die Molekularsieb Chromatographie ergab aktive Fraktionen mit einer Proteingröße von ungefähr 40 kDa. Als letzte Reinigungsschritte wurden eine hochauflösende Anionenaustausch Chromatographie auf einer MonoQ-Säule und eine Hydrophobe Interaktionschromatographie durchgeführt. In Abbildung 23 ist das SDS-Gel nach dem letzten Reinigungsschritt dargestellt. Die markierte Protein-Bande bei ca. 40 kDa wurde aus dem Gel isoliert, N-terminal ansequenziert und einem Trypsinverdau unterzogen.

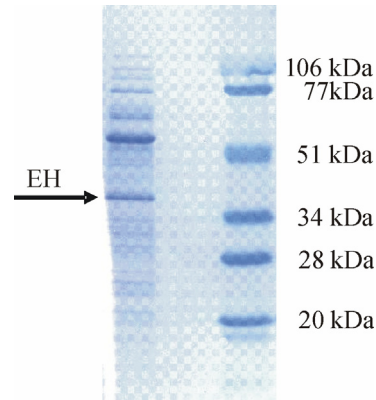


Abbildung 23

Wertfraktion nach dem letzten Reinigungsschritt.

4.3.1.2 N-terminale Ansequenzierung

Die N-terminale Ansequenzierung wurde im Labor von Herrn Dr. Friedrich bei der BASF durchgeführt.

Durch die N-terminale Ansequenzierung nach Edman konnten 18 N-terminale AS identifiziert werden:

N-Terminus-(V)EHRMVD/(V)VNC/(N)IRLH/(L)I/(N)AE/(N)E...-C-Terminus

Eine mit der Sequenz VEHRMVDVNCIRLHIAEE durchgeführte BLAST-Suche ergab signifikante Übereinstimmungen zur N-terminalen Sequenz einer EH aus *Streptomyces avermitilis* (88% Ähnlichkeit, 77% identische AS) und zur N-terminalen Sequenz einer EH aus *Arabidopsis thaliana* (83% Ähnlichkeit und 61% identische AS) (Abbildung 24).

Putative EH aus *Streptomyces avermitilis*:

Score = 37.5 bits (78), Expect = 0.033
Identities = 14/18 (77%), Positives = 16/18 (88%)

```
Query: 1 VEHRMVDVNCIRLHIAEE 18
      V HRM+DVN IRLHIAE+
Sbjct: 10 VNHRMIDVNGIRLHIAEQ 27
```

Putative EH aus *Arabidopsis thaliana*:

Score = 30.6 bits (62), Expect = 3.8
Identities = 11/18 (61%), Positives = 15/18 (83%)

```
Query: 1 VEHRMVDVNCIRLHIAEE 18
      ++HRMV VN I +HIAE+
Sbjct: 4 IDHRMVSVNGITMHIAEK 21
```

Abbildung 24

Ergebnis der BLAST-Suche mit den 18 AS der N-terminalen Sequenz.

Aus der Sequenz des N-Terminus wurde ein Primer (N-term_F) abgeleitet. Aufgrund der hohen Ähnlichkeit des 771 bp Fragments der *Streptomyces antibioticus* Tü4 EH zur putativen

EH aus *Streptomyces coelicolor* A2(3), wurde ein für die EH aus *Streptomyces coelicolor* spezifischer Reverse-Primer (Sc_R) für die PCR eingesetzt. Bei dieser PCR wurden zwei Banden mit der erwarteten Größe von ca. 1000 bp erhalten. T/A-Klonierung in den Vektor pCR[®]2.1-TOPO und anschließende Sequenzierung des Fragmentes mit M13-Primern ergaben keine Übereinstimmungen mit bekannten EHs. Wurde nur der Primer Sc_R für die PCR eingesetzt ergab sich das gleiche Ergebnis, so dass dieser Primer nicht für die Amplifikation der EH aus *Streptomyces antibioticus* Tü4 geeignet war.

4.3.1.3 Amplifikation von N-Terminus und CODEHOP-Fragment

Um das 771 bp-Fragment um den N-Terminus zu erweitern, wurden der aus dem N-Terminus abgeleitete degenerierte Primer (N-term_F) und der CODEHOP-Reverse Primer (cB5_R) verwendet. Das amplifizierte Fragment wies eine Größe von ca. 900 bp auf. T/A-Klonierung in den Vektor pCR[®]2.1-TOPO und anschließende Sequenzierung des Fragmentes mit M13-Primern ergaben 306 AS der gesuchten Sequenz.

4.3.1.4 Aufbau des C-Terminus

Die EH aus *Streptomyces antibioticus* Tü4 zeigt eine hohe Übereinstimmung mit der putativen EH aus *Streptomyces avermitilis* (NP_824987) (Abbildung 25). Daher wurde der C-Terminus dieser EH benutzt, um das 306 AS-Fragment der EH aus *Streptomyces antibioticus* (CODEHOP-Fragment + N-Terminus) zu vervollständigen. Es wurde ein Primer kreiert, der den N-Terminus der *Streptomyces avermitilis* EH und einen Überhang in die Sequenz des *Streptomyces antibioticus* EH enthielt (Abbildung 25). Als Template wurde ein Konstrukt aus CODEHOP-Fragment und N-Terminus im Vektor pCR[®]2.1-TOPO benutzt. Das amplifizierte Fragment wies eine Größe von ca. 1000 bp auf. T/A-Klonierung in den Vektor pCR[®]2.1-TOPO und anschließende Sequenzierung des Fragmentes mit M13-Primern ergaben mit 324 AS die vollständige Sequenz der EH.

```

S. antibioticus -----MVEHRMVDVNGIRLHIAEEGEGPLVLLCHFPESWHSWHRQFGPLAAAGFRV 52
S. avermitilis MAHPYARIMVNRHMIDVNGIRLHIAEQGEGPLVLLHCFPESSYSWRHQFGPLAAAGFRV 60

S. antibioticus VAPDQRCYGRSDHPEAVDAYTILHLVGDVVGLIRELGEEKAYVVCHDWCAPVAWHTALLR 112
S. avermitilis VAPDQRCYGRSDRPESVDAYTILHLVGDVIGLIHALGEETAYVVCHDWCAPVAWHTALLR 120

S. antibioticus PDLVRGVAGLSVPPPPFRGSRPPLPAMDRMFGGRFWNYFNRPGVADAEFAEDTRTALRKF 172
S. avermitilis PDVVRGVAGLSVPPPPFRGERPPLDAMDDEMFDGQFWNYFAQPGVADAEFGRDARGSLRKF 180

S. antibioticus FYWASGDAPGAGR-KQPLVDPERGWLADMDAPEVLPEWFTTEEDLDALTESFSQGFTGALN 231
S. avermitilis LYSASGDAPGAGSGRQPLVAPGRGFLDGMPPDPEVLPGWITEADLDVFAEEFAPGFTGALN 240

S. antibioticus WRLNLDNRNWELTAPWHGAVVTPPALYVYGDRDVVPAFFPGTPELIERLPSLMPNLWREPLK 291
S. avermitilis WRLNLDNRNWELTAAWQGAVSPPALYMYGDRDAVPAFFPGTAELIEKLPALMPNLRREPLV 300

S. antibioticus LAGCGHWPQQERPKG----- 306
S. avermitilis LPGCGHWTQQERPNEVNAALVDFLTGLRGEQRI 333 C-Terminus
QQERPKGVNAALVDFLTGLRGEQRI
Primer: C-Term_R

```

Abbildung 25

Alignment der EH aus *Streptomyces antibioticus* Tü4 der putativen EH aus *Streptomyces avermitilis*. Farblich gekennzeichnet wurden die AS des Oxyanionholes (blau), Protonendonatoren (grün), der katalytischen Triade (rot), das GXGXS-Motiv (violett), die Gly des GX-Motives und des Nucleophilic Elbow (oliv). In der AS Sequenz des Primers C-Term_R wurden die Sequenzbereiche der *Streptomyces antibioticus* EH (türkis) und der *Streptomyces avermitilis* EH (gelb) farblich markiert.

Abbildung 26 enthält eine schematische Zusammenfassung der Amplifikation der EH aus *Streptomyces antibioticus*.

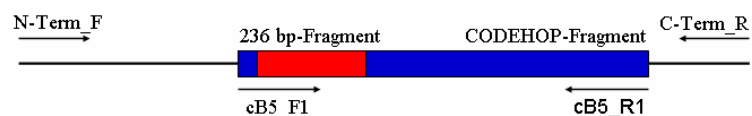


Abbildung 26

Schematische Darstellung der Amplifikation der EH aus *Streptomyces antibioticus*. Dargestellt wurden das 236 bp-Fragment (rot), das CODEHOP-Fragment (blau) und die zur Amplifikation verwendeten Primer: cB5_F1 und cB5_R1 (CODEHOP), N-Term_F (N-Terminus) und C-Term_R (C-Terminus).

4.3.1.5 Trypsin-Verdau

Als Nachweis, dass das gereinigte Protein aus *Streptomyces antibioticus* und die amplifizierte Sequenz übereinstimmen, wurde ein tryptischer Verdau durchgeführt. Der tryptische Verdau und die anschließende N-terminale Sequenzierung wurde im Labor von Herrn Dr. Friedrich bei der BASF durchgeführt und ergab folgende Fragmente:

Fragment 1: MVNL/(V)NG/(D)IR

Fragment 2 : GVAGLSVPPPPFR

Die Fragmente konnten, wie in Abbildung 27 zu sehen, in der 306 AS umfassenden Sequenz identifiziert werden.

```

N-Terminus
S.antibioticus MVEHRMVDVNGIRLHIAEEGEGPLVLLCHGFPESWHSWHRQFGPLAAAGFRVVAPDQRGY 60
                MVNVNGIR Fragment 1
S.antibioticus GRSDHPEAVDAYTILHLVGDVVGLIRELGEEKAYVVGHDWGAPVAWHTALLRPDLVRGVA 120
                GVA
S.antibioticus GLSVPPPPFRGSRPPLPAMDRMFGGRFYWNYFNRPGVADAEFAEDTRTALRKFFYWASGDA 180
                GLSVPPPPFR Fragment 2
S.antibioticus PGAGRKQPLVDPERGWLADMADPEVLPEWFTTEEDLDALTESFSQGFTGALNWYRNLDNRW 240
S.antibioticus ELTAPWHGAVVTPPALYVYGDRDVVPAFFPGTPELIERLPSLMPNLWREPLKLAGCGHWPQ 300
S.antibioticus QERPKG 306

```

Abbildung 27

306 AS umfassende Sequenz, bestehend aus N-Terminus (gelb) und CODEHOP-Fragment. Grün markiert wurden die aus dem tryptischen Verdau erhaltenen Fragmente Fragment 1 und Fragment 2.

4.3.2 Amplifikation des EH Fragments aus *Rhodococcus ruber*

Die Spezies *Rhodococcus ruber* (LU760) gehört wie die Gattung *Streptomyces* zur Ordnung der *Actinomycetales*, daher wurden die CODEHOP-Primer der Gruppe cBakterien5 verwendet, um eventuelle EH-Fragmente zu amplifizieren. Aus genomischer DNA konnten zwei Fragmente der Größen 700 und 900 bp amplifiziert werden. T/A-Klonierung und Sequenzierung dieser Fragmente (Diplomarbeit Hendrik Schewe) zeigte, dass das 700 bp-Fragment für eine putative EH kodiert. Es wurden 703 bp sequenziert, eine anschließende BLAST-Suche ergab eine eindeutige Homologie zu EHs. Wie in Abbildung 28 zu sehen, enthält die 215 AS umfassende translatierte Sequenz das Nukleophil und die katalytische Säuregruppe der katalytischen Triade, einen Teil des Oxyanionholes, und das als Protonendonator wirkende, absolut konservierte Tyr der Cap-Domäne.

```

Agrobacterium radiobacter  MTIRRPEDFKHYEVQLPDVKIHYVREGAGPTLLLLHEWPGFWWEWSKVIGPLAEH-YDVI 59
Rhodococcus ruber          -----THAIWRLQVPALAARGFRVI 20

Agrobacterium radiobacter  VPDLRGFGDSEKPDLNDSLKYSLDKAADDQAALLDALGIEKAYVVGHDFPAAIVLHKFIRK 119
Rhodococcus ruber          APDMRYCKTDAP--PAAGAYAVEFLCADILGLMDQLGLEKASLVGHDFPAGLVGVQLCMN 78

Agrobacterium radiobacter  YSDRVIKAAIFDPIQPDFGVPVYFGLGHVHESWWSQFHQLDMAVEVVGSSREVCKKYFKHF 179
Rhodococcus ruber          APQRFERFVALSTGHPAAIANAG-----VTQQMRFWYILGFLMPVVAEH--AIRANDWF 130

Agrobacterium radiobacter  FDHWSYRDELLTEEELEVHVDNCMKPDNIHGGFNYRANIRPDAALWTDLDHTMSDLPVT 239
Rhodococcus ruber          FLR----QMTRRPEQVAIWRAALEPEGRLTAALNYRANLK----LGRPHHWRPVEVPVM 182

Agrobacterium radiobacter  MIWGLGDTCVPYAPLIEFVPKYYSNYTMETIEDCGHFLMVEKPEIAIDRIKTAFR 294
Rhodococcus ruber          GVWSDRDPALGKQMLDSVHHCRAFRYERLSG----- 215

```

Abbildung 28

Alignment des CODEHOP-Fragments der *Rhodococcus ruber* EH und der EH aus *Agrobacterium radiobacter*. Farblich gekennzeichnet wurden die AS des Oxyanionholes (blau), der Protonendonatoren (grün), der katalytischen Triade (rot), des GXGXS-Motives (violett), das Gly des GX-Motives und des Nucleophilic Elbows (oliv).

4.4 Teil 2: Erstellung und Auswertung der EH/HD Datenbank

Die EH/HD Datenbank stellt eine Erweiterung des 93 EH Alignments unter Einbeziehung weiterer Enzymklassen und Strukturinformationen dar. Eine familienspezifische Datenbank bündelt und organisiert große Mengen an Sequenz-, Struktur- und Annotationsinformationen und ist daher ein nützliches Werkzeug bei der Erforschung dieser Enzyme.

Die EH/HD Datenbank ist im World Wide Web unter folgender Adresse erreichbar:

<http://www.led.uni-stuttgart.de>.

4.4.1 Die EH/HD Datenbank

Die EH/HD Datenbank beruht auf dem Algorithmus der LED (Lipase Engineering Database) (45). Sie beruht auf einem erweiterbaren Datenmodell DWARF (**D**ata **W**arehouse. for **P**rotein Families (46)) und einer relationalen Datenbank, welche Informationen betreffend Sequenz, Struktur und Annotation von Proteinen, abrufbar aus den öffentlichen Datenbanken GenBank (16) und PDB (17), einbindet.

Basierend auf phylogenetischen Analysen des 93 EH Alignments und 18, durch Stichwortsuche erhaltenen, HDs (Tabelle 23), wurden repräsentative Sequenzen ausgewählt (Tabelle 24). Mit diesen Sequenzen wurden BLAST-Suchen (3) mit einem Cutoff von $E = 10^{-10}$ durchgeführt. Die hiermit ermittelten homologen Sequenzen wurden automatisch vom System extrahiert. Nach manueller Bereinigung des Datensatzes enthält die Datenbank 397 Sequenzen unterschiedlicher α/β -Hydrolasen, 305 Protein- und 51 Struktureinträge. 48% aller Proteine sind putativ. Die Sequenzen wurden in zwei Superfamilien und 14 homologe Familien eingeteilt. Innerhalb einer Superfamilie sind die katalytischen AS konserviert, während homologe Familien außerdem eine globale Sequenzähnlichkeit aufweisen. Sequenzeinträge mit einer Sequenzidentität von mehr als 95% wurden zu einem Proteineintrag zusammengefasst.

4.4.2 Superfamilien

4.4.2.1 Cytosolische Hydrolasen Superfamilie (cHS)

Die größere der zwei Superfamilien enthält 340 Sequenzen, 267 Protein- und 57 Struktureinträge, die in 13 homologe Familien aufgeteilt sind. Die cHS besteht aus löslichen, cytosolischen Proteinen (c: cytosolisch), welche die katalytische Säuregruppe normalerweise an der β 7- bzw. in seltenen Fällen an der β 6-Position enthalten. Das Nukleophil dieser Enzyme ist entweder ein Asp oder ein Ser. Als Enzymfamilien enthalten sind: EHs, HDs, Haloacid Dehalogenasen, Non-heme Peroxidasen, 2-hydroxymuconic semialdehyd Hydrolasen, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid Hydrolasen, Prolyl-Aminopeptidasen, Esterasen, Luciferasen, welche in die folgenden homologen Familien eingeteilt wurden:

Eine Gruppe homologer Familien, die ausschließlich EHs und nicht annotierte Proteine enthalten, sind: „lösliche Säugetier EHs“, „lösliche Pflanzen EHs“, „lösliche bakterielle EHs I“ und „lösliche bakterielle EHs II“ und „lösliche EHs (β 6)“. HDs und verwandte nicht annotierte Hydrolasen sind in den homologen Familien „lösliche HDs (β 6)“ und „lösliche HDs (β 7)“ zusammengefasst. Eine weitere Gruppe von Asp-Hydrolasen wurde in der Familie lösliche Haloacid Dehalogenasen abgelegt. Die unterschiedlichen Ser-Hydrolasen und verwandte nicht annotierte Proteine wurden auf die Familien „lösliche Non-heme Peroxidasen“, „Esterasen/Lipasen/Peptidasen“, „Meta-cleavage compound Hydrolasen I+II“ verteilt. Eine Gruppe nicht zuzuordnender Proteine wurde unter dem Namen „Verschiedenes“ zusammengefasst.

4.4.2.2 Mikrosomale Hydrolasen Superfamilie (mHS)

Diese Superfamilie enthält 57 Sequenzen, 48 Protein- und 2 Struktureinträge, die in eine homologe Familie zusammengefasst wurden. Die mHS besteht aus mikrosomalen Proteinen (m: mikrosomal), welche die katalytische Säuregruppe an der β 7-Position enthalten, das Nukleophil dieser Enzyme ist ein Asp. Die mHS enthält ausschließlich als EHs annotierte Einträge.

4.4.3 Multisequenz-Alignments

Für beide Superfamilien und jede homologe Familie wurden Multisequenz-Alignments erstellt. Funktionell relevante AS wurden durch eine Farb-Codierung annotiert. Markiert wurden die AS der katalytischen Triade, des Oxyanionholes, des GXGXS-Motivs und bei EHs die AS der Protonendonatoren und für HDs die AS der Halogenid-Kralle.

4.4.4 Phylogenetische Analyse

Die Einteilung in Superfamilien und homologe Familien beruht auf einer phylogenetischen Analyse, durchgeführt mit der distanzbasierten Neighbor-Joining Methode. Der distanzbasierte phylogenetische Baum der cHS zeigt eine deutliche Aufteilung in Asp- und Ser-Hydrolasen. Die phylogenetischen Bäume der einzelnen homologen Familien wurden mit der merkmalsbasierten Methode Tree-Puzzle berechnet. Bootstrap-Werte beschreiben die Relevanz jeder Verzweigung der phylogenetischen Bäume. Diese phylogenetischen Bäume zeigen deutlich die enge Verwandtschaft unter den eukaryotischen Familien, im Gegensatz zu den mikrobiellen Familien, welche eine hohe Diversität auf Sequenzebene aufweisen.

4.4.5 Strukturelle Untersuchungen an α/β Hydrolasen

Alle Enzymfamilien, die in den Datenbanken LED und EH/HD abgelegt sind, gehören zur Familie der α/β Hydrolasen. Die einzige Schnittmenge beider Datenbanken wird durch die Non-heme Peroxidasen gebildet. Um die Gemeinsamkeiten und Unterschiede dieser großen Anzahl an verschiedenen α/β Hydrolasen zu untersuchen, wurde eine Analyse der verfügbaren Strukturdaten durchgeführt. In Tabelle 3 wurden die verwendeten PDB-Einträge aufgelistet. Untersuchungen des „ α/β Hydrolase Folds“ ergaben, dass Variationen der Anzahl von β -Strands und α -Helices in den meisten der untersuchten Ser-Hydrolase Strukturen vorkamen. Asp-Hydrolasen, gastrische Lipasen, Non-heme Peroxidasen, Prolyl-Aminopeptidasen, 2-Hydroxy-6-oxo-7-methylocta-2,4-dienoate Hydrolasen und alle Lipasen des GGGX-Types (111) wiesen einen kanonischen Fold auf. Nicht alle Strands und Helices waren betroffen. Abbildung 29 verdeutlicht das substraktive Verhalten der untersuchten Strukturen. Die betroffenen β -Strands wurden aus der Struktur entfernt, α -Helices jedoch durch Loop-Regionen ersetzt. Ausgehend vom kanonischen Fold der Asp-Hydrolasen konnten verschiedene Subtraktionsgruppen ermittelt werden. Häufig betroffen waren die β -Strands $\beta 1$ und $\beta 2$ und die α -Helices $\alpha 4$ und $\alpha 5$. Die β -Strands $\beta 1$ und $\beta 2$ enthalten keine katalytisch wichtigen AS. Nicht betroffen sind die Strands $\beta 3$ bis $\beta 8$, so dass die katalytische

Maschine strukturell in allen α/β Hydrolasen erhalten bleibt. In Tabelle 4 sind die untersuchten homologen Familien und die jeweilige Anzahl an vorkommenden β -Strands und α -Helices zusammengestellt.

Tabelle 3: Zum Strukturvergleich verwendete PDB-Einträge aus LED und EH/HD.

Homologe Familie/Enzymklasse	PDB-Eintrag
GX-Typ	
Haloalkan Dehalogenasen	1EDE
Epoxidhydrolasen	1EHY (remodelliert)
2-Hydroxy-6-oxo-7-methylocta-2,4-dienoate Hydrolasen	1IUN
Prolyl-Aminopeptidasen	1MTZ
Non-heme Peroxidasen	1BRO
<i>Burkholderia</i> Lipasen	4LIP
<i>Bacillus subtilis</i> Lipasen	1I6WA
<i>Candida antarctica</i> Lipasen	4LBS
<i>Fusarium</i> Cutinasen	2CUT
Gastrische Lipasen	1K8QA
Pankreatische Lipasen	1RP1
<i>Rhizomucor miehei</i> Lipasen	4TGL
<i>Moraxella</i> Lipasen	1JFRA
GGGX-Typ	
Acetylcholinesterasen	1C2BA
<i>Bacillus</i> Esterasen	1C7JA
„Mammalian bile salt activated“ Lipasen	1AQLA
Säugetier Carboxylesterase	1K4YA
<i>Candida rugosa</i> Lipasen	1CLEA
Hormon sensitive Lipasen	1EVQA

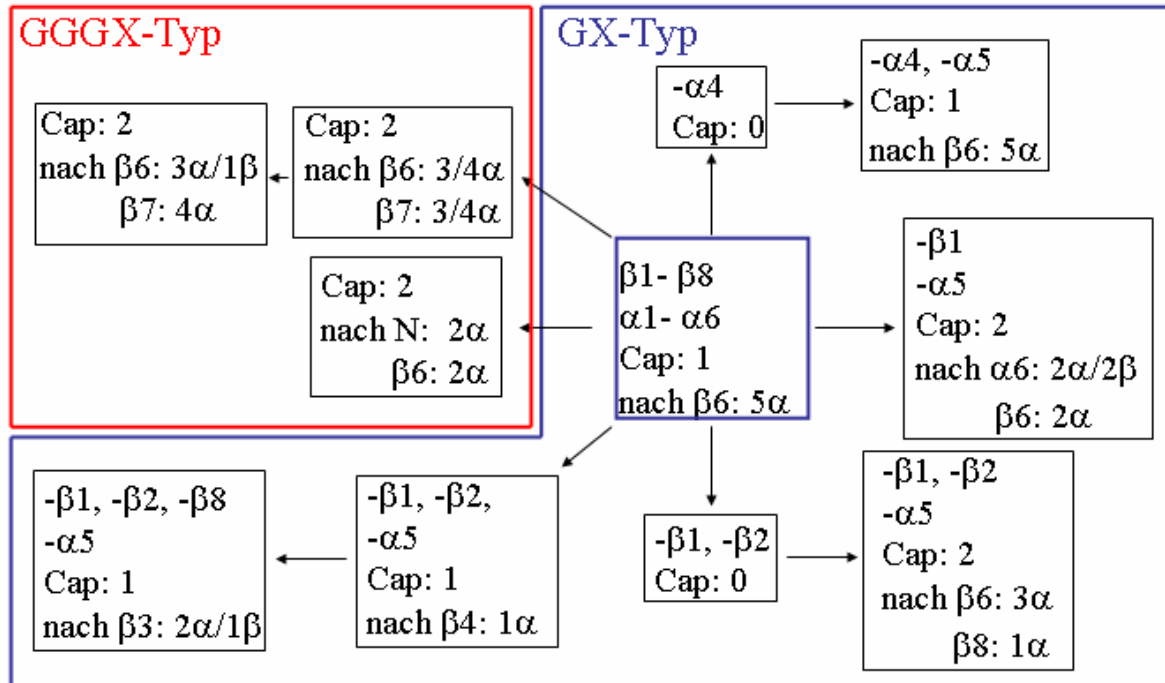


Abbildung 29

Ausgehend vom kanonischen α/β Hydrolase Fold (blaues Rechteck in der Mitte) sind die vorkommenden Subtraktionsgruppen dargestellt (schwarze Rechtecke). Blau umrandet wurden die Enzymfamilien des GX-Typs und rot umrandet die Familien des GGGX-Typs.

Tabelle 4: Variationen des α/β Hydrolase Folds der Strukturen von LED und EH/HD Datenbank.

Homologe Familie/Enzymfamilie	α/β Hydrolase Fold	Anzahl β -Strands	Anzahl α -Helices
GX-Typ			
Haloalkan Dehalogenasen	komplett	8	6
Epoxidhydrolasen	komplett	8	6
2-Hydroxy-6-oxo-7-methylocta-2,4-dienoate Hydrolasen	komplett	8	6
Prolyl-Aminopeptidasen	komplett	8	6
Non-heme Peroxidasen	komplett	8	6
<i>Burkholderia</i> Lipasen	nicht komplett	6	5
<i>Bacillus subtilis</i> Lipasen	nicht komplett	6	6
<i>Candida antarctica</i> Lipasen	nicht komplett	7	5
<i>Fusarium</i> Cutinasen	nicht komplett	5	5
Gastrische Lipasen	komplett	8	6
Pankreatische Lipasen	nicht komplett	8	4
<i>Rhizomucor miehei</i> Lipasen	nicht komplett	5	3
<i>Moraxella</i> Lipasen	nicht komplett	8	5
GGGX-Typ			
Acetylcholinesterasen	komplett	8	6
<i>Bacillus</i> Esterasen	komplett	8	6
„Mammalian bile salt activated“ Lipasen	komplett	8	6
Säugetier Carboxylesterase	komplett	8	6
<i>Candida rugosa</i> Lipasen	komplett	8	6
Hormon sensitive Lipasen	komplett	8	6

4.4.6 Topologie des Übergangszustandes

Die Hydrolasen der EH/HD Datenbank setzen unterschiedliche Substrate um und nutzen dafür zwei unterschiedliche Nukleophile (Asp und Ser). Der bei der Reaktion gebildete Übergangszustand ist für die beiden Hydrolase-Typen nicht identisch, während bei Asp-Hydrolasen das Nukleophil die Rolle einer Säure übernimmt und das Substrat als fungiert, sind die Verhältnisse bei Ser-Hydrolasen umgekehrt. Hier ist das Nukleophil ein Alkohol und das Substrat eine Säure. Trotz dieser Unterschiede, bilden beide Hydrolase-Typen ein vergleichbares Ester-ähnliches Acyl-Enzym Intermediat, welches durch die AS des Oxyanionholes stabilisiert wird. Ebenfalls hoch konserviert ist die Struktur des katalytischen Zentrums (Tabelle 5), obwohl der Abstand der Position des Oxyanions in Intermediat im Vergleich zur C_{α} -Position des Nukleophils als Konsequenz der Inversion der Esterbindung bei Asp-Hydrolasen 3 Bindungslängen und bei Ser-Hydrolasen 4 Bindungslängen beträgt (Abbildung 30).

Tabelle 5: Abstände zwischen Nukleophil, katalytischem His und Oxyanionhole

$C_{\alpha}^{\text{His}}-C_{\alpha}^{\text{Asp/Ser}}$:	$8.26 \text{ \AA} \pm 0.27 \text{ \AA}$
$C_{\alpha}^{\text{His}}-N^{\text{Ox. hole}}$:	$10.55 \text{ \AA} \pm 0.45 \text{ \AA}$
$C_{\alpha}^{\text{Asp/Ser}}-N^{\text{Ox. hole}}$:	$5.83 \text{ \AA} \pm 0.35 \text{ \AA}$

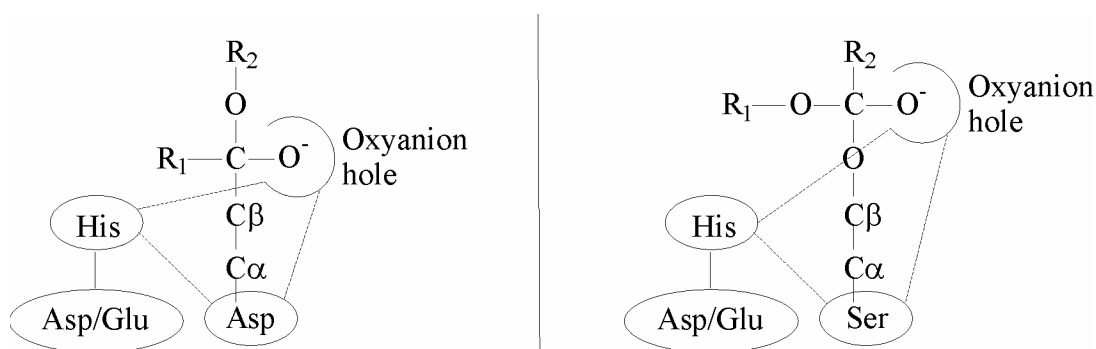


Abbildung 30

Schematische Darstellung des Übergangszustandes von Asp-Hydrolasen (links) und Ser-Hydrolasen (rechts). Die AS der katalytischen Triade sind durch Kreise gekennzeichnet. Die Abstände zwischen Nukleophil, katalytischem His und Oxyanionhole wurden durch gestrichelte Linien dargestellt.

4.4.7 Anker Konzept

Das Oxyanionhole besteht aus zwei AS, deren Backbone Amide über Wasserstoff-Brücken die negative Ladung des Übergangszustandes stabilisieren. Während die Position der zweiten, direkt auf das Nukleophil folgenden AS stabil ist, befindet sich die erste AS (HGX-Motiv) auf einem Loop. Für die Funktion des Oxyanionholes ist eine exakte Ausrichtung der beiden Backbone Amide erforderlich. Eine Fixierung des HGX-Motivs kann über Wechselwirkung eines „Ankers“ mit der Seitenkette der Oxyanionhole AS stattfinden (111). Der Anker entspricht in seinen Eigenschaften der Oxyanionhole AS. Ist diese AS hydrophob, so ist es auch der Anker, ist sie hydrophil, ist auch der Anker hydrophil. Für die Einträge der LED wurden drei verschiedene Anker-Konzepte ermittelt (Tabelle 6) (111):

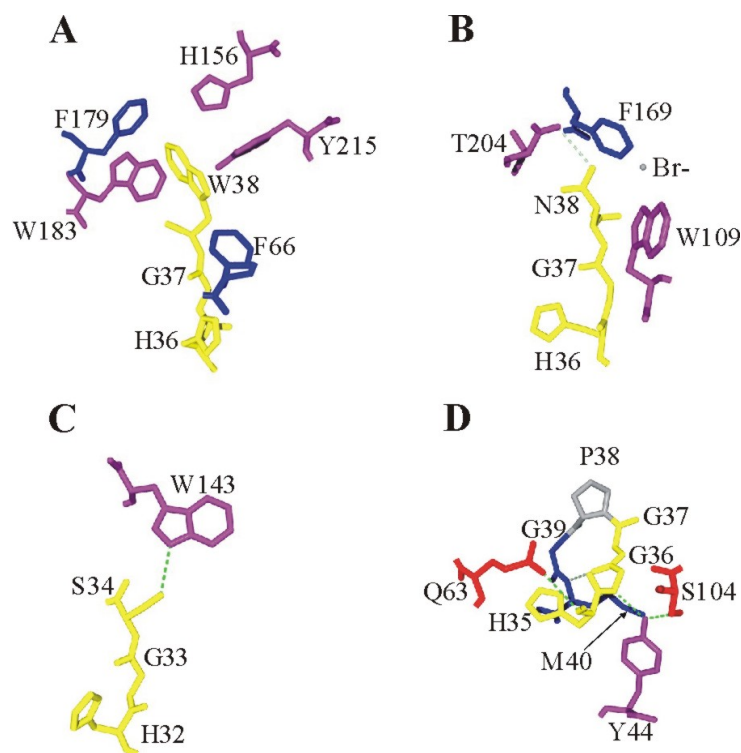
Tabelle 6: Anker-Konzepte

- 1) Hydrophiler Anker
- 2) Hydrophober Anker
- 3) Hydrophobe Tasche

Dieses Konzept wurde auf die Familien der EH/HD Datenbank angewendet. Obwohl sich die Strukturen der unterschiedlichen Enzymfamilien sehr stark ähneln, weisen sie unterschiedliche Anker-Konzepte auf. Während, wie in Abbildung 31 dargestellt, bei EHs und Non-heme Peroxidasen die Stabilisierung über eine hydrophobe Tasche realisiert ist, weisen 2-Hydroxy-6-oxo-7-methylocta-2,4-dienoate Hydrolasen einen hydrophilen Anker auf. Bei HDs befindet sich ein Backbone-Carbonyl in Wasserstoff-Brücken-Reichweite zur hydrophilen Oxyanionhole AS. Das über die Dauer der Reaktion gebundene Chlorid-Ion kann ebenfalls die Position der Oxyanionhole AS beeinflussen, somit liegt hier ein Sonderfall des hydrophilen Anker-Typs vor. In Tabelle 7 wurden die unterschiedlichen Anker-Typen zusammengefasst. Die Struktur der Prolyl-Aminopeptidase weist keine Stabilisierung der Oxyanionhole AS durch einen der Anker-Typen auf.

Tabelle 7: Anker-Typen der untersuchten Strukturen der EH/HD.

Enzym-Familie	Pdb	HGX-Motiv	Anker-AS	Anker-Typ
EHS (β 7)	1EHY	HGW	F66, H156, F179, T183, Y215	hydrophobe Tasche
HD (β 7)	1EDB	HGE	V219 (Backbone-Carbonyl), Cl ⁻	hydrophiler Anker (Sonderfall)
HD (β 6)	1IZ8	HGN	T204 (Backbone-Carbonyl), Cl ⁻	hydrophiler Anker (Sonderfall)
Non-heme Peroxidase	1BRO	HGF	F61, F159, Y160, F163, W205	hydrophobe Tasche
2-Hydroxy-6-oxo-7-methylocta-2,4-dienoate hydrolase	1IUN	HGS	Trp 143	hydrophiler Anker
Prolyl-Aminopeptidase	1MTZ	HGG		Kein Anker-Konzept

**Abbildung 31**

Dargestellt wurden die unterschiedlichen Anker-Typen der untersuchten Strukturen der EH/HD. A: EH aus *Agrobacterium radiobacter* (1EHY), hydrophobe Tasche, B: HD aus *Spingomonas paucimobilis* (1IZ8), hydrophiler Anker, C: 2-Hydroxy-6-oxo-7-methylocta-2,4-dienoate Hydrolase aus *Pseudomonas fluorescens* (1IUN), hydrophiler Anker und D: Prolyl-Aminopeptidase aus *Thermoplasma acidophilum* (1MTZ).

Farbig dargestellt wurden: das HGX-Motiv (gelb), hydrophobe AS (blau), polare AS (rot), hydrophobe und polare AS (violett). Die gestrichelten Linien symbolisieren Wasserstoff-Brücken.

In Tabelle 8 sind die Anker-Typen der LED Enzymfamilien aufgelistet (111).

Tabelle 8: Anker-Typen der LED Enzymfamilien.

Homologe Familie/Enzymfamilie	Anker Konzept
„Non-heme“ Haloperoxidase	hydrophobe Tasche
<i>Burkholderia</i> Lipasen	hydrophober Anker
<i>Bacillus subtilis</i> Lipasen	hydrophober Anker
<i>Candida antarctica</i> Lipasen	hydrophiler Anker
<i>Fusarium</i> Cutinasen	hydrophiler Anker
Gastrische Lipasen	hydrophober Anker
Pankreatische Lipasen	hydrophober Anker
<i>Rhizomucor miehei</i> Lipasen	hydrophiler Anker
<i>Moraxella</i> Lipasen	hydrophober Anker

4.4.8 Anker-Struktur Beziehung

Die Struktureigenschaften und die Anker-Typen der untersuchten Strukturen wurden in Abbildung 32 zusammengefasst. Hydrophobe und hydrophile Anker-Typen treten bei vollständigem sowie bei unvollständigem α/β Hydrolase Fold auf. Auffällig ist, dass hydrophobe Taschen nur bei Proteinen mit vollständigem Fold auftreten, die kein Lid aufweisen. Bei unvollständigem Fold enthielten alle Proteine mit einer geraden Anzahl an β -Strands, hydrophobe Anker eine, während alle Protein mit ungerader Anzahl an β -Strands hydrophile Anker aufwiesen. Des weiteren konnte keine Korrelation mit Substraten oder Funktion des jeweiligen Proteins gefunden werden.

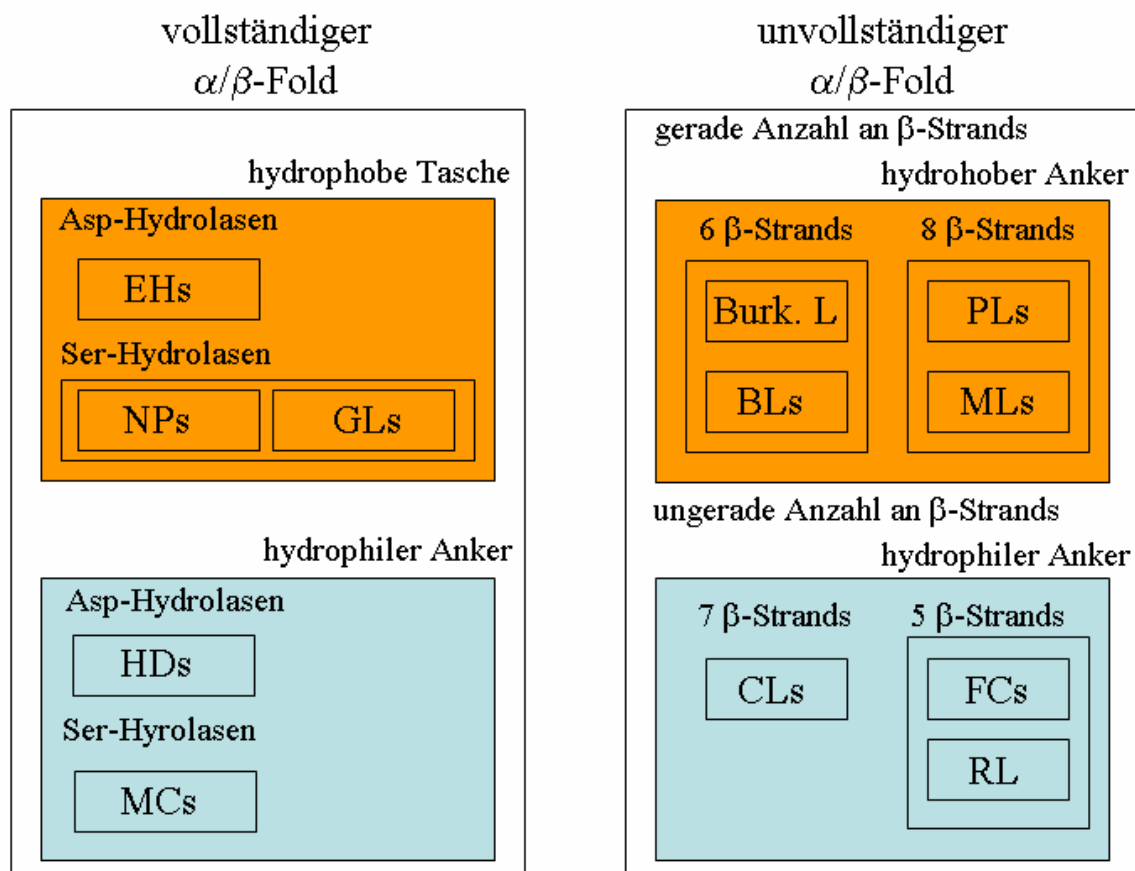


Abbildung 32

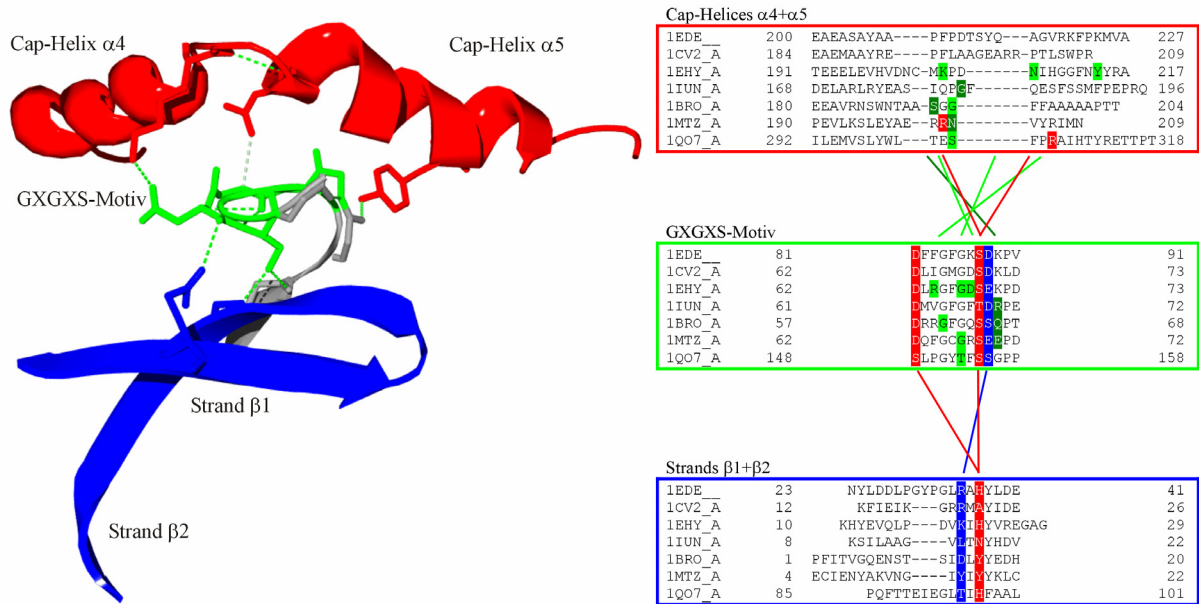
Zusammenhang zwischen Vollständigkeit des α/β Hydrolase Folds und der unterschiedlichen Anker-Typen. Die unterschiedlichen Anker-Typen wurden farbig markiert: hydrophober Anker (orange), hydrophiler Anker (blau). Vollständiger Fold, hydrophob: EHs: Epoxidhydrolasen, NPs: Non-heme Peroxidasen, GLs: Gastrische Lipasen. Vollständiger Fold, hydrophil: HDs, Haloalkan Dehalogenasen, MC: 2- Hydroxy-6-oxo-7-methylocta-2,4-dienoate Hydrolasen. Unvollständiger Fold, hydrophob: Burk.L: *Burkholderia* Lipasen, BLs: *Bacillus subtilis* Lipasen, PLs: Pankreas Lipasen, MLs: *Moraxella* Lipasen. Unvollständiger Fold, hydrophil: CLs: *Candida antarctica* Lipase, FCs: *Fusarium* Cutinase, RL: *Rhizomucor miehei* Lipase.

4.4.9 Das GXGXS-Motiv

Das GXGXS-Motiv (92) ist hoch konserviert und bildet einen strukturell ebenfalls hoch konservierten Loop nach β -Strand β_4 des zentralen β -Faltblattes aus. Der Loop befindet sich in der Struktur zwischen des β -Strands β_1 und β_2 und den α -Helices α_4 und α_5 der Cap-Domäne, wie eine Münze im Münzschaft [„Coin in a coin slot“] (12).

Die Funktion des GXGXS-Motivs scheint an die hoch konservierte Struktur der Enzymfamilien der EH/HD Datenbank gekoppelt zu sein. Die drei Elemente β_1 und β_2 , GXGXS-Motiv und α_4 und α_5 der Cap-Domäne sind durch Wasserstoffbrücken miteinander verbunden. Die β -Strands β_1 und β_2 sind durch Wasserstoffbrücken an das GXGXS-Motiv gekoppelt die von fünf strukturell konservierten AS gebildet werden (Abbildung 33): zwei AS auf Strand β_2 und drei AS im GXGXS-Motiv. Die Verbindung zur der Cap-Domäne ist bei den Strukturen der EH/HD Datenbank variabler gestaltet. Die beteiligten AS befinden sich jedoch immer am Übergang zwischen Cap-Helix α_4 und α_5 und an unterschiedlichen Positionen im GXGXS-Motiv. Eine mögliche Erklärung für diese größere Variabilität ist die funktionelle Diversität der Cap-Domäne: Cap-Domänen enthalten oft AS die an der Substratbindung beteiligt sind (118, 124) und weisen eine größere Variabilität als der hoch konservierte α/β Hydrolase Fold auf.

Das GXGXS-Motiv wirkt wie ein Schloss, das Cap-Domäne und N-terminale katalytische Domäne verbindet (Abbildung 33) und scheint daher eine wichtige strukturelle Rolle zu spielen bei der Stabilisierung und Ausrichtung der Cap-Domäne.

**Abbildung 33**

Links: GXGXS-Motiv (grün), die Strands $\beta 1$ und $\beta 2$ (blau) und die Cap-Helices $\alpha 4$ und $\alpha 5$ (rot) (*Agrobacterium radiobacter* EH). Wasserstoffbrücken wurden durch unterbrochene Linien dargestellt. Rechts: Multisequenz-Alignment von 1EDE (HD), 1CV2_A (HD), 1EHY_A (EH), 1IUN_A (2-hydroxyomuconic semialdehyd Hydrolase), 1BRO_A (Non-heme Bromoperoxidase), 1MTZ (Prolyl-Aminopeptidase) und 1QO7_A (EH). Die Rahmenfarben des Alignments sind identisch zur Färbung der Struktur (links). Im Alignment haben durch Wasserstoffbrücken verbundene AS die gleiche Farbe (Wasserstoffbrücken wurden durch Linien dargestellt).

5 Diskussion

5.1 Systematische Analyse von EHs

5.1.1 Rolle der modularen Architektur

Der α/β Hydrolase Fold ist eines der häufigsten Faltungsmuster und beinhaltet EHs und andere Hydrolasen wie Lipasen, Acetylcholinesterasen, Carboxypeptidasen und Haloperoxidasen (62). Obwohl diese Enzyme unterschiedliche Arten von Reaktionen katalysieren und unterschiedliche Substrate akzeptieren, enthalten alle α/β Hydrolasen eine konservierte Struktur und eine vergleichbare katalytische Triade bestehend aus der Kombination: Nukleophil-His-Säuregruppe (62).

Anhand von EH-Strukturen wurde der modulare Aufbau dieser Enzyme in N-terminale Domäne, N-terminale katalytische Domäne, NC-Loop, Cap-Domäne, Cap-Loop und C-terminale katalytische Domäne eingeteilt (13). Trotz des hoch konservierten Zusammenspiels aus Architektur und Geometrie der katalytischen Maschine, wurden eine weite Bandbreite an Variationen entdeckt. EHs z.B. zeigen zwei solcher Variationen im katalytischen Zentrum: Unterschiede in der Position der katalytischen Säuregruppe, welche sich entweder nach den β -Strands β_6 oder β_7 des α/β Hydrolase Folds befinden kann und die Existenz eines zweiten für die Substratbindung relevanten Tyr in der Cap-Domäne.

Die Cap-Domänen von EHs und HDs enthalten AS, die katalytisch relevant und an der Substratbindung (8), bzw. bei HDs auch an der Halogenidbindung (31) beteiligt sind. Für HDs wurde sogar vorgeschlagen, dass die Cap-Domäne direkt an der Reaktion beteiligt und für die Substratspezifität relevant ist (62).

Cytosolische und mikrosomale EHs unterscheiden sich in ihren Substratspektren (48). Hierfür könnte ein enger Substratkanal verantwortlich sein, der zwischen NC- und Cap-Loop sichtbar ist. Diese Beobachtung wird unterstützt durch die hohe Beweglichkeit des NC-Loops in der Struktur der *Agrobacterium radiobacter* EH, gefolgert aus hohen B-Faktoren für diesen Bereich (96) und der hohen Flexibilität der NC-Loops von Haloalkan Dehalogenasen in molekulardynamischen Simulationen (104). Aufgrund dieser Flexibilität könnte sich der NC-Loop zur Seite bewegen und so den Zugang in das Molekül ermöglichen. Zusätzlich wird vermutet, dass die Cap-Helix α_3 , welche sich oberhalb des NC-Loops und neben den Cap-Loop befindet, ebenfalls beweglich ist, wie es für die strukturell ähnlichen Haloalkan Dehalogenasen beschrieben wurde (110).

NC- und Cap-Loop unterscheiden sich beide deutlich in ihrer Länge für die drei bekannten EH-Strukturen. Unterschiede in den Längen von NC- und Cap-Loop könnten direkte Auswirkungen auf die Form und Erreichbarkeit des katalytischen Zentrums haben und die unterschiedlichen Substratspektren von EHs verursachen. Während mikrosomale EHs mit langen NC-Loops fähig sind voluminöse polyzyklische aromatische Kohlenwasserstoffe umzusetzen (48), setzen cytosolische Säugetier und Pflanzen EHs epoxidierte Fettsäuren um. Diese EHs enthalten einen langen Cap-Loop.

5.1.2 Klassifizierung und Substratspektrum

Eine Klassifizierung für EHs, außerhalb der Einteilung in mikrosomale und cytosolische EHs, wurde bisher nicht beschrieben.

Die Einteilung in Superfamilien auf Basis der Sequenzähnlichkeit und die Gruppierung in Cluster aufgrund von Loop-Längen korreliert mit Substratspezifität und metabolischer Funktion der EHs (Tabelle 9): 1) Alle **cytosolischen EHs höherer Organismen aus Cluster I** haben lange Cap-Loops und NC-Loops mittlerer Länge. Diese Enzyme zeigen Aktivität gegen eine gemeinsame Substratklasse, aliphatische Epoxide und sind am Fettsäure-Metabolismus beteiligt (48). Die homologe Familie der Säugetier EHs ist an den Metabolismen von Xenobiotica und von im Organismus erzeugten epoxidierten Fettsäuren beteiligt. Die Familie der Pflanzen EHs hingegen spielt eine zentrale Rolle in der Biosynthese essentieller Bestandteile der Cuticula, wie epoxidierte Stearinsäure (23) und bei der Entgiftung von epoxidierten Fettsäuren im Keimling (4). 2) Alle **bakteriellen EHs aus Cluster II** haben kurze Cap- und NC-Loops. Für mehrere dieser EHs wurden kleine Substrate wie Styroloxid, Mono- und Disubstituierte Epoxide nachgewiesen (130). 3) **Mikrosomale EHs** haben kurze bis mittel lange Cap-Loops und mittel bis lange NC-Loops. Diese Enzyme sind am Metabolismus polyzyklischer aromatischer Kohlenwasserstoffe beteiligt. Die Familie der Insekten EHs ist durch den Abbau von polyzyklischen Sesquiterpenestern in die Regulation der Juvenil Hormone involviert (147), während die gut untersuchten Säugetier EHs an der Aktivierung von karzinogenen polyzyklischen Kohlenwasserstoffen und der Entgiftung epoxidierte Stoffwechselintermediate mitwirken (48). Die meisten mikrosomalen EHs besitzen kurze Cap-Loops, mikrosomale Säugetier EHs jedoch enthalten Cap-Loops mittlerer Länge. Interessanterweise, ist diese Familie, wie cytosolische EHs höherer Organismen, fähig epoxidierte Fettsäuren umzusetzen (132). Auf Grundlage dieser Daten, scheint es, dass lange Cap-Loops die Fähigkeit für den Umsatz

aliphatischer Substrate bedingen und lange NC-Loops die Fähigkeit für den Umsatz polyzyklischer aromatischer Kohlenwasserstoffe verleihen, während EHs mit kurzen Cap- und NC-Loops kleine Substrate bevorzugen. Gemäß dieser Regel, kann die Substratspezifität nicht-charakterisierter bakterieller EHs von Cluster I und III vorhergesagt werden: Cluster I EHs enthalten lange Cap-Loops und sollten daher epoxidierte Fettsäuren umsetzen, und Cluster III EHs enthalten lange NC-Loops und sollten epoxidierte polyzyklische aromatische Kohlenwasserstoffe umsetzen. Dieser beobachtete Zusammenhang zwischen Loop-Länge und Substratspezifität wird weiterhin gestützt durch experimentelle Daten, erhalten durch Mutanten von Haloalkan Dehalogenasen (114). Bei diesem strukturell sehr ähnlichen Enzym führten Variationen der Länge des NC-Loops zur Umsetzung von Chloralkanen unterschiedlicher Kettenlängen.

Tabelle 9: Klassifikation von EHs durch Loop-Länge und Substratspezifität

EH Superfamilie	Organismen	Cluster	NC-Loop	Cap-Loop	Substrate
cytosolisch	Pflanzen/ Säugetiere	I	kurz-medium	lang	epoxidierte Fettsäuren
	Bakterien	I	kurz-medium	lang	epoxidierte Fettsäuren ^P
	Bakterien	II	kurz	kurz	kleine Substrate ^c
mikrosomal	Insekten	III	lang-sehr lang	kurz	polyzyklische aromatische Kohlenwasserstoffe ^c
	Säugetiere	III	lang-sehr lang	medium	polyzyklische aromatische Kohlenwasserstoffe ^c
	Pilze/Hefen	III	lang-sehr lang	kurz	polyzyklische aromatische Kohlenwasserstoffe ^c
	Bakterien	III	lang	kurz	polyzyklische aromatische Kohlenwasserstoffe ^P

^c: biochemisch charakterisiert

^P: putativ (Vorhersage durch Loop-Länge)

5.1.3 Funktionalität

Obwohl EH Aktivität gegen epoxidierte Fettsäuren und polyzyklische aromatische Kohlenwasserstoffe essentiell ist für einen Organismus, enthalten nur Säugetiergenome cytosolische und mikrosomale EHs. Im Vergleich dazu kodieren die Genome von Pilze/Hefen, Insekten und Pflanzen nur jeweils für eine der beiden EH Superfamilien. Obwohl Pflanzen ein breites Spektrum an epoxidierten Steroiden und Diterpenoiden (1) enthalten, fehlen mikrosomale EHs. In Pflanzen wie z.B. *Arabidopsis thaliana* und *Zea mays* wurde jedoch das Enzym Glutathion S-Transferase (GST) gefunden (82), welche ebenfalls Epoxidringe öffnen können. Dies könnte ein pflanzen-spezifischer Abbauweg für epoxidierte polyzyklische aromatische Kohlenwasserstoffe sein. In ähnlicher Weise kodieren Pilze/Hefen ausschließlich für mikrosomale EHs. Die Aktivität der fehlenden cytosolischen EHs wird möglicherweise durch andere Enzyme übernommen: *Saccharomyces cerevisiae* z.B. hat die Fähigkeit die epoxidierte Fettsäure Leukotrien mit Hilfe einer Leukotrien A(4) Hydrolase umzusetzen (81). Leukotrien A(4) Hydrolasen haben keine Sequenzähnlichkeit zu EHs und gehören zur Familie der Zink-Metalloproteasen, sie zeigen EH und Protease Aktivität (56). Interessanterweise enthalten die Genome einiger Bakterien Klassen (*Mycobacterium*) oder Spezies (*Caulobacter crescentus*, *Mesorhizobium loti*, *Sinorhizobium meliloti*) EHs beider Superfamilien. Die weite Verbreitung bakterieller EH Sequenzen reflektiert das breite Substratspektrum, welchem Bakterien ausgesetzt sind (90). Im Allgemeinen zeigen Bakterien eine hohe Diversität, ausgelöst durch die biologischen Habitate, die durch diese Organismen bewohnt werden (21).

5.1.4 Homologiemodellierung

Mit paarweisem Sequenzvergleich, ist Homologiemodellierung begrenzt auf hohe Sequenzähnlichkeiten zwischen Target- und Template-Protein (11). Daher könnte nur ein kleiner Prozentsatz von 14% der diversen Proteinfamilie der EH-Strukturen vorhergesagt werden. Durch Nutzung der Ergebnisse der systematischen Analyse kann dieser Wert auf 78% der 93 EHs gesteigert werden. Unter der Annahme, dass EHs mit ähnlichen Loop-Längen eine ähnliche Struktur aufweisen, wurde die Loop-bezogene Einteilung in die drei Cluster und das Multisequenz-Alignment der 93 EHs zur Erstellung von Homologiemodellen von fünf repräsentativen EHs aus Cluster I-III genutzt: *Glycine max* EH, *Mycobacterium tuberculosis* EH, *Corynebacterium* sp. EH, *Mesorhizobium loti* EH and *Streptomyces coelicolor* EH. Die Anzahl der vorhersagbaren Strukturen variiert für die drei Cluster:

Während 94% aller Cluster I EHs und 100% aller Cluster II EHs vorhersagbar sind, können nur 19% der mikrosomalen Cluster III EHs komplett modelliert werden. Infolge ihrer medizinischen Relevanz sind mikrosomale Säugetier EHs gut untersuchte Enzyme (48), welche sehr lange NC-Loops enthalten. Eine Kristallstruktur eines solchen Enzyms würde dringend benötigt, um auch Strukturen mit sehr langen NC-Loops modellieren zu können. Zusammenfassend können zur Zeit die Strukturen aller α/β Hydrolase Domänen, 78% der NC- und 89% der Cap-Loops der 93 EH Sequenzen durch Homologiemodellierung vorhergesagt werden.

Obwohl EHs nur eine geringe Sequenzähnlichkeit aufweisen, ist ihre Struktur hoch konserviert. Daher ist Strukturvorhersage möglich, trotz der Tatsache, dass die Sequenzähnlichkeiten oftmals unterhalb des Grenzwertes für verlässliche Homologiemodellierungen von 30% liegen. Zusätzlich scheint die Substratspezifität durch die Länge von zwei variablen Loops dominiert zu werden. Die hier vorgestellte systematische Analyse zeigt die modulare Architektur von EHs auf, welche den Weg zu einem tieferen Verständnis der Struktur und Funktion von EHs und auch anderer homologer α/β Hydrolasen ebnet.

5.2 Familienspezifische degenerierte Primer

5.2.1 Screening nach unbekanntem EHs

Die Anzahl neuer Biokatalysatoren für den Einsatz in der chemischen Industrie ist in den letzten Jahren stark angestiegen. Heute ist es möglich, Enzyme mit bestimmten Substratspektren und -Selektivitäten so zu verändern, dass sie ausreichend stabil für die organische Synthese werden (35). Um den Bedarf an neuen Enzymen decken zu können, wurde eine große Anzahl an Screening Methoden entwickelt. Für ein solches Screening werden Enzymquellen benötigt: kommerziell erhältliche Enzymbibliotheken, Stammsammlungen, Genbanken aus kultivierbaren oder nicht kultivierbaren Organismen. Jede dieser Quellen hat ihre Vor- und Nachteile. Die kommerziell erhältlichen Enzymbibliotheken (z.B. Sigma, Roche Molecular Biochemicals) haben den Vorteil, Enzyme in großer Menge und gleich bleibender Qualität liefern zu können. Der Nachteil ist ein begrenztes Angebot an verschiedenen Enzymen und es besteht normalerweise keine Möglichkeit für Mutationsexperimente. Stammsammlungen (z.B. ATCC, DSMZ) sind, da sie komplette Organismen anbieten, eine Quelle für neue unentdeckte Katalysatoren, ein Nachteil

ist die Limitierung auf den jeweiligen Organismus. Genbanken haben die Vorteile Quellen für neue unentdeckte Katalysatoren sein zu können, Expression von Enzymen die im Original-Wirt nicht exprimiert werden, zu bieten, Fermentationen sind möglich und die Einführung von Mutationen ist einfach. Nachteile sind die notwendige „Wartung“ dieser Bibliotheken, und die Gefahr, dass viele Proteine im verwendeten Wirtsorganismus nicht exprimiert werden. Um Enzymquellen auf Aktivität screenen zu können werden Screeningsysteme benötigt. Im Fall der Epoxidhydrolasen ist entweder die Detektion des entstehenden Diols oder des Epoxids sehr aufwendig (38, 150) oder ungenau (149) oder das Substrat unbekannt. Daher wurde das Screening nach unbekanntem EHs auf DNA-Ebene durchgeführt. Mit Hilfe der in der systematischen Analyse durchgeführten phylogentischen Analyse ermittelten homologen Familien, wurden mit dem Programm CODEHOP (120) degenerierte Primer für das Screening entwickelt und erfolgreich eingesetzt.

5.2.1.1 CODEHOP-Primer

Die meisten Anwendungen der PCR (93) nutzen Primer, die exakt mit einer bekannten Target-Sequenz übereinstimmen. In einigen Situationen jedoch, werden Primer für die Amplifikation unbekannter Gene benötigt, z.B. bei der Amplifikation von Genen, die für Proteine aus bekannten Proteinfamilien kodieren (36, 119). In solchen Fällen basiert das Primer Design in der Regel auf reverser Translation konservierter Bereiche überlagerter Sequenzen. Aufgrund der Degeneriertheit des genetischen Codes kann diese Art des Primer Designs Schwierigkeiten bereiten. Die degenerierten Positionen eines Primers erschweren die Bestimmung geeigneter Annealing Temperaturen und Primerlängen. Der Degenerierungsgrad limitiert die Amplifikation fremder Gene, da leicht Fehlamplifikationen aufgrund zu geringer Übereinstimmung zwischen Primer und Target-Sequenz auftreten können.

Zwei Strategien für das Amplifizieren unbekannter Gene werden verbreitet angewendet: Synthese eines Pools an degenerierten Primern, mit den oben beschriebenen Problemen und die Erstellung von Consensus-Primern, welche zwar nicht degeneriert sind, aber ebenfalls Unterschiede zur zu amplifizierenden Sequenz aufweisen, und somit ebenfalls zu Fehlalignments führen können.

CODEHOP (**C**onsensus **D**egenerate **H**ybrid **O**ligonucleotide **P**rimers) (120) sind Hybrid-Primer, die sich aus einem 3' degenerierten Core und einer 5' nicht-degenerierten Consensus Clamp zusammensetzen. Reduktion der Länge des degenerierten Cores auf ein Minimum, verringert die Anzahl degenerierter Primer im Primer-Pool. Die Hybridisierung des

degenerierten Cores mit der Target-Sequenz wird durch die nicht-degenerierte Consensus Clamp stabilisiert. Dies erlaubt höhere Annealing Temperaturen ohne eine größere Anzahl degenerierter Primer.

5.2.1.2 Amplifikation der CODEHOP-Fragmente

In dieser Arbeit wurden CODEHOP-Primer verwendet, um die EH zu identifizieren, die für die bekannte EH Aktivität von *Streptomyces antibioticus* Tü4 gegen Styroloxid verantwortlich ist (150). Es wurden CODEHOP-Primer für die Gruppe mit eukaryotischen EHs verwandten bakteriellen EHs, entwickelt. Mit diesen Primern war es möglich 771 bp mit einer signifikanten Homologie zu EHs zu amplifizieren. Die Translation ergab eine 284 AS lange Sequenz, in der alle katalytisch relevanten AS vorhanden waren (Abbildung 22). Mit demselben Primersatz konnten zusätzlich aus genomischer DNA von *Rhodococcus ruber* 703 bp mit signifikanter Homologie zu EHs amplifiziert werden. Die Translation ergab eine 215 AS lange Sequenz, die ebenfalls katalytisch relevante AS enthielt (Abbildung 28).

5.2.1.3 Vervollständigen der Sequenz *Streptomyces antibioticus* Tü4 EH

Die Sequenz der *Streptomyces antibioticus* EH wurde N- und C-Terminal vervollständigt. Durch Reinigung und N-terminale Ansequenzierung (durchgeführt im Labor von Herrn Dr. Friedrich, BASF AG) konnte der N-Terminus der EH bestimmt werden. Dieser wurde über PCR and das CODEHOP-Fragment angefügt. Da die *Streptomyces antibioticus* EH große Ähnlichkeit mit einer EH aus *Streptomyces avermitilis* zeigte (Abbildung 25) wurde der C-Terminus dieses Enzyms mittels PCR and das Konstrukt aus CODEHOP-Fragment und N-Terminus angefügt.

Auf eine Expression der *Streptomyces antibioticus* EH wurde aus zeitlichen Gründen verzichtet.

5.3 Datenbanken

Im April 2003 wurde die Sequenzierung des Humanen Genoms vollendet (28). Die freie und umfassende Verfügbarkeit einer großen Vielfalt an Daten ist eine fruchtbare Grundlage für Biologen aller Disziplinen (14). Die größten öffentlichen Datenbanken enthalten eine gewaltige Menge an Daten. Zur Zeit beinhaltet die GenBank (16) > 17 Milliarden Nukleotidbasen, welche > 14 Millionen Sequenzen in 100 000 Spezies vertreten. Die

Möglichkeiten, die ein solch großer Satz an Daten bietet sind überwältigend. Der Zuwachs an neuen Daten wird weiterhin zunehmen, allein die GenBank verdoppelt ihre Zuwachsraten innerhalb von 14 Monaten. Ebenso wichtig sind spezialisierte Datenbanken, die einen einfacheren Zugang zu einer Teilmenge der Daten gewähren (14). Spezialisierte Datenbanken sind von enormem Wert für die Forschung, da sie die große Datenmenge reduzieren und die Einbeziehung anderer Arten von Daten erlauben. Diese Datenbanken bieten häufig nicht nur sequenzbasierte Informationen, sondern auch z.B. Daten über Genexpressionen, makromolekulare Wechselwirkungen oder Informationen über biologische Stoffwechselwege. Einer der wichtigsten Punkte ist aber, dass diese kleineren Datenbanken von Spezialisten unterhalten werden und die Daten häufig experimentell verifiziert sind, was bedeutet, dass sie den neuesten Stand der Forschung auf dem jeweiligen Gebiet darstellen.

5.3.1 Die EH/HD Datenbank

Die EH/HD Datenbank (<http://led.uni-stuttgart.de>) wurde, wie die Lipase Engineering Database (LED) (45) vor ihr, entwickelt, um Daten für die systematische Analyse von Sequenz-, Struktur- und Funktionsbeziehungen der enthaltenen Proteinfamilien bereitzustellen. Die EH/HD Datenbanken basieren auf einem erweiterbaren Modell einer relationalen Datenbank, welche Sequenz-, Protein- und Strukturinformationen aus den öffentlichen Datenbanken GenBank (16) und PDB (17) extrahiert. Die EH/HD Datenbank enthält auf Proteinebene 397 Sequenz-, 305 Proteineinträge und 51 Strukturdatensätze für 19 dieser Proteine, welche hierarchisch in Superfamilie, homologe Familie, Protein- und Sequenzeintrag geordnet sind. Die EH/HD Datenbank enthält neben Asp-Hydrolasen wie EHs, HDs und Haloacid Dehalogenasen auch Ser-Hydrolasen wie Non-heme Peroxidasen, 2-hydroxymuconic semialdehyd Hydrolasen, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid Hydrolasen, Prolyl-Aminopeptidasen, Esterasen und Luciferasen, welche in 2 Superfamilien und 14 homologe Familien eingeteilt wurden. Für die systematische Analyse werden annotierte Sequenz-Alignments und phylogenetische Analysen für jede homologe Familie und Superfamilie bereitgestellt. Für strukturelle Vergleiche wurden alle experimentell bestimmten Strukturen überlagert. Zusätzlich wurden die Homologiemodelle der fünf EHs aus *Glycine max*, *Mycobacterium tuberculosis*, *Corynebacterium* sp., *Mesorhizobium loti* und *Streptomyces coelicolor* in die Datenbank aufgenommen.

Die EH/HD Datenbank wurde durch Dang Sy Hai um eine integrierte DNA-Datenbank erweitert (32). Diese Datenbank enthält zu jedem Proteineintrag der EH/HD Datenbank die

zugehörige kodierende DNS-Sequenz. Außerdem werden annotierte DNA-Alignments jeder homologen Familie und Informationen über Codon-Usage und eventuell enthaltene „Rare Codons“ bereitgestellt.

5.3.2 Reduktiver Fold und Anker-Prinzip

Dass der α/β Hydrolase Fold Variationen zulässt ist bekannt (59). In den frühen 1990er Jahre wurde bei fünf Proteinen das gleiche Faltungsmuster festgestellt: der α/β Hydrolase Fold (101). Die Sequenzen waren scheinbar nicht verwandt, die Proteine setzten unterschiedliche Substrate um und es wurde nicht einmal das gleiche Nukleophil benutzt. Die bisher aufgelösten Strukturen zeigen, dass der α/β Hydrolase Fold wohl eines der variabelsten Proteinfaltungsmuster ist. Dieser Fold toleriert sogar große Insertionen, so dass die Größe einer Domäne zwischen 197 AS für die *Fusarium solani* Cutinase bis hin zu 582 AS für die Acetylcholinesterase aus *Mus musculus* betragen kann. Das Verständnis dieser Plastizität ist für Protein Design sowie Proteinklassifizierungen auf Sequenzebene unerlässlich (59).

Der kanonische α/β Hydrolase Fold wird beschrieben als eine aus acht Strands bestehende, meist parallele α/β Struktur (101). Diese Beschreibung wurde von mehreren Datenbanken übernommen wie z.B. SCOP (94), CATH (102) und die α/β Hydrolase Fold Datenbank ESTHER (29). Insertionen, Cap, Lid oder Flap genannt, definieren Form und Erreichbarkeit der Bindungstasche und treten an unterschiedlichen Positionen im Fold auf (62). Insertionen können nach einem oder mehreren Strands auftreten, meist sind die Strands β_3 , β_4 , β_6 und β_7 betroffen (59). Substratbinde-Domänen sind meist aus mehreren Insertionen aufgebaut und bilden ein Cap oberhalb des aktiven Zentrums, sie können aber auch durch große individuelle Domänen, wie die C-terminale Domäne pankreatischer Lipasen, gebildet werden.

Analysen von 19 Strukturen verschiedener homologer Familien der LED und EH/HD Datenbank betätigen diese Aussagen (Tabelle 10).

Tabelle 10: Variationen in α/β Hydrolase Fold und Anker-Prinzip der Strukturen von LED und EH/HD Datenbank.

Homologe Familie/Enzymfamilie	α/β Hydrolase Fold	Anzahl β -Strands	Anzahl α -Helices	Anker-Typ
GX-Typ				
Haloalkan Dehalogenasen	komplett	8	6	hydrophil. A.
Epoxidhydrolasen	komplett	8	6	hydrophob. T.
2-Hydroxy-6-oxo-7-methylocta-2,4-dienoate Hydrolasen	komplett	8	6	hydrophil. A.
Prolyl-Aminopeptidasen	komplett	8	6	-
Non-heme Haloperoxidasen	komplett	8	6	hydrophob. T.
<i>Burkholderia</i> Lipasen	nicht komplett	6	5	hydrophob. A.
<i>Bacillus subtilis</i> Lipasen	nicht komplett	6	6	hydrophob. A.
<i>Candida antarctica</i> Lipasen	nicht komplett	7	5	hydrophil. A.
<i>Fusarium</i> Cutinasen	nicht komplett	5	5	hydrophil. A.
Gastrische Lipasen	komplett	8	6	hydrophob. A.
Pankreatische Lipasen	nicht komplett	8	4	hydrophob. A.
<i>Rhizomucor miehei</i> Lipasen	nicht komplett	5	3	hydrophil. A.
<i>Moraxella</i> Lipasen	nicht komplett	8	5	hydrophob. A.
GGGX-Typ				
Acetylcholinesterasen	komplett	8	6	
<i>Bacillus</i> Esterasen	komplett	8	6	
„Mammalian bile salt activated“ Lipasen	komplett	8	6	
Säugetier Carboxylesterasen	komplett	8	6	
<i>Candida rugosa</i> Lipasen	komplett	8	6	
Hormon sensitive Lipasen	komplett	8	6	

Weiterhin wurde die Stabilisierung des Oxyanionholes (Anker-Konzept (111)) der Hydrolasen des GX-Typs (111) beider Datenbanken untersucht (Tabelle 10). Das HGX-Motiv der GX-Hydrolasen ist eine Loop-Region, die einen Teil des Oxyanionholes trägt. Die Position der Backbone Amide der Oxyanionhole AS muss absolut fixiert sein, um die Funktion des Oxyanionholes während der Reaktion zu gewährleisten. Die Position auf einer flexiblen Loop-Region erzwingt eine zusätzliche Stabilisierung dieser AS. Dies wird bewerkstelligt durch Wechselwirkungen der Seitenkette dieser AS mit der Seitenkette einer stabilisierenden (verankernden) AS, dem Anker. Gestützt wird diese Idee durch Mutationsexperimente durchgeführt mit der *Pseudomonas cepacia* Lipase. In diesem Fall führte ein Austausch der Seitenkette der Oxyanionhole AS von hydrophob nach hydrophil ohne Änderung des Ankers zu einem Verlust der Aktivität (111), nach einer zusätzlichen

Mutation, die dem Anker einen ebenfalls hydrophilen Charakter verlieh, konnte wieder Aktivität nachgewiesen werden (44). Mögliche Anker-Typen sind: hydrophiler Anker (hydrophil. A.), hydrophober Anker (hydrophob. A.) und hydrophobe Tasche (hydrophob. T.). Untersuchungen an 13 Strukturen unterschiedlicher homologer Familien aus LED und EH/HD Datenbank (Tabelle 10), zeigen, dass nicht nur die strukturell diversen Proteine der LED unterschiedliche Anker-Typen aufweisen, sondern auch die strukturell sehr ähnlichen Proteine der EH/HD Datenbank. Aus der systematischen Analyse konnten keine Hinweise auf eine strukturell bedingte Präferenz für einen bestimmten Anker-Typ ermittelt werden.

5.3.3 Das GXGXS-Motiv

Für die Rolle des hoch konservierten GXGXS-Motivs wurden mehrere Vorschläge gemacht. Als Funktion wurde eine Stabilisierung des Oxyanionholes durch die erste, meist aromatische, variable AS des GXGXS-Motivs vorgeschlagen (116). Die HD aus *Xanthobacter autrophicus* z.B. enthält an dieser Position ein Phe. Der aromatische Ring des Phe (GFGXS) ist in einer „T-Anordnung“ (Abbildung 34) auf den Imidazolring des His (HGX) ausgerichtet, bei welchem der positiv geladene His-Ring aufrecht auf der leicht negativ geladenen Oberfläche des Phe steht (116). Zwar wurden keine Wasserstoffbindungen oder Salzbrücken entdeckt, eine His-Phe Wechselwirkung kann jedoch einen erheblichen stabilisierenden Effekt haben (9, 85). Des weiteren spielt das GXGXS-Motiv bei der Bildung der hydrophoben Tasche des Ankers eine Rolle (Abbildung 34). Diese Wechselwirkungen lassen sich jedoch nicht bei allen untersuchten Strukturen der EH/HD-Datenbank nachweisen und erklären daher nicht die hohe Konservierung des GXGXS-Motivs. Alle Hydrolasen des GX-Typs enthalten ein konserviertes HGX-Motiv (111) im Oxyanionhole, aber nur die Mitglieder der EH/HD Datenbank weisen das GXGXS-Motiv auf. Daher wird eine andere Funktion des GXGXS-Motivs vorgeschlagen: in Hydrolasen, welche die β -Strands $\beta 1$ und $\beta 2$ und eine den EHs ähnliche Cap-Domäne enthalten, stabilisiert das GXGXS-Motiv die Position von $\beta 1$ und $\beta 2$ und stellt die Verbindung zwischen der α/β Hydrolase Domäne und der Cap-Domäne dar.

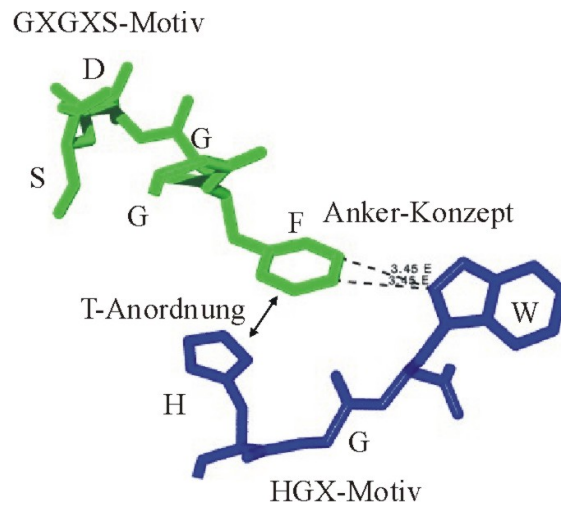


Abbildung 34

Darstellung der Wechselwirkungen zwischen HGX-Motiv (blau) und GXGXS-Motiv (grün) der *Agrobacterium radiobacter* EH. Die T-Anordnung des His-Imidazolrings (H) des HGX-Motivs und des aromatischen Ringes des Phe (F) des GXGXS-Motivs wurde durch einen Doppelpfeil markiert, der Abstand zwischen Phe (F) und Trp (W) des HGX-Motivs wurde durch unterbrochene Linien veranschaulicht.

6 Zusammenfassung

Epoxidhydrolasen (EHs) besitzen häufig die Fähigkeit industriell wichtige Epoxide enantioselektiv umzusetzen. Viele enantiomerenreine Epoxide und vicinale Diole sind Ausgangsstoffe für die pharmazeutische und agrochemische Industrie (51).

Ziel dieser Arbeit war es durch eine systematische Analyse verfügbarer EH Sequenzen und Strukturen Übereinstimmungen und Unterschiede dieser sehr diversen Enzymfamilie zu ermitteln.

Über Stichwort- und BLAST-Suche wurden 93 unterschiedliche EH Sequenzen ermittelt. Eine phylogenetische Untersuchung dieser Sequenzen zeigte die deutliche Trennung der beiden Superfamilien in cytosolische (cEHs) und mikrosomale (mEHs) EHs und die Aufteilung in separate homologe Familien innerhalb jeder Superfamilie. Bakterielle EHs zeigen im Gegensatz zu eukaryotischen EHs eine große Variabilität und bilden kaum deutlich voneinander abgegrenzte homologe Familien. EHs gehören zur Familie der α/β Hydrolasen. Untersuchungen der drei bekannten EH Strukturen aus *Agrobacterium radiobacter*, *Mus musculus* und *Aspergillus niger* führten zur Einteilung der modularen Strukturen in N-terminale Domäne, N-terminale katalytische Domäne, NC-Loop, Cap-Domäne, Cap-Loop und C-terminale katalytische Domäne. Trotz großer Unterschiede in der äußeren Form der drei Strukturen, verursacht durch die variable N-terminale Domäne, weisen EHs eine hoch konservierte Struktur im Bereich des α/β Hydrolase Folds auf. Unterbrochen wird diese hohe Konservierung nur durch zwei variable Loops: NC- (verbindet N-terminale katalytische Domäne und Cap-Domäne) und Cap-Loop (befindet sich innerhalb der Cap-Domäne). Die bakterielle EH aus *Agrobacterium radiobacter* enthält kurze Loops (NC-Loop: 19 AS, Cap-Loop: 9 AS), die EH aus *Mus musculus* enthält einen NC-Loop mittlerer Länge (23 AS) und einen langen Cap-Loop (35 AS) und die mikrosomale EH aus *Aspergillus niger* enthält einen langen NC-Loop (35 AS) und einen kurzen Cap-Loop (8 AS). Die Loop-Längen aller EHs wurden aus dem Multisequenz-Alignment der 93 EHs bestimmt und in einem Diagramm gegen einander aufgetragen. Hierbei bildeten sich drei nach Superfamilien und Organismengruppen geordnete Cluster: Cluster I (cEHs aus Säugetieren, Pflanzen und zu eukaryotischen EHs verwandte bakterielle cEHs) enthält die strukturbekannte EH aus *Mus musculus*, Cluster II (cEHs aus Bakterien) enthält die strukturbekannte EH aus *Agrobacterium radiobacter* und Cluster III (mEHs) enthält die strukturbekannte EH aus *Aspergillus niger*.

EHs besitzen zwar eine hoch konservierte Struktur, zeigen aber nur geringe Sequenz Identitäten, welches die Erstellung verlässlicher Homologiemodelle erschwert oder

unmöglich macht. Mit der Annahme, dass EHs mit ähnlichen Loop-Längen eine ähnliche Struktur aufweisen, wurde die Loop-bezogene Einteilung in die drei Cluster und das Multisequenz-Alignment der 93 EHs zur Erstellung von Homologiemodellen von fünf repräsentativen EHs aus Cluster I-III genutzt. Mit dem üblichen Verfahren, dem paarweisen Alignment zwischen Target- und Template-Sequenz, könnten, aufgrund der geringen Sequenzähnlichkeit maximal 14% aller EH Strukturen vorhergesagt werden, Loop-bezogen ließe sich dieser Wert auf 78% aller EHs steigern.

Des weiteren lässt sich die Loop-bezogene Einteilung in Cluster zur Vorhersage der Substratspezifität nutzen. Betrachtet man die Substratspektren der einzelnen EHs, fällt auf, dass EHs mit langem Cap-Loop (cyt. und mik. Säugetier und Pflanzen EHs) die Fähigkeit haben epoxidierte Fettsäuren umzusetzen (22, 48, 52), während EHs mit langen NC-Loops (mik. Säugetier, Insekten EHs und EHs aus Pilzen/Hefen) epoxidierte polyzyklische aromatische Kohlenwasserstoffe als Substrate akzeptieren (19, 34, 48). Für EHs mit kurzen Cap-Loops wird der Umsatz von Fettsäuren nicht beschrieben. Der lange Cap-Loop bildet einen hydrophoben Tunnel von der Proteinoberfläche bis zum Nukleophil der katalytischen Triade, der bei EHs mit kurzem Cap-Loop nicht erkennbar ist. Dieser hydrophobe Tunnel führt zur Bildung einer L-förmigen Bindungstasche (8) und ähnelt der Fettsäure-Bindestelle in Lipasen (112) und würde der hydrophoben Alkankette einer Fettsäure Platz bieten. EHs mit kurzem Cap-Loop und ohne hydrophoben Tunnel und L-förmiger Bindungstasche, dürften also nicht die Fähigkeit haben epoxidierte Fettsäuren umzusetzen, wohingegen die bakteriellen EHs aus Cluster I, die eine Verwandtschaft zu eukaryotischen EHs zeigen und lange Cap-Loops besitzen, Aktivität zeigen sollten.

Die aus der phylogentischen Analyse erhalten Einteilung in homologe Familien wurde zur Entwicklung familienzuspezifischer Primer verwendet, welche an zwei bakteriellen Organismen getestet wurden. *Streptomyces antibioticus* Tü4 zeigt eine EH-Aktivität gegen Styroloxid (150). Ein bereits aus dieser Spezies amplifiziertes 236 bp Fragment siedelte sich im phylogenetischen Baum bei den zu eukaryotischen EHs verwandten bakteriellen EHs des Clusters I an. Mit den für diese Gruppe entwickelten Primern konnte ein 771 bp Fragment mit signifikanter Homologie zu EHs amplifiziert werden. Nach Reinigung, N-terminaler Ansequenzierung und Trypsin-Verdau durch die BASF AG, Ludwigshafen, konnte das Fragment N-terminal komplementiert werden. Der C-Terminus wurde mittels PCR nach Vorlage der EH aus *Streptomyces avermitilis* aufgebaut. Als zweites Bsp. wurden aus genomischer DNA der Spezies *Rhodococcus ruber* (LU760), die wie die Gattung *Streptomyces* zur Ordnung der *Actinomycetales* gehört, unter Verwendung des selben

Primersatzes ein 700 bp-Fragment amplifiziert, welches ebenfalls als Fragment einer putativen EH identifiziert werden konnte.

Der zweite Teil dieser Arbeit befasst sich mit der Erstellung und Auswertung einer familienspezifischen Datenbank. Die Epoxidhydrolase / Haloalkan Dehalogenase Datenbank (EH/HD) stellt eine Erweiterung des 93 EH Alignments unter Einbeziehung weiterer Enzymfamilien und Strukturinformationen dar. Für die Erstellung der EH/HD Datenbank wurde das automatische Datenextraktionssystem der Lipase Engineering Database (LED) benutzt (45). Ein Satz repräsentativer Sequenzen, resultierend aus den 93 EHs der systematischen Analyse und 18 durch Stichwortsuche ermittelten HDs, wurde zur Erstellung der Datenbank genutzt. Die gesammelten Daten wurden durch das Datenbanksystem DWARF (46) hierarchisch in Proteineintrag, homologe Familie und Superfamilie geordnet. Die Datenbank enthält 397 Sequenzeinträge, 305 Proteineinträge und für 19 dieser Proteineinträge 51 Strukturdatensätze. Neben EHs und HDs beinhaltet die Datenbank weitere Enzymfamilien wie z.B. Non-heme Peroxidasen, 2-hydroxymuconic semialdehyd Hydrolasen, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid Hydrolasen, Prolyl-Aminopeptidasen, Esterasen und Luciferasen, welche zu 14 homologen Familien und zwei Superfamilien zusammengefasst wurden.

Um die Gemeinsamkeiten und Unterschiede der großen Anzahl an verschiedenen α/β Hydrolasen aus LED und EH/HD Datenbank zu untersuchen, wurde eine Analyse der verfügbaren Strukturdaten durchgeführt. Während die Strukturen der EH/HD Datenbank einen kanonischen, also vollständigen α/β Hydrolase Fold aufwiesen, wiesen die Strukturen der Hydrolasen der LED häufig Reduzierungen auf. Diese beschränkten sich jedoch auf die β -Strands $\beta 1$ und $\beta 2$, sowie die α -Helices $\alpha 4$ und $\alpha 5$, welche keine katalytisch wichtigen AS tragen. Trotz massiver struktureller Unterschiede ist die katalytische Maschine strukturell in allen α/β Hydrolasen erhalten.

Für Lipasen wurde ein sog. Anker-Konzept (111) entwickelt, welches auch bei den Strukturen der EH/HD Datenbank greift. Trotz der hohen Konservierung der Strukturen der EH/HD Datenbank wurde die Stabilisierung des Oxyanionholes unterschiedlich realisiert, sie erfolgt entweder über eine hydrophobe Tasche oder einen hydrophilen Anker. Im Fall der Prolyl-Aminopeptidasen konnte keiner der bekannten Anker-Typen nachgewiesen werden.

Das GXGXS-Motiv (92) ist hoch konserviert und bildet einen strukturell ebenfalls hoch konservierten Loop aus. Das GXGXS-Motiv tritt nur in Strukturen der EH/HD Datenbank auf und befindet sich zwischen den ersten beiden β -Strands des α/β Hydrolase Folds und den

Cap-Helices $\alpha 4$ und $\alpha 5$, die drei Strukturteile sind durch Wasserstoffbrücken miteinander verbunden. Die Positionen, der die Wasserstoffbrücken ausbildenden AS, sind auf $\beta 2$ und dem GXGXS-Motiv absolut konserviert. Die Wasserstoffbrückenbindungen zur Cap-Domäne sind flexibler, befinden sich aber alle in der Loop-Region zwischen Cap-Helix $\alpha 4$ und $\alpha 5$. Das GXGXS-Motiv stabilisiert die Position der Strands $\beta 1$ und $\beta 2$ und wirkt wie ein Schloss, das Cap-Domäne und N-terminale katalytische Domäne verbindet.

Durch systematische Analyse von Sequenzen und Strukturen dieser großen und diversen Proteinfamilie, wurden strukturell und funktionell relevante Motive und die modulare Architektur analysiert. Diese Untersuchungen lieferten einen tiefen Einblick in die Sequenz-Struktur-Funktions Beziehungen der α/β Hydrolasen. Das Verständnis der modularen Struktur erlaubt zuverlässige Vorhersagen der biochemischen Eigenschaften für „functional genomics“ und „protein engineering“.

7 Material und Methoden

7.1 Theoretische Arbeiten

7.1.1 Software und Hardware

7.1.1.1 Software

Tabelle 11: Verwendete Software

Einsatzgebiet	Software
Homologiesuche	BLAST (3)
Sequenzvergleich	ClustalX (136)
Visualisierung/Bearbeitung von Strukturdaten	Spdb-Viewer (54) InsightII, Accelrys
Sekundärstrukturvorhersage	DSSP (66)
Homologiemodellierung	Swiss-Model Server (54)
Molekulardynamische Simulation	Amber6 (106) Titra (89)
Phylogenetische Analyse	Tree-Puzzle (126)
Visualisierung phylogenetischer Bäume	Phylo dendron Tree-View (105)
Erstellung degenerierter Primer	CODEHOP (120)
Erstellung Datenbank	DWARF (46)

7.1.1.2 Datenbanken

Tabelle 12: Tabelle der verwendeten Datenbanken

Thema	Datenbank
Proteine/Nukleotide	Swiss-Prot (24)
	Genbank (16)
	PIR (148)
	EMBL (131)
	TrEMBL (24)
	DDBJ (135)
Strukturen	PDB (17)
Literatur	NCBI (http://www.ncbi.nlm.nih.gov)

7.1.1.3 Hardware

Tabelle 13: verwendete Hardware

Komponenten	Hardware
Beowulf-Cluster, BioCortex (Linux)	
Prozessoren	128x Athlon 1800+ Dualprozessoren (AMD)
Verbindung zwischen den Rechnern	High Speed Myrinet Interconnect
Verbindung zum Server	1 Gbit-Netzwerk
Speicher-Kapazität	1TB
SGI, Octane2 (Irix 6.5)	
Prozessor	R12000
Arbeitsspeicher	256 MB

7.1.2 Bioinformatische Methoden

7.1.2.1 Sequenz-Alignments

Alignments sind die Grundlage aller Sequenzanalysen. Das einfachste Alignment, das paarweise Alignment, besteht aus zwei Sequenzen, die aufgrund der Position ihrer Nukleotide bzw. AS aneinander ausgerichtet werden. Ziel ist es dabei, möglichst viele identische Positionen nebeneinander in den zu vergleichenden Sequenzen zu finden. Die Bewertung dieses Vergleichs erfolgt mit Hilfe einer Substitutionsmatrix.

Das Multisequenz-Alignment, also das gleichzeitige analysieren mehrerer Sequenzen, liefert im Vergleich zum paarweisen Sequenz-Alignment genauere Informationen zur AS Verteilung an einzelnen Positionen. Alle Multisequenz-Alignments wurden mit ClustalX unter Verwendung der Voreinstellungen durchgeführt.

Bei der Berechnung und der Darstellung eines Alignments gibt es zwei Parameter, die bestimmt werden müssen: die Identität und die Ähnlichkeit. Die Ähnlichkeit wird anhand der verwendeten Substitutionsmatrix angegeben.

7.1.2.1.1 Globale Alignments

In den frühen 70er Jahren entwickelten Saul Needleman und Christian Wunsch einen effizienten Algorithmus zur Generierung eines global optimalen Alignments (97). Sie konstruierten zunächst eine Matrix der zwei zu vergleichenden Sequenzen. Jedes Alignment lässt sich als Pfad durch diese Matrix beschreiben, deren Zellen die Paare des Alignments beschreiben. Daher werden solche dynamischen Programmierungsalgorithmen auch als Pfad-

Matrix-Methoden bezeichnet. Der Pfad für ein globales Alignment, einem Vergleich über die gesamte Länge der Sequenzen, beginnt an einem Rand der Matrix und endet am diagonal gegenüberliegenden Rand. Für jede Zelle werden drei Werte berechnet, der sog. Score für den Vergleich der beiden AS und die Scores für den Fall eines Gaps in der einen bzw. anderen Sequenz an dieser Position. Der beste dieser drei Scores wird als Wert in die Zelle eingesetzt. Die Information, auf welche Weise dieser Wert ermittelt wurde, geht in das sog. „Backtracing“ ein, hierbei wird der mögliche Rück-Pfad durch die Matrix und somit das endgültige, optimale Alignment festgelegt.

Bei drei oder mehr Sequenzen würden Rechenzeit und Speicherbelastung exponentiell ansteigen. Als Bsp.: Ein MSA aus n Proteinen der Länge 100 AS, würde zu 100^n Sequenzvergleichen führen. Zur Reduktion des Rechenaufwands wurden sog. progressive Alignment-Methoden entwickelt. Progressive Alignment-Methoden arbeiten mit einer Abfolge von paarweisen Alignments, indem zwei ausgewählte Sequenzen verglichen werden und anschließend nach und nach die restlichen Sequenzen diesem Alignment zugefügt werden. Progressive Alignment Methoden sind heuristisch (Griechisch: *Heureka!*-Ich habe es gefunden), sie gehören zu den sog. „greedy algorithms“ (Englisch: *greedy*-habgierig, gefräßig), welche ein Problem in Fragmente zerlegen und dann für jedes Fragment nach der besten Lösung suchen, ohne jedoch das Gesamtproblem zu berücksichtigen. Im Falle von MSAs wird das komplette Alignment in mehrere paarweise Alignments aufgeteilt. Da progressive Methoden heuristisch sind, ist nicht garantiert, dass hiermit das best mögliche Alignment ermittelt wird. In der Praxis hat sich jedoch gezeigt, dass progressive Alignment Methoden effizient sind und biologisch sinnvolle Alignments erzeugen. Die unterschiedlichen progressiven Alignment Methoden unterscheiden sich in vielerlei Hinsicht: (1) Festlegung der Sequenzen des ersten paarweisen Alignments; (2) Art wie die restlichen Sequenzen dem Alignment zugefügt werden und (3) Art der Bewertung der individuellen Alignments in Bezug auf das vorhergehende Alignment. Ein Bsp. eines allgemein verwendeten Programms für progressive MSAs ist ClustalX (136).

7.1.2.1.2 Lokale Alignments

Die gängigsten Programme für Sequenz-Alignments beruhen auf einem sog. lokalen Alignment. Lokale Alignment Methoden suchen das optimale Alignment zwischen lokalen Regionen zweier Sequenzen. Lokale Alignment Methoden sind angebracht bei Sequenzen, die lokal ähnlich Regionen aufweisen. Sie werden benutzt, um Sequenzmotive, Domänen und

andere Sequenzwiederholungen zu finden, vor allem zur Datenbanksuche. Der Rechenaufwand, die große Menge an zu durchmusternden Sequenzen einer Datenbank mit globalen Methoden zu bewerkstelligen, wäre enorm. Um den Aufwand zu reduzieren wird nicht die komplette Sequenz, sondern nur Bruchstücke für das Alignment eingesetzt. Ein dynamischer Programmieralgorithmus, der lokale Alignments zweier Sequenzen durchführt, ist der Smith-Waterman Algorithmus (128), welcher dem Needleman-Wunsch Algorithmus (97) ähnelt. Ein lokales Alignment muss erweiterbar sein, um die Sequenz von Anfang bis Ende zu vergleichen. Der Score eines solchen Alignments ist kumulativ. Falls dieser Score unter einen bestimmten Wert fällt, besteht die Möglichkeit das Alignment abubrechen und ein neues Alignment zu beginnen. Auch kann das Alignment irgendwo innerhalb der Matrix enden. Weit verbreitete Programme, die lokale Alignments verwenden, sind BLAST (3) und FASTA (107).

7.1.2.2 Substitutionsmatrizen

Identität und Ähnlichkeit von Sequenzen sind wichtige Parameter bei der Berechnung von Alignments. Anhand dieser Parameter lässt sich eine Aussage über die Homologie der Sequenzen machen. Die Identität ist ein Maß für die Anzahl der identischen Positionen eines Alignments, die Ähnlichkeit wird in Abhängigkeit von der verwendeten Substitutionsmatrix angegeben. In Substitutionsmatrizen werden alle möglichen Zeichen (die 4 Basen bei Nukleotid Sequenzen, die 20 AS bei AS Sequenzen) entlang der X- und Y-Achsen notiert. Jeder Zelle einer Matrix wird ein Wert zugeordnet, der sich nach der Stärke der Ähnlichkeit zwischen den beiden Zeichen richtet.

Im wesentlichen gibt es zwei wichtige Substitutionsmatrizen, die für Sequenz-Alignments verwendet werden: **PAM** und **BLOSUM**.

Die **PAM**-Matrize (33) wurde in den 70er Jahren von Margret Dayhoff aus etwas 100 verschiedenen Proteinsequenzen, die untereinander zu mehr als 85% identisch waren, berechnet. Aus globalen Alignments wurde ermittelt, wie häufig eine AS ausgetauscht wurde, und vor allem gegen welche. Das Ergebnis war eine Matrix mit 400 Feldern (20 AS auf jeder Achse), in deren Feldern die Anzahl der beobachteten Austausche aufgeführt wurde. **1PAM** (**P**ercent **A**ccepted **M**utation) bedeutet ein Austausch pro 100 AS. **PAM**-Matrizen sind extrapolierbar und können so auf die benötigten evolutiven Abstände der Sequenzen angepasst werden.

Die **BLOSUM**-Matrix (**B**locks **S**ubstitution **M**atrix) wurden von Jorja und Steven Henikoff (61) entwickelt und basiert auf einem lokalen Alignment. Für jede Matrize wurden Alignments von Sequenzen verwendet, die zu einem bestimmten Prozentsatz identisch sind. Die Matrix BLOSUM62 wurde mit einem Alignment aus Sequenzen mit 62% Identität berechnet, die Nummer der Matrix steht also für die prozentuale Identität der Sequenzblöcke. Die **BLOSUM**-Matrizen sind nicht extrapolierbar.

Durchführung:

Alle Multisequenz-Alignments wurden unter Verwendung der voreingestellten Parameter durchgeführt.

7.1.2.3 Homologie Suche

BLAST (**B**asic **L**ocal **A**lignment **S**earch **T**ool) ist ein heuristischer Suchalgorithmus der 1990 am NCBI entwickelt wurde (3). Heuristische Verfahren sind eine Annäherung an die genaue Berechnung von Sequenz-Alignments mit dem Smith & Waterman (128) oder Needleman & Wunsch Algorithmus (97). Sie ermöglichen es, innerhalb kurzer Zeit ganze Datenbanken nach ähnlichen Sequenzen zu durchsuchen. Der BLAST-Suchalgorithmus besteht aus drei Schritten: Auffinden von Hits, Suche nach einem zweiten Hit und Ausdehnung der Hits. Im ersten Schritt, dem Auffinden von Hits wird, mit Sequenzabschnitten der Länge W (Word size) nach Treffern gesucht, die einen Score größer T (Threshold) aufweisen. Die Voreinstellungen für W liegen für Nukleotidsequenzen bei 11 und für Proteinsequenzen bei 3. Im zweiten Schritt, der Suche nach einem zweiten Hit, wird die Nachbarschaft der Treffer untersucht. Befindet sich in direkter Nachbarschaft eines Hits mit einem maximalen Abstand A (Fensterlänge) ein zweiter Hit, werden die beiden Hits im weiteren Verlauf der Suche berücksichtigt. Im dritten Schritt, der Ausdehnung der Hits, werden die Hits solange bidirektional ausgedehnt, bis sich der Score nicht mehr erhöhen lässt.

7.1.2.4 Score und E-Wert

Der **Score** (S) eines Alignments ist abhängig von der verwendeten Substitutionsmatrix und der Höhe der Strafpunkte für die Entstehung und Verlängerung von Gaps (Englisch: *Gaps-Lücken*) im Alignment. Zur Normalisierung wird der Scores mit Hilfe der, für die Substitutionsmatrizen spezifischen Karlin-Altschul-Parameter λ und K (68), in den Bit Score (S') umgerechnet.

$$S' = \frac{\lambda S - \ln K}{\ln 2}$$

Der Bit Score allein sagt jedoch nichts über die statistische Signifikanz eines Alignments aus. Der E-Wert (Erwartungswert) gibt die Zahl falsch positiver Treffer in einer zufälligen Datenbank gleicher Größe an. In seine Berechnung wird die Länge der Suchsequenz (m) und die Summe der Sequenzlängen aller in der Datenbank vorhandenen Längen (n) einbezogen.

$$E = mn 2^{-S'}$$

Der E-Wert sollte so klein wie möglich sein, denn umso geringer ist die Wahrscheinlichkeit, dass das Alignment mit diesem Bit Score durch Zufall entstanden ist.

Durchführung:

7.1.2.4.1 EH Sequenzen für die systematische Analyse

In einem ersten Schritt wurden über eine Stichwortsuche („epoxide“) in GenBank 71 unterschiedliche EH, putative EH oder Hydrolase Sequenzen ermittelt. Mit diesen Sequenzen wurden ein Multisequenz-Alignment (ClustalX) und eine phylogenetische Analyse (Tree-Puzzle) durchgeführt. Anhand dieser Analyse wurden 13 repräsentative Sequenzen gewählt (Tabelle 21) und eine BLAST-Suche unter Verwendung der Voreinstellungen durchgeführt. Die Suche ergab 75 unterschiedliche Sequenzen. Beim Vergleich mit den Sequenzen der Stichwortsuche ergaben sich insgesamt 93 unterschiedliche Sequenzen (Tabelle 22).

7.1.2.4.2 BLAST-Suchen kurzer AS-Sequenzen

Für BLAST-Suchen mit kurzen AS-Sequenz (ca. 20 AS) wurde die Substitutionsmatrix PAM30 (30% ausgetauschte AS) verwendet.

7.1.2.4.3 EH/HD Datenbank

Aus der phylogenetischen Analyse von 93 EH Sequenzen und weiterer 18, durch Stichwortsuche ermittelter, HD Sequenzen (Tabelle 23) wurden 13 Sequenzen als „Seed“ für eine BLAST-Suche zur Erstellung der Datenbank eingesetzt (Tabelle 24). Als Cutoff wurde $E = 10^{-10}$ festgesetzt. Für die Erstellung der EH/HD Datenbank wurde das automatische

Datenextraktionssystem der LED benutzt (45). Die gesammelten Daten wurden vom System auf Grundlage ihrer Sequenzähnlichkeit bewertet und in Proteineintrag, homologe Familie und Superfamilie klassifiziert. Sequenzeinträge mit Sequenzidentitäten größer als 95% wurden zu einem Proteineintrag zusammengefasst. Erhältliche Strukturinformationen wurden aus der PDB extrahiert und die Sekundärstruktur-Informationen durch das Programm DSSP (66) erstellt. Für die Überlagerung dieser Strukturen, wurden strukturell konservierte AS, wie die AS der katalytischen Triade und des Oxyanionholes verwendet.

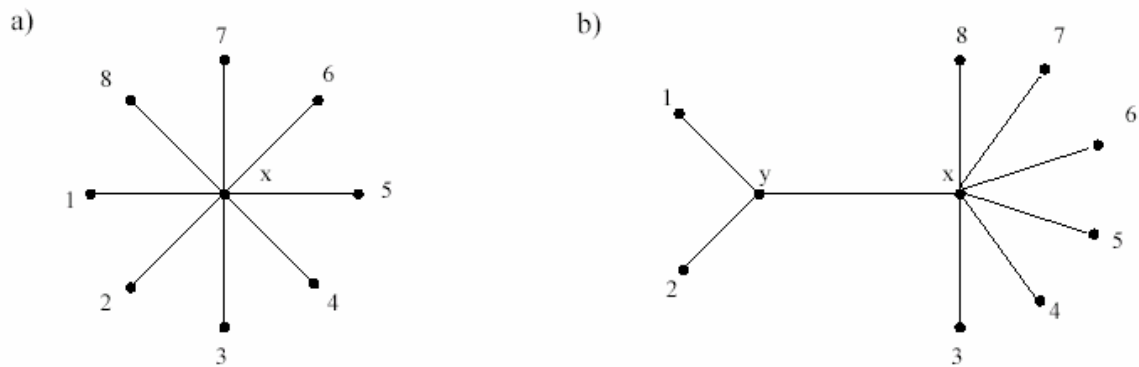
7.1.2.5 Phylogenetische Analyse

Phylogenetische Analysen versuchen, die evolutionären Beziehungen, zwischen Organismen, bzw. Nukleotid- oder Proteinsequenzen zu beschreiben.

Die Methoden zur Berechnung phylogenetischer Bäume werden in zwei Gruppen eingeteilt: distanzbasierte und merkmalsbasierte Methoden. Verwendet wurde als distanzbasierte Methode der Neighbor-Joining Algorithmus (ClustalX) und als merkmalsbasierte Methode Maximum-Likelihood (Tree-Puzzle).

7.1.2.5.1 Neighbor-Joining

Die Neighbor-Joining-Methode (NJ) (121) errechnet anhand einer Distanzmatrix einen phylogenetischen Baum. Die Distanzen werden mit paarweisen Alignments bestimmt. Ausgehend von einer Distanzmatrix werden zu Beginn alle Sequenzen an einem sternförmigen Baum verteilt. Die Astlängen orientieren sich an den Distanzen. Der erste Schritt besteht darin, einen inneren Knoten für ein Sequenzpaar so einzufügen, dass die Gesamtdistanz möglichst gering bleibt. In Abbildung 32 handelt es sich um das Sequenzpaar 1 und 2. Beim Einfügen des inneren Knotens Y muss für jedes Paar, an dem der Knoten Y plaziert werden kann, die Gesamtdistanz des Baumes berechnet werden, um die geringste Gesamtdistanz zur Positionierung von Y auszuwählen. Nach dem Einfügen muss die Distanzmatrix neu bestimmt werden. Anstatt der Distanzen zu den beiden Sequenzen des ausgewählten Paares, enthält die Matrix nun die Distanzen aller nicht ausgewählten Taxa zum inneren Knoten Y.

**Abbildung 35**

Entstehung eines Neighbor-Joining-Baumes

Für das obige Bsp. würden sich die neuen Distanzen zu dem inneren Knoten Y D_{Yk} berechnen durch:

$$D_{Yk} = \frac{D_{1k} + D_{2k}}{2}, k \in \{3,4,5,6,7,8\}$$

Die Schritte einen neuen inneren Knoten einzufügen und die Distanzmatrix neu zu bestimmen werden nun so oft wiederholt, bis ein phylogenetischer Baum entstanden ist. Nach jedem Einfügen neuer Knoten verringert sich die Anzahl der zu untersuchenden Knoten um eins. Der Baum ist dann vollständig erstellt, wenn die Anzahl der zu untersuchenden Knoten gleich drei ist.

Durchführung:

Phylogenetische Analysen nach der Neighbor-Joining-Methode wurden für die Einteilung in Superfamilien und homologe Familien bei der Erstellung der EH/HD Datenbank verwendet. Zur Berechnung der phylogenetischen Bäume mit ClustalX wurden die voreingestellten Parameter verwendet.

7.1.2.5.2 Maximum-Likelihood

Die Maximum-Likelihood-Methode (43) ist eine merkmalsbasierte Methode und bezieht sich, im Gegensatz zu distanzbasierten Methoden, nicht auf eine Distanzmatrix, sondern direkt auf das Alignment der verwendeten Sequenzen. Ziel dieses Verfahrens ist es, unter allen alternativen Bäumen, die für einen Datensatz konstruierbar sind, diejenigen zu finden, der

mit der größten Wahrscheinlichkeit die Entstehung der terminalen Sequenzen erklärt. Die Gesamtwahrscheinlichkeit („Likelihood“) eines Baumes setzt sich aus Übergangswahrscheinlichkeiten zusammen, die besagen, wie wahrscheinlich ein AS Austausch an einer bestimmten Position ist. Grundlage der Berechnung ist eine Matrix, die die Übergangswahrscheinlichkeiten der Proteine beschreibt (20x20 Punkte für AS). Unter der Voraussetzung, dass alle Mutationen unabhängig von ihrer Position in der Sequenz erfolgen, kann für jede Position im Alignment ein eigener Baum berechnet werden. Um einen Baum zu berechnen, werden die Einzelwahrscheinlichkeiten für jede einzelne Position im Alignment berechnet. Die Wahrscheinlichkeit L für den endgültigen Baum ist das Produkt der Einzelwahrscheinlichkeiten:

$$L = L_{(1)} * L_{(2)} * L_{(3)} * \dots * L_{(n)} = \prod_{i=1}^n L_{(i)}$$

Die Einzelwahrscheinlichkeiten eines Baumes, berechnen sich z.B. für die dritte Position im Alignment aus den Wahrscheinlichkeiten für die möglichen Austausche P . Für einen Baum, bestehend aus vier Sequenzen, ergeben sich 16 Möglichkeiten:

$$L_{(3)} = P_{(1)} + P_{(2)} + P_{(3)} + \dots + P_{(16)} = \sum_{i=1}^{16} P_{(i)}$$

Die Wahrscheinlichkeit für den endgültigen Maximum-Likelihood Baum wird als Logarithmus angegeben:

$$\ln L = \ln L_{(1)} + \ln L_{(2)} + \ln L_{(3)} + \dots + \ln L_{(n)} = \sum_{i=1}^n \ln L_{(i)}$$

Durchführung:

Phylogenetische Analysen nach der Maximum-Likelihood-Methode wurden für die systematische Analyse der 93 EHs und die Berechnung der phylogenetischen Bäume der Superfamilien und homologen Familien der EH/HD Datenbank eingesetzt. Zur Berechnung der phylogenetischen Bäume mit Tree-Puzzle wurden die voreingestellten Parameter verwendet.

7.1.2.6 Erstellung degenerierter Primer mit CODEHOP

Das Design familienspezifischer Primer basiert auf Multisequenz-Alignments. Hierbei werden konservierte Regionen des Proteins revers translatiert und degenerierte Primer entwickelt.

Für die Erstellung familienspezifischer Primer existieren zwei Strategien:

1) Bildung eines Pools degenerierter Primer, die über die komplette Länge der Sequenz degeneriert sind. Diese Degenerierung kann zu Fehlamplifikationen, sog. Artefakten führen. Der Schmelzpunkt solcher Primer ist schwer bestimmbar.

2) Design eines einzelnen Consensus-Primers über die hoch konservierte Region. Diese Primer sind sehr spezifisch, bei entfernt verwandten Sequenzen bindet der Primer, aufgrund zu vieler nicht mit der Template-Sequenz übereinstimmender AS, nicht.

Eine sehr erfolgreiche Strategie mittels PCR unbekannte Gene mit geringer Sequenzähnlichkeit zu bekannten Sequenzen zu amplifizieren, ist der Einsatz von Hybrid-Primern. Sie umgehen die oben genannten Probleme durch eine Kombination beider Strategien. Diese bestehen am 3'-Ende aus einem relativ kurzen degenerierten „Core“ und am 5'-Ende aus einer nicht-degenerierten „Consensus Clamp“. Der kurze, degenerierte „Core“, erniedrigt die Gesamtzahl der individuellen Primer im degenerierten Primer Pool. Die Hybridisierung des degenerierten „Cores“ mit der Template-Sequenz wird durch die nicht-degenerierte „Consensus-Clamp“ stabilisiert, welche höhere Annealing Temperaturen ohne gleichzeitiges Ansteigen des Degenerierungsgrades ermöglicht. Diese Art des Primer Designs wird „COsensus-DEgenerate Hybride Oligonucleotide Primers“ oder CODEHOP (120) genannt.

Zu Beginn der Amplifikation wird nur ein kleiner Anteil (1/Degenerierungsgrad) der degenerierten „Cores“ der Primer vollständig mit der zu amplifizierenden Sequenz übereinstimmen. Nach mehreren PCR-Zyklen, wenn Amplifikate vorheriger Runden als Template genutzt werden, erhöht sich dieser Anteil der bindenden Primer, da jetzt die, bei allen Primern identische „Consensus-Clamp“ vorhanden ist.

Durchführung:

Die für das Primer-Design verwendete Software ist im Netz unter folgender Adresse verfügbar: <http://bioinformatics.weizmann.ac.il/blocks/codehop.html>

Für die Berechnung der Primer wurden Schmelztemperaturen von ca. 60°C und ein maximaler Degenerierungsgrad von 128 verwendet.

Die Accession Codes der für das Primer Design verwendeten Sequenzen befinden sich im Anhang Tabelle 29.

7.1.2.7 Homologiemodellierung

Eine Methode, um die Tertiärstruktur eines Proteins ausgehend von der AS-Sequenz zu modellieren, ist die Homologiemodellierung. Diese Methode sagt die Struktur eines Proteins vorher, indem seine Sequenz (Target) mit der Sequenz eines homologen Proteins bekannter Struktur (Template) verglichen wird. Der erste Schritt einer Homologiemodellierung ist die Erstellung eines verlässlichen Sequenz-Alignments, so dass äquivalente AS beider Proteine übereinstimmen. Die Qualität des Alignments ist der entscheidende Faktor bei der Bestimmung der Qualität des späteren Modells. Je niedriger die Identität zwischen Target und Template, desto schlechter ist das Alignment der beiden Sequenzen und folglich auch das Modell. Aufgrund des Alignments wird der Verlauf des Backbones nach Vorgabe der Template-Struktur festgelegt. Insertionen oder Deletionen von AS, welche hauptsächlich in den variablen Loop-Regionen der Proteine vorkommen, werden anschließend eingefügt. Stabile Protein-Strukturen befinden sich in einem Zustand mit möglichst geringer potentieller Energie. Um qualitativ hochwertige Modelle zu erhalten, wird die im Modell enthaltene Energie mittels verschiedener Techniken minimiert. Ein gutes Homologiemodell kann die biologischen und biochemischen Eigenschaften eines Proteins (Funktion, Mechanismus, Substratspezifität, Inhibition) imitieren.

Durchführung:

Für die Remodellierung der EH aus *Agrobacterium radiobacter* sowie für die Homologiemodellierung der EHs aus *Glycine max*, *Mycobacterium tuberculosis*, *Corynebacterium* sp., *Streptomyces coelicolor* und *Mesorhizobium loti* wurden das Programm Spdb-Viewer für die Erstellung der Alignments und der Input-Projekt-Dateien, sowie der Swiss-Model Server für die Berechnung der Modelle verwendet.

7.1.2.7.1 Remodellierung

Als Template-Struktur wurde die Non-heme Peroxidase aus *Streptomyces aurefaciens* (1BRO), als Target-Struktur das Monomer A des PDB-Eintrages der EH aus *Agrobacterium radiobacter* (1EHY_A) verwendet. Die zu modellierenden Loop-Regionen wurden aus der PDB-Datei gelöscht und die beiden Strukturen überlagert und als Input-Projekt-Datei für den

Swiss-Model Server gespeichert. Die Berechnung wurde im „Optimise (project) mode“ durchgeführt.

7.1.2.7.2 Homologiemodelle

Verwendet wurde jeweils die passende Template-Struktur und die Sequenz der Target EH. Mit dem Spdb-Viewer wurde automatisch ein Alignment erzeugt, welches mit dem 93 EH Alignments als Vorlage korrigiert, und eine Input-Projekt-Datei für den Swiss-Model Server erstellt wurde.

7.1.2.8 Molekulardynamische Simulation

Der Begriff molekulardynamische (MD) Simulationen beschreibt die Untersuchung von Molekülen und molekularen Systemen anhand von Computermodellen, mit dem Ziel das natürliche System zu imitieren. Die Modelle bestehen aus einem Satz an Atomen, definiert als Punktmassen, mit Raumkoordinaten. Alle Moleküle eines solchen molekularen Systems weisen Wechselwirkungen untereinander auf, welche Funktionen von Position und Eigenschaften der einzelnen Atome sind. Diese Wechselwirkungen sind die zentrale Größe in einer MD Simulation und werden durch sog. Force-Fields beschrieben.

7.1.2.8.1 Ladungsberechnung

Titra (89) ist ein Programm, für elektrostatische Berechnungen in Proteinen. Für Proteine, deren Raumkoordinaten im Format der Protein Data Bank (PDB) abgelegt sind, berechnet das Programm pKa-Werte und Titrationskurven der titrierbaren AS. Informationen über die Wasser-Zugänglichkeit der ionischen AS werden mit dem Programm ACC_RUN (im Programm-Paket enthalten) berechnet. Die Protonierung eines Proteins bei bestimmten pH ist bedingt durch die pKa-Werte der titrierbaren AS (Asp, Glu, Lys, Arg, His, C- und N-Terminus).

Durchführung:

Verwendet wurde die remodellierte Struktur der *Agrobacterium radiobacter* EH (ehy_mod.pdb).

Mit dem Programm ACC_RUN wurden zwei Eingabe-Dateien erstellt:

ehy_mod.acc: enthält die initialen pKa-Werte und

ehy_mod.atom.acc: enthält Daten über die Wasser-Zugänglichkeit in \AA^2 .

Die pKa-Werte für die EH wurden im Bereich zwischen pH_{min} 6 und pH_{max} 9 mit einer Schrittweite von 0,5 mit Titra berechnet.

Als Ergebnis wurden folgende Dateien ausgegeben:

ehy_mod.tcv: enthält die Daten der Titrationskurve

ehy:mod.pks: enthält Informationen über den jeweiligen pH-Wert, AS-Name, AS-Nummer, $\text{pK}_{\text{intrinsisch}}$, $\text{pK}_{\text{Protein}}$ und Partialladung.

Die Berechnung der Partialladungen ergab, dass His22 und His275 eine Ladung von ca. 1 aufwiesen. Das Amber-Force-Field umschreibt mit HIP ein positiv geladenes His, daher wurden im PDB-File die entsprechenden His in HIP umbenannt und unter dem Namen ehy_mox.pdb gespeichert.

7.1.2.8.2 Minimierung

Die Minimierung wird durchgeführt, um deformierte Bindungslängen- und Winkel zu korrigieren. Dadurch wird die in der Struktur enthaltene Energie reduziert und das Protein befindet sich in einem lokalen Energieminimum.

Die Minimierung wurde im Vakuum mit dem Programm Sander-Classic (Amber6) mit der folgenden Eingabe-Datei durchgeführt:

```
Minimization
&cntrl
    imin=1, ntp=10, maxcyc=500, ncyc=20, ntb=0, idiel=0, scee=1.2,
&end
```

imin=1: Durchführung einer Minimierung.

ntp=20: Energiedaten wurden alle 20 Schritte in die Output-Datei geschrieben.

maxcyc=500: Die Minimierung dauerte maximal 500 Zyklen.

ncyc=20: Die Minimierung erfolgte die ersten 20 Zyklen mit der Methode „Steepest decent“ anschließend wurde die Methode „conjugate gradient“ verwendet.

ntb=0: Kein konstantes Volumen („periodic boundary“).

idiel=0: Distanzabhängige dielektrische Funktion für Vakuum-Simulationen.

scee=1.2: Multiplikationsfaktor für Partialladungen, Default-Einstellung

7.1.2.8.3 Solvatisierung und Neutralisierung des Proteins

Um das Protein mit einer Wasserhülle zu umgeben, wurde das Programm xLEaP (Amber6) verwendet. Für die Solvatisierung wurde das Protein mit einer Wasserbox (watbox216) aus TIP3 Wasser umgeben. Die Wasserbox hatte die Form eines „truncated ocathedron“ und befand sich im Abstand von 10 Å von der Proteinoberfläche. Um die Ladung des Proteins von -16 zu neutralisieren, wurden Na⁺-Ionen zugegeben. Anschließend wurde das System als Koordinaten-Datei (ehy.xyz) und Parameter-Datei (ehy.parm) gespeichert.

7.1.2.8.4 Aufheizen

Das Aufheizen der Struktur auf 300 K wurde in Wasser mit dem Programm Sander (Amber6) durchgeführt.

```

Heating
&cntrl
    nmropt=1, ntp=100, ntwr=1000, ntwx=200, ntf=2, ntc=2, ntb=1,
    nstlim=20000, ntt=1, tol=0.00000001,
&end
&wt
    type='TEMP0', istep1=0, istep2=5000, value1=10., value2=10.,
&end
&wt
    type='TEMP0', istep1=5001, istep2=15000, value1=10., value2=300.,
&end
&wt
    type='TEMP0', istep1=15001, istep2=20000, value1=300., value2=300.,
&end
&wt
    type='TAUTP', istep1=0, istep2=5000, value1=0.2, value2=0.2,
&end
&wt
    type='TAUTP', istep1=5001, istep2=20000, value1=0.4, value2=0.4,
&end
&wt
    ttype='END'
&end

```

nmropt=1: Beschränkung für zulässige Distanzen und Winkel.

ntp=100: Die berechneten Energiewerte wurden alle 100 Schritte in die Dateien „mdout“ und „mdinfo“ geschrieben.

ntwr=1000: Alle 1000 Schritte wurde eine Restart-Datei geschrieben.

ntwx=200:	Die berechneten Koordinaten wurden alle 200 Schritte in das Datei „mdcrd“ geschrieben.
ntf=2:	Keine Bindungswechselwirkungen mit Wasserstoffen.
ntc=2:	SHAKE, Bindungen mit Wasserstoff haben feste eine Größe.
ntb=1:	Konstantes Volumen („periodic boundary“).
nstlim=20000:	Anzahl der durchgeführten Schritte.
tol=0.00000001:	Relative geometrische Toleranz für das Zurücksetzen der Koordinaten bei SHAKE in Å.
TEMP0:	Änderung die Zieltemperatur.
istep1/istep2:	Grenzen Änderungen ein.
TAUTP:	“Pressure relaxion time” in ps.

7.2 Experimentelle Arbeiten

7.2.1 Material

7.2.1.1 Chemikalien und Enzyme

Tabelle 14: Verzeichnis der Hersteller der verwendeten Chemikalien

Hersteller	Chemikalie
ARK Scientific Biosystems, Darmstadt	PCR-Primer, Sequenzier-Primer
Bio Labs New England, Beverly, USA	Restriktionsenzyme
BIO-RAD Laboratories, München	LMW-Standard für SDS-PAGE
Boehringer-Mannheim, Mannheim (Roche-Diagnostics)	RNAse, Restriktionsendonukleasen, Alkalische Phosphatase (CIAP), X-Gal
Clonetec	M13-Primer
Difco-Laboratories, Detroit, USA	Trypton, Malzextrakt
Eppendorf	<i>Taq</i> -Polymerase, <i>Taq</i> -Puffer
Gibco BRL GmbH, Eggenstein	1 kb-, 1+ kb- und 100 mb-Ladder (Standard für Agarosegele), Restriktionsendonukleasen, TSB
Interactiva Biotechnologie, Ulm	PCR-Primer
Invitrogen Life Technologies	TOPO TA Cloning [®] Kit
Fluka Chemie, uchs, Schweiz	Agar, Ampicillin (Natriumsalz), Brom-phenolblau, Coomassie Brilliant Blue R250, DMF, DMSO, EDTA, Ethanol, Hefeextrakt, Lysozym, Magnesiumsulfat (wasserfrei), Natriumacetat, Natriumchlorid, PEG 6000, Saccharose, SDS, Styroloxid, Tris-(hydroxymethyl)-aminomethan, TEMED
MBI	Restriktionsendonukleasen, T4-Ligase
Merck, Darmstadt	Eisessig, DC-Platten
Qiagen, Hilden	Spin-Mini-Plasmid-Präparations-Kit, Gel-Extraktions-Kit zur Extraktion von DNA aus Agarosegelen
Perkin-Elmer, Weiterstadt	Big Dye [™] Terminator Cycle Sequencing ready Reaction Kit
Pharmacia, LKB, Freiburg	Ultrapure dNTP-Set für DNA-Sequenz-Reaktionen
Roth GmbH, Karlsruhe	Rotiphorese [®] NF Acrylamidlösung/bis-Lösung 30% für SDS-Gele
Serva Feinbiochemikalien GmbH, Heidelberg	Agarose

7.2.1.2 Geräte

Tabelle 15: Verzeichnis der Hersteller der verwendeten Geräte

Anwendung	Geräte	Hersteller
Agarose-Gelelektrophorese	DNA Sub Cell™, Mini Sub™ DNA Cell, Mini Sub™ Cell GT	BioRad
	Video Copy Processor P66E	Mitsubishi
	BWM 9X Monitor	Javelin Electronics
	UV-Leuchttisch	MWG-Biotech
Brutschränke	WTE	Binder
DNA-Sequenzierer	377A DNA Sequencer	Applied Biosystems
	Macintosh Quadra 650	Apple Macintosh
Fermenter	Bioengineering	Bioengineering AG
Steuersoftware	IRIS NT Version 4.11	Infors
Gaschromatograph	HRGC MEGA series 2	Fison Instruments, Rodano, Italien
	EL930 FID	
	FS-Cyclodex β -I/P CS Fused Silica	CS-ChromatographieService, Langerwehe
Mikroskop	Axiolab E	Zeiss
Objektive	Achroplan 10x/0,25, 40x/0,65, 100x/1,0	Zeiss
Mikrowellengerät	Micro-Chef FM A935	Moulinex
Netzgeräte	Power Pac 3000, Power Pac 300, Modell 200/2.0 Power Supply	BioRad
PAGE	Minigel-Twin G42	Biometra®
	Model 583 Gel Dryer	BioRad
PCR-Geräte	Mastercycler gradient	Eppendorf
Photometer	BioPhotometer	Eppendorf
	UV/Vis-Spektrophotometer	Pharmacia Biotech
Proteinreinigung	Akta Explorer 100	Amersham Bioscience
	Säulen	Amersham Bioscience
Inkubatoren	HT-Schüttelinkubator	Infors AG
Termomixer	Thermomixer 5436	Eppendorf
Vortexer	Vortex Genie 2	Scientific Industries
Waagen	Basic, MCI Research RC 210 D	Sartorius
	Precision Advanced	OHAU®
Zentrifugen	Eppendorf Centrifuge 5417 C	Eppendorf
	Eppendorf Centrifuge 5415 R	Eppendorf
	KR 22 i (Rotor: AK 500-11, 155 mm)	Jouan
	Sorvall RC - B (Rotor SA 600)	Du Pont Instruments

7.2.1.3 Plasmide

Tabelle 16: verwendete Plasmide

Plasmid	Größe	Charakteristik	Lieferant/Referenz
pUC18	2,7 kb	Amp ^r , lacZ-Promotor	MBI-Fermentas (140)
pCR [®] 2.1-TOPO	3,9 kb	Amp ^r , lacZ-Promotor	Invitrogen

7.2.1.4 Mikroorganismen

Tabelle 17: Verwendete Mikroorganismen

Spezies	Stamm	Eigenschaften	Referenz
<i>Escherichia coli</i>	Dh5 α	supE44 Δ lacU169(Φ 80lacZ Δ M15) <i>hsdR17</i> <i>recA1 end A1 gyrA9 thi-1relA1</i>	Clontech (57)
	One-shot	F' { <i>lacI^q Tn10 (Tet^R) mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80lacZ Δ M15 Δ lacX74 <i>recA1 deoR</i> <i>araD139</i> Δ (<i>ara-leu</i>)7697 <i>galU galK rpsL</i> (Str ^R) <i>endA1 nupG</i>	Invitrogen
<i>Streptomyces antibioticus</i> (<i>Streptomyces eurythermus</i>)	Tü4	WT	ETH 6677, NRRL 2539, DSM40014
<i>Streptomyces coelicolor</i>	A3(2)	WT	DSM65783
<i>Rhodococcus ruber</i>	LU760	WT	DSM44319

7.2.1.5 Medien, Puffer und Lösungen

Alle Angaben beziehen sich auf 1 l H₂O. Für Festmedien wurden, wenn nicht anders angegeben, 16 g Agar verwendet.

HA-Medium (Fermentation von <i>S. antibioticus</i> Tü4)	Glucose	4,0 g
	Malzextrakt	10,0 g
	Hefeextrakt	4,0 g
	Leitungswasser	ad 1000 ml
	1 M Calciumchlorid-Lsg	1 ml/l
		pH 7,3
LB-Medium (122) (modifiziert nach Luria-Bertani)	Trypton	10,0 g
	Hefeextrakt	5,0 g
	Natriumchlorid	5,0 g
		pH 7,5
SOC-Medium	Trypton	2% (M/V)
	Hefeextrakt	0,5% (M/V)
	Natriumchlorid	10 mM
	Kaliumchlorid	2,5 mM
	Magnesiumchlorid	10 mM
	Magnesiumsulfat	10 mM
	Glucose	20 mM
TSB-Medium (Anzucht von Streptomyceten)	TSB	30 g
	Leitungswasser	ad 1000 ml
TSS (27)	PEG 6000	10,0 g
	DMSO	5 ml
	1 M Magnesiumchlorid-Lsg	5 ml
	LB-Medium	ad 100 ml
		pH 6,5
Medienzusatz		
Antibiotikum (Endkonzentration)	Ampicillin	100 µg/ml

7.2.1.6 Puffer für die Agarose-Gel-Elektrophorese

50x TAE-Puffer	Tris	242 g
	Eisessig	57 ml
	0,5 M EDTA, pH 8,0	100 ml
	H ₂ O	ad 1000 ml
1%-Agarosegel	Agarose	4 g
	TAE 1x	ad 400 ml
2%-Agarosegel	Agarose	8 g
		ad 400 ml
6x DNA-Auftragspuffer	Glycerin	30% (M/V)
	Bromphenolblau	0,2% (M/V)
	EDTA (pH 7,5)	5 mM

7.2.1.7 Puffer und Lösungen für die SDS-PAGE

Lower Tris 4x	Tris	36,46 g
	SDS	0,8 g
	dH ₂ O	ad 200 ml pH 8,8
Upper Tris 4x	Tris	12,11 g
	SDS	0,8 g
	dH ₂ O	ad 200 ml pH 6,8
Trenngel (12,5%)	Acrylamidlg 30%	3,33 ml
	Lower Tris 4x	2 ml
	dH ₂ O	2,67 ml
	TEMED	4 µl
	Ammoniumpersulfat 10%	40 µl

Sammelgel (4%)	Acrylamidlg 30%	0,52 ml
	Upper Tris 4x	1 ml
	dH ₂ O	2,47 ml
	TEMED	4 µl
	Ammoniumpersulfat 10%	40 µl
5x Laufpuffer	Tris	15 g
	Glycin	72 g
	SDS	10% (M/V)
	H ₂ O	ad 1000 ml
		pH 8,3
5x Probenpuffer	Tris/HCl (pH 6,8)	320 mM
	Glycerin	50% (M/V)
	SDS	10% (M/V)
	β-Mercaptoethanol	25% (M/V)
	Bromphenolblau	0,1% (M/V)

7.2.1.8 Lösungen für die Coomassie-Färbung von SDS-Gelen

Färbelösung	Coomassie-Brilliant Blue R-250	0,1% (M/V)
	Methanol	30% (V/V)
	Eisessig	10% (V/V)
Entfärbelösung	Methanol	30% (V/V)
	Eisessig	10% (V/V)
	H ₂ O	60% (V/V)

7.2.1.9 Puffer für die Schnellisolierung von Plasmid-DNA

Resuspendierungs-Puffer	Tris/HCl (pH 7,5)	100 mM
	EDTA	10 mM
	RNaseA	250 µg/ml
Lyse-Puffer	Natriumhydroxid	0,2 M
	SDS	1,0% (M/V)
Neutralisierungs-Lsg	Kaliumacetat	29,4 g
	Eisessig	11,5 ml
	H ₂ O	28,5 ml

7.2.1.10 dNTP-Mix für die DNA-Sequenzierung

dNTP-Mix (25 mM)	dATP (100 mM)	2,5 µl
	dTTP (100 mM)	2,5 µl
	dCTP (100 mM)	2,5 µl
	dGTP (100 mM)	2,5 µl
	H ₂ O _d	90 µl

Phenol/Chloroform Phenol (pH 8) im Verhältnis 1:1 mit Chloroform mischen

LI-Lsg (73)	Tris-HCl	25 mM
	EDTA	25 mM
	Saccharose	300 mM

TE-Puffer	Tris	10 mM
	EDTA	1 mM
		pH 8

7.2.2 Molekularbiologische Methoden

7.2.2.1 Stammhaltung und Kultivierung von *E. coli*-Stämmen

7.2.2.1.1 Stammhaltung auf festen Nährböden

Um Reinkulturen zu erhalten (122) wurden Verdünnungsausstriche auf LB-Agarplatten mit Zusatz von entsprechenden Antibiotika zur Plasmidselektion hergestellt, die über Nacht bei 37°C inkubiert und anschließend mit Parafilm verschlossen bei 10°C im Kühlraum gelagert wurden. Zur langfristigen Aufbewahrung der Stämme wurden 1,5 ml einer Übernachtskultur abzentrifugiert, das Pellet in einem Gemisch aus 500 µl LB-Medium und 500 µl Glycerin (100%) aufgenommen und bei -80°C gelagert.

7.2.2.1.2 Anzucht in Flüssigmedien

Zur Durchführung von Plasmidpräparationen im kleinen Maßstab wurden 5 ml LB-Medium mit Selektionsantibiotikum mit einer Einzelkolonie oder 10 µl Glycerinkultur beimpft und über Nacht bei 37°C und 200 Upm im Schüttler inkubiert. Zur Isolierung größerer Mengen an Plasmid-DNA bzw. zur Präparation kompetenter Zellen wurden 50 ml im 250 ml Erlenmeyerkolben 1:100 angeimpft und im Schüttler bei 37°C bis zur gewünschten Zelldichte inkubiert.

7.2.2.2 Stammhaltung und Kultivierung von *Streptomyces*-Stämmen

7.2.2.2.1 Anzucht in Flüssigmedien

5 µl einer Glycerinkultur wurden in 5 ml TSB-Medium im Reagenzglas bei 30°C für 2-3 d im Schüttler inkubiert. Zur besseren Durchlüftung wurde eine Metall-Feder zugegeben. Anschließend wurde ein Schikane-Erlenmeyerkolben mit 500 ml TSB-Medium 1:100 mit dieser Vorkultur angeimpft und wiederum für 2-3 d bei 30°C im Schüttler inkubiert.

7.2.2.2.2 Fermentation

Für die Vorkultur wurden 0,5 l TSB-Medium für 2-3 Tage bei 30 °C im Schikane-Kolben kultiviert. Für die Fermentation von *Streptomyces antibioticus* Tü4 wurden die in Tabelle 18 angegebenen Parameter verwendet. Es wurden zwei Fermentationen mit Ausbeuten von je ~900 g feuchter Zellmasse durchgeführt. Während der Fermentation erfolgte keine Probennahme, die Reinheit der Kultur und Aktivität des Enzyms wurden erst nach Beendigung der Fermentation überprüft. Auf Kontaminationen wurde mikroskopisch geprüft, die Aktivität der fermentierten *Streptomyces antibioticus* Tü4 Zellen wurde mittels GC überprüft. Die Zellen wurden durch Zentrifugation geerntet.

Tabelle 18: Parameter der Fermentation

Parameter	
Rührertyp	Scheibenrührer
Rührergeschwindigkeit	300 Upm
Belüftung	17 l/min
pH-Wert	7,3
Temperatur	30°C
Volumen	27 l
Fermentationsdauer	~72 h
Medium	HA-Medium + 1 mM Ca ₂ Cl

7.2.2.3 Zellaufschluss von *E. coli*-Zellen

Um die Zellen aufzuschließen wurden 50 ml eines Kultivierungsansatzes 20 min bei 4000 Upm zentrifugiert, in 2 ml 50 mM Natriumphosphatpuffer pH 7,2 resuspendiert, mit Lysozym versetzt und für 45 min auf Eis inkubiert. Anschließend wurden die Zellen mittels Ultraschall 2x 30 s mit 1 min Pause behandelt (80 W, Pulse 35% s-1). Die aufgeschlossenen Zellen wurden bei 4°C und 4000 Upm für 20 min zentrifugiert, der Überstand abgenommen und auf Eis zur weiteren Verwendung aufbewahrt.

7.2.2.4 Zellaufschluss von *Streptomyces*-Zellen

7.2.2.4.1 Zellaufschluss mit Ultraschall

Um die Zellen aufzuschließen wurden 1 g feuchte Zellmasse mit 1 ml 0,5 M Natriumphosphatpuffer pH 8 resuspendiert. Für den Aufschluss wurde die Probe 5x 1 min mit je 1 min Pause mit Ultraschall behandelt (80 W, Pulse 35% s-1). Der Aufschluss erfolgte auf Eis.

7.2.2.4.2 Zellaufschluss für die Proteinreinigung

230 g feuchte Zellmasse von *Streptomyces antibioticus* Tü4 wurden ad 430 ml mit 50mM Kaliumphosphat, pH 8,0 suspendiert. Der pH des Gemisches wurde mit 1 M Kaliumhydroxid-Lösung auf pH 8 eingestellt. Zur Homogenisierung wurde anschließend im Ultraturrax bei voller Leistung suspendiert. Das Suspendat wurde für 15 min bei 12000 Upm bei 4°C zentrifugiert. 530 ml Überstand (Leitfähigkeit 5,4 mS/cm) wurden mit 1 M Kaliumhydroxid-Lsg auf pH 8 eingestellt und als Auftrag für die Anionenaustausch-Chromatographie eingesetzt.

7.2.2.5 Zellaufschluss von *Rhodococcus*-Zellen

6 g feuchte Zellmasse wurden in 30 ml Kaliumphosphatpuffer pH 8 resuspendiert, mit 60 ml Glaskugeln (\varnothing 0,1-0,2 mm) in eine 100 ml-Schottflasche gegeben und ad 100 ml mit Kaliumphosphatpuffer pH 8 aufgefüllt. Die Mischung wurde für 15 min auf einer Eis-Wasser-Mischung gekühlt und anschließend mit einem Scheibenrührer auf voller Leistung für 20 min gerührt. Um die Glaskugeln von den aufgeschlossenen Zellen zu trennen, wurde bei 4000 Upm und 4°C für 15 min zentrifugiert.

7.2.2.6 DNA-Isolierung

7.2.2.6.1 Gesamt-DNA-Isolierung aus Actinomyceten

(modifiziert nach Altenbuchner und Cullum, 1984(2))

Die Zellen wurden 2 Tage in TSB-Medium und 1% Glycerin angezogen. 10-20 ml der Kultur wurden 5 min bei 8000 Upm abzentrifugiert. Das Pellet wurde in 5 ml LI-Lösung unter Zugabe von 10 mg/ml Lysozym resuspendiert und 30-45 min bei 37°C unter gelegentlichem Mischen inkubiert. Anschließend wurden 2 ml SDS 5% zugegeben, die Mischung kurz gevortext und 10 min bei 60°C inkubiert. Nachdem die Lsg auf RT abgekühlt war, wurden 2 ml Phenol/Chloroform (pH 8) zugegeben. Nach vorsichtigem Durchmischen wurde die Suspension 5 min auf Eis gestellt. Anschließend wurde 15 min bei 10.000 Upm zentrifugiert und der Überstand abgenommen. Bei Trübung des Überstandes musste die Phenolisierung wiederholt werden. Zu dem Überstand wurden 1/10 Vol. 3 M Natriumacetat und 1Vol. Isopropanol (-20°C) zugegeben und vorsichtig gemischt. Die ausgefallene DNA wurde mit einer Pipettenspitze gefischt und in 70% Ethanol gewaschen. Die DNA wurde in 0,5-1 ml TE resuspendiert und bei 4°C aufbewahrt.

7.2.2.6.2 Gesamt-DNA-Isolierung mit Anionenaustauschersäulen

Zur Isolierung genomischer DNA wurde das QIAmp-Kit, Qiagen verwendet. Es wurden 2 ml einer *Streptomyces antibioticus*-Kultur eingesetzt und nach Angaben des Herstellers über Anionentauscher-Säulen gereinigt.

7.2.2.7 Plasmidisolierung

7.2.2.7.1 Isolierung von Plasmid-DNA aus *E. coli*

Mit dieser Methode wurde Plasmid-DNA für Anwendungen isoliert, die weder besonders viel noch besonders reine DNA erforderten (z.B. analytischer Restriktionsverdau). Sie beruht auf dem Prinzip der alkalischen Lyse (20). Nach vollständiger Lyse der Zellmembran mit Hilfe von SDS wird die entfaltete RNA durch Natriumhydroxid und RNaseA hydrolysiert, während chromosomale und Plasmid-DNA denaturiert werden. Durch Absenken des pH-Wertes renaturiert die Plasmid-DNA, während die chromosomale DNA zusammen mit den Proteinen ausfällt und abzentrifugiert werden kann. Die Plasmid-DNA kann anschließend durch Isopropanolfällung konzentriert und gereinigt werden.

Durchführung:

1,5 ml einer Übernachtskultur von *E. coli* wurden 1 min bei 14.000 Upm abzentrifugiert. Das entstandene Pellet wurde in 200 µl Resuspendierungspuffer aufgenommen. Durch die Zugabe von 200 µl Lyse-Puffer und anschließendem vorsichtigen Mischen wurde die Lsg. klar. Nach Zugabe des Neutralisationspuffers wurden die Proteine und die chromosomale DNA gefällt. Anschließend wurde für 5 min bei RT und 14.000 Upm zentrifugiert. Der Überstand wurde abgenommen und daraus die Plasmid-DNA durch Zugabe von 420 µl Isopropanol gefällt. Nach 10 min Inkubation bei RT wurde die Probe für 20 min bei 4°C und 14.000 Upm zentrifugiert, das Pellet mit 200 µl 70%igem Ethanol gewaschen, zentrifugiert und im Vakuum getrocknet. Das Pellet wurde in 20 µl TE-Puffer resuspendiert und bei -20°C aufbewahrt.

7.2.2.7.2 Isolierung reiner Plasmid-DNA aus *E. coli*

Die Spin-Mini-Plasmidisolation (Spin-Mini, Qiagen) wurde durchgeführt, um bis zu 20µg sehr reiner Plasmid-DNA zu gewinnen (Sequenzierung, PCR). Sie beruht ebenfalls auf dem Prinzip der alkalischen Lyse. Nach Abtrennung der restlichen Zellbestandteile wird die Plasmid-DNA über eine Ionenaustauschersäule gereinigt.

Durchführung:

2 ml einer Übernachtskultur von *E. coli* wurden 1 min bei 14.000 Upm zentrifugiert. Die Spin-Mini-Plasmidpräparation wurde mit dem QIAprep Spin Miniprep Kit (Qiagen) nach Angaben des Herstellers durchgeführt.

7.2.2.7.3 Isolierung größerer Mengen an Plasmid-DNA aus *E. coli*

Für die Isolierung größerer Mengen an DNA wurden 15 ml einer *E. coli*-Kultur 10 min bei 4000 Upm zentrifugiert. Die Mini-Plasmid-Präparation wurde mit dem Peqlab Mini-Kit II nach Angaben des Herstellers durchgeführt.

7.2.3 Enzymatische Methoden

7.2.3.1 Verdau von DNA mit Restriktionsendonukleasen

Zum Restriktionsverdau wurden 0,5-1 µg DNA in einem Reaktionsvolumen von 10-50 µl mit 1-2 Units des jeweiligen Restriktionsenzym in dem, den Angaben des Herstellers entsprechenden Reaktionspuffer umgesetzt. Die Inkubationszeit der Spaltungsansätze betrug 1-2 h bei der für das jeweilige Enzym empfohlenen optimalen Temperatur.

7.2.3.2 Ligation mit T4-Ligase

Die T4-Ligase katalysiert die Bildung von Phosphodiesterbindungen zwischen benachbarten 3'-Hydroxy- und 5'-Phosphat-Enden doppelsträngiger DNA. Die T4-Ligase erkennt außerdem getrennte Phosphodiesterbindungen in doppelsträngiger DNA und kann diese schließen („nick repair“).

Zur Ligation wurde die Insert-DNA im Vergleich zur Vektor-DNA etwa in 3-5 fachem Überschuss zugegeben. Die Inkubation wurde für 1 h bei 37°C oder ÜN bei RT durchgeführt.

7.2.3.3 T/A-Klonierung

Bei der sogenannten T/A-Klonierung wird die Eigenschaft z.B. der *Taq*-Polymerase, ein Adenin an das 3'-Ende des amplifizierten Fragments anzuhängen (terminale Transferase Aktivität), ausgenutzt. Das TOPO TA Cloning[®]-Kit (Invitrogen) stellt einen linearisierten Vektor (pCR[®]2.1-TOPO) zur Verfügung, der an seinen Enden, zum Adenin komplementäre 3'-Thymidin-Überhänge besitzt. Die kovalent an den Vektor gebundene Topoisomerase verbindet Insert und Vektor und führt so die Ligation durch (127).

Durchführung

Der Reaktionsansatz wurde, wie in Tabelle 15 dargestellt pipettiert und für 30 min bei RT inkubiert. Anschließend wurde der Ansatz bis zur Transformation auf Eis gelagert. Die Transformation erfolgte nach der Heatshock-Methode in kompetente One Shot[®] *E. coli* Zellen (Invitrogen).

Tabelle 19: Reaktionsansatz für die T/A-Klonierung

Reagenz	Volumina
PCR-Produkt	4 µl
Salz-Lsg (Invitrogen)	1 µl
pCR [®] 2.1-TOPO-Vektor	1 µl
Endvolumen	6µl

7.2.3.4 Herstellung kompetenter Zellen

Zur Herstellung kompetenter Zellen wurde eine ÜN-Kultur 1:100 mit LB-Medium verdünnt und bei 37°C im Schüttler bis zu einer OD₅₇₈ von 0,4-0,7 inkubiert. Anschließend wurden die Zellen bei 3000 Upm und 4°C für 15 min zentrifugiert. Das Pellet wurde in 2 ml kaltem TSS-Puffer resuspendiert. Zur Lagerung wurden die aliquotierten kompetenten Zellen (200 µl) in flüssigem Stickstoff eingefroren und bei -80°C gelagert.

7.2.3.5 Transformation nach der Heatshock-Methode

7.2.3.5.1 *E. coli* DH5α-Zellen

Die kompetenten Zellen wurden auf Eis aufgetaut. 200 µl kompetente Zellen wurden mit der zu transformierenden DNA gemischt und für 20 min auf Eis inkubiert. Anschließend wurde der „Heatshock“ für 40 sek in einem 42°C warmen Wasserbad durchgeführt. Die Zellen wurden danach unmittelbar auf Eis gebracht und mit 800 µl LB vermischt. Die Inkubation erfolgte für 1h bei 37°C. Die Zellen wurden für 1 min bei 8000 Upm zentrifugiert, der Überstand bis auf 50-100 µl verworfen, das Pellet resuspendiert, auf LB/Amp⁺-Platten ausgestrichen und über Nacht bei 37°C inkubiert.

7.2.3.5.2 One Shot[®]-Zellen (Invitrogen)

Die kompetenten Zellen wurden auf Eis aufgetaut. Zu einem Aliquot kompetenter Zellen wurden 2 µl Reaktionsansatz (siehe 7.2.3.3) gegeben und für 30 min auf Eis inkubiert. der „Heatshock“ wurde für 30 sek in einem 42°C warmen Wasserbad durchgeführt und die Zellen anschließen sofort auf Eis gebracht und mit 250 µl SOC-Medium vermischt. Die Inkubation erfolgte für 1 h bei 37°C. 50-100 µl des Transformationsansatzes wurden auf LB-Amp⁺-Platten ausgestrichen und über Nacht bei 37°C inkubiert.

7.2.3.6 Polymerase-Kettenreaktion

Die PCR (engl. Polymerase Chain Reaction), entwickelt 1985 von K. Mullis, dient zur spezifischen Amplifikation einer Ausgangs-DNA (Template) (93). Dabei wird ein Zyklus von drei Reaktionsschritten bis zu 40 mal wiederholt. Der erste Reaktionsschritt eines jeden Zyklus ist die Denaturierung des Template-DNA-Doppelstranges bei ca. 95°C. Dabei entstehen DNA-Einzelstränge, an die sich im zweiten Reaktionsschritt, dem Annealing bei 45-65°C (abhängig von der Schmelztemperatur der jeweiligen Primer/Template-Kombination), jeweils ein Primer anlagert. Im dritten Schritt, der Extension, werden durch Verlängerung der Primer am 3'-Ende mittels thermostabiler Polymerasen bei 72°C zwei neue Stränge synthetisiert, die bei den folgenden Zyklen wieder als Template zur Verfügung stehen. Die Menge an gewünschtem DNA-Fragment steigt exponentiell über die Reaktion, weshalb eine geringe Menge an Ausgangs-DNA ausreichend ist.

Durchführung

Als Ausgangs-DNA wurde entweder genomische DNA, eine Spin-Mini-Plasmidpräparation oder eine Kolonie eingesetzt. Die Primer wurden in einer Konzentration von 100 pm/μl in sterilem H₂O gelöst. Die vom Hersteller der verwendeten Polymerasen empfohlenen Reaktionspuffer wurden benutzt. Die Zusammensetzung typischer PCR-Ansätze ist in Tabelle 20 dargestellt.

Tabelle 20: Zusammensetzung eines typischer PCR-Ansätze und Kolonie-PCR-Ansätze

Komponente	Volumen pro Reaktion (gen. DNA)	Volumen pro Reaktion (Kolonie-PCR)
10 x Polymerase Puffer	5 μl	2,5 μl
DMSO	5 μl	-
dNTP-Mix (je 2,5 mM)	2 μl	1 μl
Primer 1	2 μl	1 μl
Primer 2	2 μl	1 μl
Ausgangs-DNA	1-2 μl	Kolonie
DNA-Polymerase	0,5 μl	0,2 μl
H ₂ O	ad 50 μl	ad 25 μl

7.2.3.7 Sequenzierung

Die Sequenzierung von DNA erfolgt nach der von Frederick Sanger entwickelten enzymatischen Kettenabbruch-Methode (123). Dabei werden unterschiedlich lange DNA-Fragmente durch kontrollierte Unterbrechung der Synthese mittels spezifischer Terminatoren (2',3'-Didesoxynukleosid-5'-triphosphate, ddNTPs), die zusätzlich zu den 2'-Desoxynukleotiden zum Reaktionsansatz gegeben werden, erzeugt. Da die ddNTPs an ihrer 3'-Position keine OH-Gruppe besitzen, kann ein entstehendes DNA-Fragment nach dem statistischen Einbau dieses Terminators nicht mehr durch Ausbildung weiterer Phosphodiester-Bindungen zum nächsten anzuhängenden 2'-Desoxynukleotid verlängert werden und es kommt zum Kettenabbruch.

Bei der automatischen DNA-Sequenzierung mit dem *377-DNA-Sequencer* von Applied Biosystems sind die vier ddNTPs mit vier unterschiedlichen Fluoreszenzfarbstoffen markiert, und werden zusammen in einer Spur zusammen auf einem Gel aufgetrennt (47, 115). Die Fluoreszenzfarbstoffe werden durch zwei Argonlaser mit Emissionsbanden von 488 und 514 nm angeregt und detektiert. Diese Vierfarbentechnik erhöht sowohl den Durchsatz als auch die Geschwindigkeit der Datenerfassung, die bereits während des Gellaufs durch einen angeschlossenen Rechner erfolgt.

Die modernste Methode bedient sich der AmpliTaq[®]-Polymerase, FS und wird als „Cycle Sequencing“ bezeichnet. Diese Methode erlaubt die zur PCR (93) analoge zyklische Sequenzierung, bei der die enzymatische Sequenzierung mit einem Primer über mehrere Zyklen durchgeführt werden, was geringere Mengen an Ausgangs-DNA erfordert, die Automatisierung vereinfacht und die Sequenzierung doppelsträngiger DNA ermöglicht. Die AmpliTaq[®]-Polymerase,FS ist ein hitzestabiles Enzym aus dem Bakterium *Thermus aquaticus*, FS ist die Abkürzung für Fluoreszenz-Sequenzierung. Es handelt sich bei diesem Enzym um eine Doppelmutante der Wildtyp AmpliTaq[®]-Polymerase. Durch eine Mutation (G46D) verliert das Enzym seine 5',3'-Nuklease-Aktivität. Die zweite Mutation im aktiven Zentrum des Enzyms (F667Y) führt zu einer Veränderung seiner Akzeptanz gegenüber Didesoxynukleotiden. Diese können dadurch leichter eingebaut werden.

7.2.3.7.1 DNA-Sequenzierung mit dem 377A-DNA-Sequencer

Gießen des Polyacrylamid-Gels

Um die exakte Erfassung der Fluoreszenzsignale während des Gellaufs zu gewährleisten, wurden die Glasplatten vor dem Gießen des Gels gründlich mit einem Detergenz (Alkanox) gewaschen und mit dest. Wasser gespült. Danach wurden sie mit Isopropanol (p.a.) nachbehandelt und mit einem fusselreifen Zellstoff (Kimwipe) trockengerieben. Anschließend wurden die Abstandhalter plaziert und die Glasplatten zusammengesetzt.

Tabelle 21: 6%iges Polyacrylamidgel

Rotiphorese [®] NF-Harnstoff	18 g
Rotiphorese [®] NF 40%ige Acrylamidlg (19:1)	6,6 ml
Rotiphorese [®] NF-10x TBE-Puffer	5,0 ml
H ₂ O _{bidest.}	ad 50 ml

Der Harnstoff (Tabelle 21) wurde durch Rühren gelöst und anschließend durch einen 0,2 µm Filter filtriert. Die Lsg wurde in ein Becherglas überführt, 24 µl TEMED und 180 µl Ammoniumpersulfatlg (10%) zugegeben, vermischt und anschließend sofort zwischen die durch Abstandhalter getrennten Gasplatten gegossen, der „Vorkamm“ eingesetzt und die Platten mit Klammern während der zweistündigen Polymerisationszeit abgedichtet. anschließend wurden die Klammern entfernt, der Vorkamm durch einen „Haifisch-Kamm“ ersetzt und das Gel in den Sequenzierer eingebaut. Auf die Überprüfung des Basisdetektionssignals folgte ein einstündiger Vorlauf.

Mit dem Dye Terminator Cycle Sequencing Kit mit AmpliTaq[®] DNA Polymerase wurde folgender Reaktionsansatz für die Sequenzierung verwendet (Tabelle 22).

Tabelle 22: Zusammensetzung der Sequenzierungsreaktion

Komponente	Volumen pro Reaktion	Konzentration/Menge
Terminator Ready Reaction Mix	4-8 µl	1x
Primer	1 µl	3,3 pmol
DNA	5 µl	33-500 ng
H ₂ O	ad 20 µl	

Die Sequenzierungsreaktion wurde in einem Thermocycler unter folgenden Bedingungen durchgeführt (Tabelle 23).

Tabelle 23: PCR-Programm der Sequenzierung

Programmschritt	Denaturierung	Annealing	Extension
1	95°C, 7 min		
2 (25x)	95°C, 40 sek	55°C, 30 sek	60°C, 4 min
3			60°C, 4 min

Probenaufarbeitung

Der PCR-Ansatz durch eine Sephadex-Matrix gefällt. Die DNA wurde getrocknet und anschließend in 4 µl Formamid/25 mM EDTA (4:1) aufgenommen. Die Proben wurden 3 min bei 95°C denaturiert, auf Eis abgekühlt und nach kurzem Zentrifugieren in die Taschen des in das Sequenziergerät eingebauten Polyacrylamidgels pipettiert. Die Elektrophorese erfolgte nach dem vom Hersteller empfohlenen Programmparametern (2500 V, 40 mA, 30 W) für 12 h. Die bereits vom Rechner erfassten und prozessierten Sequenzdaten wurden anschließend mittels geeigneter Software ausgewertet.

7.2.4 Analytische Methoden

7.2.4.1 Bestimmung der Epoxidhydrolase-Aktivität

7.2.4.1.1 Chirale Kapillar-Gaschromatographie

Prinzip

Die Gaschromatographie ist eine sehr empfindliche Trennmethode, mit der sich Gase und flüchtige Substanzen bestimmen lassen, insofern sie sich unzersetzt verdampfen lassen. Ein inertes Trägergas strömt durch eine thermostabile Säule, die zu analysierenden Substanzen werden im Trägergasstrom detektiert. Typische Trägergase sind Helium und Stickstoff, die stationäre Phase besteht üblicherweise aus Aluminium- oder Siliziumoxid.

Durchführung

Die Messungen der Aktivitätstests wurden mit einem HRGC MEGA series 2 Gaschromatographen mit einem EL930 FID (Temperatur: 200°C) durchgeführt. Als Trennsäule wurde eine FS-Cyclodex β -I/P CS Fused Silica Kapillarsäule (Heptakis-(3-*O*-acetyl-2,6-di-*O*-pentyl)- β -cyclodextrin) mit einer Länge von 50 m und einem Innendurchmesser von 0,25 mm verwendet. Als Trägergas wurde Wasserstoff verwendet. Die Injektion erfolgte durch einen Splitinjektor (Temperatur: 200°C; Splitverhältnis: 1:100). Durch den Vergleich mit den Retentionszeiten von Standardsubstanzen konnten die Peaks den Enantiomeren des Substrats Styroloxid und des Produkts 1,2-Phenylethandiol zugeordnet werden. In Tabelle 24 sind die verwendeten Temperaturprogramme für Styroloxid und 1,2-Phenylethandiol und die eingestellten Gasrückdrücke aufgezeigt.

Tabelle 24: GC-Temperaturprogramme für Styroloxid und 1,2-Phenylethandiol

Substanz	Temperaturprogramm	Gasrückdruck
Styroloxid	75°C, isotherm	130 kPa
1,2-Phenylethandiol	150°C, isotherm	100 kPa

Reaktionsgemisch

1,5 ml Rohextrakt wurden mit 50 μ l DMSO und 5 μ l Styroloxid vermischt und ÜN bei 30°C unter dem Abzug geschüttelt. Für die gaschromatographische Analyse wurde eine Extraktion mit 2 x 300 μ l Diethylether durchgeführt. Jeder Ansatz wurde zur Trocknung mit einer Spatelspitze wasserfreiem Natriumsulfat versetzt.

7.2.4.1.2 In der Mikrotiterplatte

Prinzip:

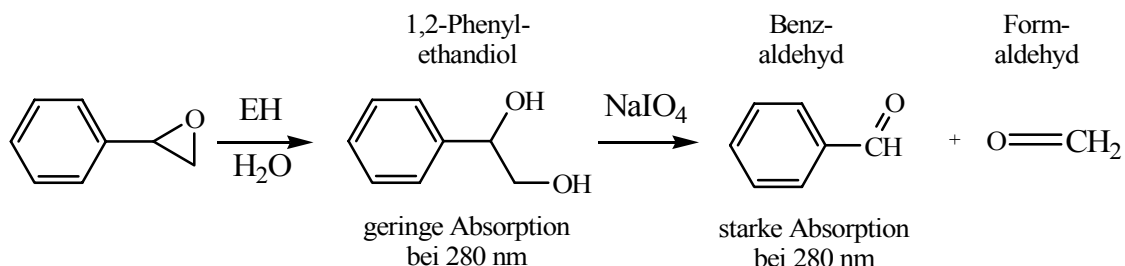


Abbildung 36:

Reaktionsprinzip des Aktivitätstests in der Mikrotiterplatte

Der Test auf EH-Aktivität (38) beruht auf dem unterschiedlichen Absorptionsverhalten des Produkts, 1,2-Phenylethandiol und Benzaldehyd bei einer Wellenlänge von 280 nm. Aus Styroloxid wird durch die Aktivität der EH das schwach absorbierende 1,2-Phenylethandiol gebildet, anschließend wird dieses durch die reduzierende Wirkung von Natriumperiodat in das stark absorbierende Benzaldehyd und Formaldehyd gespalten.

Durchführung:

In einer Mikrotiterplatte (MTP) wurden zu 90 µl einer Probe, bzw. für die Negativkontrolle 1 M Kaliumphosphat-Puffer pH 8,0, 5 µl 1 M Kaliumsulfatpuffer pH 8,0 sowie 5 µl einer 0,25 M Styroloxidlösung in DMSO gegeben. Anschließend wurde für 2-20 h bei RT inkubiert. Nach der Inkubation wurden je 20 µl/Well entnommen, in eine neue UV-stabile MTP überführt und mit 100 µl 0,1 M Natriumacetat pH 5 versetzt. Die gesamte MTP wurde für die Ermittlung der Null-Werte verwendet und bei einer Wellenlänge von 280 nm vermessen. Anschließend wurden 5 µl einer 10 mM Natriumperiodatlösung in 0,1 M Natriumacetat pH 5,0 zugegeben, die MTP für 1 min inkubiert und wiederum bei einer Wellenlänge von 280 nm vermessen. Durch den Vergleich der Messwerte vor und nach der Natriumperiodatzugabe konnte die Zunahme an Benzaldehyd und somit die EH-Aktivität ermittelt werden.

7.2.4.2 Proteinreinigung

Die Reinigung der EH aus *Streptomyces antibioticus* Tü4 wurde durch die BASF AG, Ludwigshafen durchgeführt. Es wurden Ionenaustausch- und Ausschluss-Chromatographie verwendet.

Prinzip:

7.2.4.2.1 Ionenaustausch-Chromatographie

Die Grundlage für die Ionenaustausch-Chromatographie (engl. ion exchange chromatography, IEC) ist die kompetitive Wechselwirkung geladener Ionen. Ein Probenmolekül konkurriert mit Salz-Ionen um die geladenen Positionen auf einer Ionenaustauscher-Matrix. In einem ersten Schritt bindet das Molekül an die fixierten Ladungen der stationären Phase, im zweiten Schritt erfolgt die Verdrängung und Elution des Proteins durch die steigende Salzkonzentration des Eluenten. Für die Reinigung der EH aus *S. antibioticus* Tü4 wurde eine Anionenaustausch-Chromatographie verwendet.

7.2.4.2.2 Hydrophobe Interaktionschromatographie

Bei der hydrophoben Interaktionschromatographie (engl. hydrophobic interaction chromatography, HIC) adsorbieren die unpolaren Oberflächenregionen von Proteinen bei hoher Salzkonzentration an die schwach hydrophoben Liganden einer stationären Phase. Die Elution der Proteine erfolgt durch eine Verringerung der Salzkonzentration. Die HIC kombiniert die Eigenschaften einer nicht-denaturierenden Salzpräzipitation mit der Separationskraft der Chromatographie.

7.2.4.2.3 Molekularsieb-Chromatographie

Die Ausschluss-Chromatographie trennt gelöste Moleküle nach ihrer Größe und basiert auf der unterschiedlichen Permeation der Analyten in ein poröses Trägermaterial mit kontrollierter Porengröße. Für das Trennverhalten ist das hydrodynamische Volumen der Probenmoleküle verantwortlich. Moleküle ab einer bestimmten Größe können nicht in die Poren des Trenngels eindringen und eluieren zusammen mit der Lösungsmittelfront im Ausschlussvolumen V_0 . Kleinere Moleküle bewegen sich nicht nur zwischen den einzelnen Teilchen der stationären Phase, sondern dringen außerdem auch in ihre Poren ein. Dadurch erfahren sie eine Verzögerung und ihr Elutionsvolumen V_m entspricht der Summe des

internen Porenvolumens und des Partikelzwischenraums. Die kleinsten Komponenten haben somit die längste Aufenthaltsdauer in den Poren und werden zuletzt eluiert. Die mobile Phase dient nur als Lösungsmittel und hat keinen unmittelbaren Einfluss auf die Trennung.

Durchführung:

Anionenaustausch-Chromatographie

Eine Q-Sepharose FF-Säule (Pharmacia, Durchmesser: 2,6 cm, Volumen: 100 ml) wurde in 50mM Kaliumphosphat, pH 8,0 äquilibriert. (Fließgeschwindigkeit 10ml/min, Fraktionsgröße 1 Minute) 530 ml Zellaufschluss (mit Kaliumhydroxidpuffer auf pH 8 eingestellt) wurden als Auftrag auf die Säule gegeben und die Säule mit 600 ml 50mM Kaliumphosphat, pH 8,0 gewaschen. Für die Elution wurde ein Gradient von 500 ml 50mM Kaliumphosphat, pH 8,0 nach 500 ml 50mM Kaliumphosphat, pH 8,0, 0,5M NaCl angelegt. Anschließend wurde die Säule mit 300 ml 50mM Kaliumphosphat, pH 8,0, 0,5M NaCl gewaschen. Die aktiven Fraktionen 101-110 mit einem Gesamtvolumen von 82 ml wurden gesammelt.

Molekularsieb-Chromatographie

Die aktiven Fraktionen der Anionenaustausch-Chromatographie (82 ml) wurden auf pH 7,5 korrigiert und mit 42,3 g Ammoniumsulfat (80% Sättigung) gefällt. Der pH wurde während der Fällung bei 7,5 konstant gehalten. Nach 10 min rühren bei RT wurde die Trübung bei 1200 Upm (Sorval) und 4°C für 20 min abzentrifugiert und der Überstand verworfen. Das Pellet wurde in 8,2 ml 50mM Kaliumphosphat, pH 8,0 gelöst. Eine Superdex 200-Säule (Pharmacia, Durchmesser: 2,6 cm, Volumen: 240 ml) wurde in 240ml 50mM Kaliumphosphat, pH 8,0 äquilibriert (Fließgeschwindigkeit 4ml/min, Fraktionsgröße 1min). Die aktiven Fraktionen 25-27 mit einem Gesamtvolumen von 23,5 ml wurden gesammelt.

Anionenaustausch-Chromatographie (Mono Q)

Die aktiven Fraktionen der Ausschluss-Chromatographie (23,5 ml) wurden auf drei Aliquots mit Volumina von 7,5-8 ml aufgeteilt. Eine Mono-Q Chromatographie-Säule (Pharmacia, Volumen: 1 ml) wurde mit 50mM Kaliumphosphat, pH 8,0 äquilibriert. Ein Aliquot wurde aufgegeben und mit einem linearen Gradienten der Länge 75 ml nach 50mM Kaliumphosphat, pH 8,0, 0,5M NaCl (Fließgeschwindigkeit: 0,75 ml/min) eluiert. Die aktiven Fraktionen der drei Mono-Q-Läufe, jeweils Fraktion 59-67, mit einem Gesamtvolumen von 18 ml wurden gesammelt.

Hydrophobe Interaktions-Chromatographie (TSK Phenyl)

Die aktiven Fraktionen der Ausschluss-Chromatographie (18 ml) wurden wieder auf drei Aliquots mit Volumina von 5 ml aufgeteilt. Dazu wurde je 1,13g Ammoniumsulfat gegeben. Dadurch wurde eine 49%ige Sättigung der Lösung erreicht. Diese Lösung wurde auf eine TSK Phenyl Chromatographie-Säule (Tosohaas, Volumen: 3,8 ml), äquibriert in 20mM Kaliumphosphat, pH 7,5, 40%ige Sättigung an Ammoniumsulfat, gegeben. Die Säule wurde in einem linearen Gradienten nach 20mM Kaliumphosphat, pH 7,5 eluiert (0,75ml/min, 100 Minuten, 1 min Fraktionen). Die aktiven Fraktionen 89-92 wurden gesammelt.

7.2.4.3 In-Gel Verdau der Proteinbande mit Trypsin

Die Coomassie-Blau gefärbte Bande wurde aus dem Agarose-Gel ausgeschnitten, zerkleinert und in ein Eppendorfgefäß überführt. Dann wurde das ca. 10 fache Bandenvolumen Wasser zugegeben und 30 min geschüttelt. Der Überstand wurde verworfen, der Vorgang einmal wiederholt. Anschließend wurde das ca. 10 fache Bandenvolumen Methanol/Wasser zugegeben, 5min geschüttelt und zentrifugiert. Nach 1 min. Ultraschallbad wurde der Überstand verworfen, der Schritt einmal wiederholt. Im nächsten Schritt wurde wieder ca. das 10 fache Bandenvolumen Wasser zugegeben, 5min geschüttelt, zentrifugiert und der Überstand verworfen. Dann erfolgte die Zugabe eines ca. 10fachen Bandenvolumens Acetonitril/Wasser, es wurde 5min geschüttelt und zentrifugiert. Nach 1 min. Ultraschallbad wurde der Überstand verworfen, und der Vorgang einmal wiederholt. Anschließend wurden die Gelstücke im einer Speed Vac ca. 10min getrocknet. Die Gelstücke wurden mit Spaltpuffer (Trypsin und Protein im Verhältnis 1:1 in einer Mischung aus 10µl 1%iger Essigsäure gelöst, 100µl 100mM Ammoniumhydrogencarbonat und 5%Acetonitril bei pH 8.5 gelöst) gerade überschichtet. Gut geschüttelt, zentrifugiert und über Nacht bei 37°C inkubiert. Für die folgenden Pipettiervorgänge des Überstandes wurde die gleiche Pipettenspitze verwendet. Der Überstand 1 wurde in ein neues Eppendorfgefäß überführt. Anschließend wurden 50µl 70%iges Acetonitril und 1% (V/V) TFA zu den Gelstücken gegeben, geschüttelt und nach 5 min. Ultraschallbad, zentrifugiert. Überstand 2 wurde abgenommen und zu Überstand1 gegeben. Die beiden letzten Schritte wurden jeweils einmal wiederholt. Die vereinigten Überstände wurden in einer Speed Vac bis auf eine leichte Feuchte einrotiert.

7.2.4.4 N-terminale Sequenzanalyse: Edman-Abbau

Der Edman-Abbau ist ein zyklischer Prozess, bei dem in jedem Reaktionszyklus vom N-terminalen Ende die endständige AS abgespalten und identifiziert wird. Die Reaktion besteht aus drei Schritten: Kupplung, Spaltung und Konvertierung.

Im ersten Schritt, der Kupplung, wird an die freie N-terminale Aminogruppe das „Edman-Reagenz“ Phenylisothiocyanat (PITC) gekoppelt, es entsteht ein disubstituierter Thioharnstoff (Phenylthiocarbamoylpeptid, PTC-Peptid). Im zweiten Schritt, der Spaltung, wird das getrocknete PTC-Peptid mit wasserfreier Säure behandelt. Hierbei greift der Schwefel des PITC nukleophil an der Carbonylgruppe der ersten AS an. Bei diesem Vorgang spaltet sich die AS als heterozyklisches Derivat einer Anilinothiazolinon (ATZ)-AS ab. Im dritten Schritt, der Konvertierung, wird die chemisch instabile ATZ-AS zu einem stabilen Derivat, der Phenylthiohydantoin-AS umgesetzt und mittels HPLC und durch Vergleich mit einer Standard-AS-Mischung identifiziert.

Die N-terminale Sequenzierung wurde durch die BASF AG, Ludwigshafen durchgeführt.

8 Anhang

8.1 Accession Codes der verwendeten Sequenzen

Tabelle 25: Ausgangssequenzen mikrosomaler (mik.) und cytosolischer (cyt.) EHs für die BLAST-Suche der systematischen Analyse.

Homologe Familie	Accession Code	Organismus	
Insekten, mik.	AAB18243	<i>Trichoplusia ni</i>	
<i>C. elegans</i> , mik.	NP_505811	<i>Caenorhabditis elegans</i>	
Säugetier, cyt.	NP_001970	<i>Homo sapiens</i>	
Pflanzen1, cyt.	AAB02006	<i>Nicotiana tabacum</i>	
Pflanzen2, cyt.	CAA55294	<i>Glycine max</i>	
Bakterien1, cyt.	1EHY	<i>Agrobacterium radiobacter</i>	
Bakterien2, cyt.	NP_103292	<i>Mesorhizobium loti</i>	
Bakterien3, cyt.	C83216	<i>Pseudomonas aeruginosa</i>	putativ
Bakterien4, cyt.	CAC37878	<i>Streptomyces coelicolor</i>	putativ
Bakterien5, cyt.	NP_334552	<i>Mycobacterium tuberculosis</i> CDC1551	
Bakterien6, cyt.	NP_396231	<i>Agrobacterium tumefaciens</i> C58 (Cereon)	
Bakterien7, cyt.	CAA11900	<i>Corynebacterium</i> sp.	
Bakterien8, cyt.	NP_107141	<i>Mesorhizobium loti</i>	

Tabelle 26: 93 EH Sequenzen der systematischen Analyse

Accession Code	Organismus	Accession Code	Organismus
cytosolische EHs		mikrosomale EHs	
Säugetier EHs		Säugetier EHs	
NP_001970	<i>Homo sapiens</i>	AAA52389	<i>Homo sapiens</i>
AAG14968	<i>Homo sapiens</i>	CAA68486	<i>Homo sapiens</i>
AAG14967	<i>Homo sapiens</i>	P04068	<i>Oryctolagus cuniculus</i>
A47504	<i>Mus musculus</i>	BAB31132	<i>Mus musculus</i>
P34914	<i>Mus musculus</i>	NP_034275	<i>Mus musculus</i>
BAB22083	<i>Mus musculus</i>	P07687	<i>Rattus norvegicus</i>
NP_031966	<i>Mus musculus</i>	NP_036976	<i>Rattus norvegicus</i>
NP_075225	<i>Rattus norvegicus</i>	BAA19200	<i>Sus scrofa</i>
pflanzliche EHs		Insekten EHs	
AAK50099	<i>Arabidopsis thaliana</i>	CT34977	<i>Drosophila melanogaster</i>
AAL38771	<i>Arabidopsis thaliana</i>	NP_611385	<i>Drosophila melanogaster</i>
AAL49778	<i>Arabidopsis thaliana</i>	NP_611386	<i>Drosophila melanogaster</i>
NP_180242	<i>Arabidopsis thaliana</i>	XP_080944	<i>Drosophila melanogaster</i>
NP_190669	<i>Arabidopsis thaliana</i>	Q25489	<i>Manduca sexta</i>
NP_567228	<i>Arabidopsis thaliana</i>	AAB18243	<i>Trichoplusia ni</i>
CAA55294	<i>Glycine max</i>	AAB88192	<i>Trichoplusia ni</i>
T07145	<i>Glycine max</i>	C. elegans EHs	
Momo	<i>Momordica charantia</i> (BASF)	NP_504650	<i>Caenorhabditis elegans</i>
AAB02006	<i>Nicotiana tabacum</i>	T33960	<i>Caenorhabditis elegans</i>
AAL58264	<i>Oryza sativa</i>	Pilze/Hefen EHs	
AAL58266	<i>Oryza sativa</i>	IQO7B	<i>Aspergillus niger</i>
AAL58275	<i>Oryza sativa</i>	CAB59812	<i>Aspergillus niger</i>
AAL58278	<i>Oryza sativa</i>	AAF64646	<i>Rhodotorula glutinis</i>
AAL58281	<i>Oryza sativa</i>	AAF18956	<i>Xanthophyllomyces dendrorhous</i>
BAA84626	<i>Oryza sativa</i>	bakterielle EHs	
BAA84627	<i>Oryza sativa</i>	AAK25053	<i>Caulobacter crescentus</i>
T07043	<i>Solanum tuberosum</i>	AAK78696	<i>Clostridium acetobutylicum</i>
AAA81891	<i>Solanum tuberosum</i>	NP_107140	<i>Mesorhizobium loti</i>
T07044	<i>Solanum tuberosum</i>	AAL17927	<i>Mycobacterium smegmatis</i>
T07048	<i>Solanum tuberosum</i>	CAB99443	<i>Stigmatella aurantiaca</i>
C. elegans EHs			
NP_497268	<i>Caenorhabditis elegans</i>		
bakterielle EHs			
AAK90870	<i>Agrobacterium tumefaciens</i>		
AAL46181	<i>Agrobacterium tumefaciens</i>		

CAA73331	<i>Agrobacterium tumefaciens</i>		
NP_396231	<i>Agrobacterium tumefaciens</i>		
NP_532494	<i>Agrobacterium tumefaciens</i>		
NP_534160	<i>Agrobacterium tumefaciens</i>		
1EHY	<i>Agrobacterium radiobacter</i>		
NP_388739	<i>Bacillus subtilis</i>		
AAF78792	<i>Bradyrhizobium japonicum</i>		
NP_420043	<i>Caulobacter crescentus</i>		
CAA11900	<i>Corynebacterium</i> sp.		
H75259	<i>Deinococcus radiodurans</i>		
BAB49078	<i>Mesorhizobium loti</i>		
NP_103292	<i>Mesorhizobium loti</i>		
NP_104428	<i>Mesorhizobium loti</i>		
NP_107141	<i>Mesorhizobium loti</i>		
NP_109474	<i>Mesorhizobium loti</i>		
NP_302492	<i>Mycobacterium leprae</i>		
CAA17992	<i>Mycobacterium tuberculosis</i>		
NP_218134	<i>Mycobacterium tuberculosis</i>		
NP_334552	<i>Mycobacterium tuberculosis</i>		
NP_336446	<i>Mycobacterium tuberculosis</i>		
NP_488261	<i>Nostoc</i> sp.		
C83216	<i>Pseudomonas aeruginosa</i>		
F83145	<i>Pseudomonas aeruginosa</i>		
G83279	<i>Pseudomonas aeruginosa</i>		
H83384	<i>Pseudomonas aeruginosa</i>		
AAL19286	<i>Salmonella typhimurium</i>		
NP_436749	<i>Sinorhizobium meliloti</i>		
CAB38503	<i>Streptomyces coelicolor</i>		
CAB62690	<i>Streptomyces coelicolor</i>		
CAB76359	<i>Streptomyces coelicolor</i>		
CAB94628	<i>Streptomyces coelicolor</i>		
CAB95984	<i>Streptomyces coelicolor</i>		
CAC37878	<i>Streptomyces coelicolor</i>		
T36559	<i>Streptomyces coelicolor</i>		
S77308	<i>Synechocystis</i> sp.		

Tabelle 27: 18 HD Sequenzen, verwendet für eine phylogenetische Analyse, für die Auswahl der repräsentativen HDs für die BLAST-Suchen zur Erstellung der EH/HD Datenbank.

Accession Code	Organismus	
A49896	<i>Sphingomonas paucimobilis</i>	
CAB45532	<i>Mycobacterium bovis</i>	
NP_106032	<i>Mesorhizobium loti</i>	
AAC15838	<i>Rhodococcus rhodochrous</i>	
CAB65289	<i>Mycobacterium</i> sp. GP1	
NP_536182	<i>Agrobacterium tumefaciens</i> str. C58 (U. Washington)	
AAK91024	<i>Agrobacterium tumefaciens</i> str. C58 (Cereon)	
NP_640591	<i>Xanthomonas axonopodis</i>	
AAM35127	<i>Xanthomonas axonopodis</i>	
NP_635611	<i>Xanthomonas campestris</i>	
NP_299247	<i>Xylella fastidiosa</i>	
AAL01057	<i>Photobacterium profundum</i>	putativ
AAK46638	<i>Mycobacterium tuberculosis</i> CDC1551	
CAC41377	<i>Mycobacterium avium</i>	
NP_419991	<i>Caulobacter crescentus</i> CB 15	putativ
B43718	<i>Xanthobacter autrophicus</i>	
NP_336339	<i>Mycobacterium tuberculosis</i> CDC1551	
NP_484237	<i>Nostoc</i> sp. PCC 7120	

Tabelle 28: Ausgangssequenzen für die BLAST-Suchen zur Erstellung der EH/HD Datenbank und ihre ursprüngliche Klassifizierung in Superfamilien und homologe Familien.

Superfamilie	homologe Familie	Gattung	Spezies	Datenbank	Accession Code
cyt. EH	bacterial	<i>Bacillus</i>	<i>subtilis</i>	REF	NP_388739
cyt. EH	bacterialII	<i>Agrobacterium</i>	<i>radiobacter</i>	REF	NP_532494
cyt. EH	plantsI	<i>Arabidopsis</i>	<i>thaliana</i>	GB	AAL38771
cyt. EH	plantsII	<i>Oryza</i>	<i>sativa</i>	GB	AAL58278
cyt. EH	C. elegansI	<i>Caenorhabditis</i>	<i>elegans</i>	REF	NP_497268
cyt. EH	mammaliaI	<i>Mus</i>	<i>musculus</i>	PIR	A47504
mik. EH	fungi_bacteria	<i>Mesorhizobium</i>	<i>loti</i>	REF	NP_107140
mik. EH	insects	<i>Manduca</i>	<i>sexta</i>	SP	Q25498
mik. EH	C. elegansII	<i>Caenorhabditis</i>	<i>elegans</i>	PIR	T33960
mik. EH	mammaliaII	<i>Rattus</i>	<i>norvegicus</i>	SP	P07687
HD	HF1	<i>Xanthobacter</i>	<i>autropicus</i>	SP	P22643
HD	HF2	<i>Sphingomonas</i>	<i>paucimobilis</i>	PIR	A49896
HD	HF3	<i>Photobacterium</i>	<i>profundum</i>	GB	AAL01057

Tabelle 29: Homologe Familien, die zur Erstellung der familienspezifischen Primer verwendet wurden.

Primer Gruppe	Accession Code	Primer Gruppe	Accession Code	Primer Gruppe	Accession Code
mSäuger	AAA52389	cSäuger	A47504	cBakterien1	AAF78792
	AAA59580		AAG14967		NP_104428
	BAA19200		AAG14968		NP_107141
	BAB31132		BAB22083		NP_436749
	CAA68486		NP_001970	cBakterien2	H75259
	NP_034275		NP_031966		NP_388739
	NP_036976		NP_075225		NP_488261
	P04068		NP_497268		NP_497268
	P07687		P34914	cBakterien3	CAA17992
mInsekten	AAB18243	cPflanzen1	AAB02006		CAB38503
	AAB88192		AAL58264		NP_302492
	CT34977		AAL58266	cBakterien4	AAK90870
	NP_611385		AAL58275		AAL46181
	NP_611386		AAL58278		CAB62690
	Q25489		AAL58281		CAB94628
	XP_080944	cPflanzen2	AAA81891		H83384
mPilze	AAF18956		AAK50099		NP_103292
	AAF64646		AAL38771		NP_109474
	CAB59812		AAL49778		NP_396231
mBakterien	AAK25053		BAA84626	cBakterien5	NP_218134
	AAK78696		CAA55294		NP_336446
	AAL17927		NP_180242		NP_420043
	NP_107140		NP_190669		NP_532494
	NP_436753		NP_567228		NP_534160
			T07043		T36559
			T07044		
			T07048		
			T07045		

8.2 Sequenzen

8.2.1 Primer

Tabelle 30: Verwendete Primer: Alle Primer wurden in 5'-3'-Ausrichtung aufgeführt, unter Verwendung der üblichen Bezeichnungen der vier Basen und degenerierter Positionen.

Anwendung	Primer	Sequenz
250 bp-Fragmentes	250_F	cgggatcccgcacggcttcccgga
	250_R	cgggatcccgscgccccagttgtg
CODEHOP: cBakterien5	cB5_F	tggctgctgtgccaygggnttycng
	cB5_R	gggcctctctgctggrdcartgncc
Reverse-Primer <i>S. coelicolor</i> EH	Sc_R	gctctagagcaggaggagccatgccgcag
N-Terminus	N-Term_F	atggtsgagcacmgsatggtsgac
Reverse-Primer zum Aufbau des C-Terminus	C-Term_R	tcagatcctctgctgccccgcagcccggtcaggaagt cgaccagcggcggttcacccccttcggctgctcctgctg
M13-Sequenzier-Primer	M13_F	caggaaacagctatgac
	M13_R	actggccgctggtttac

8.2.2 Sequenzen der CODEHOP-Fragmente

8.2.2.1 *Streptomyces antibioticus* Tü4

8.2.2.1.1 Sequenz des 250 bp-Fragments (236 bp)

EcoRI

GGATCCCGCACGGCTTCCCGGAGTCGGGGTACGCTGGCGGCACCAGCTGGCGGCACTCGCCTCGG
CCGGTTACCGCGGTGGCGATCGATGTGCGGGGGTACGGCCGTTTCGTCCAAGCCGGAGGCCGTGGA
GGCGTACCGGATGCTCGACCTGATCGCGGACAACGTGCCGTGGTGGAGGCGCTGGGCCAGCGTTCC
GCCGTGGTTCGTCGGCCACAACCTGGGGCGCCGGATCC

EcoRI

8.2.2.1.2 DNA-Sequenz des CODEHOP-Fragments (771 bp)

TTTGGTGCTGCTGTGCCATGGGTTTCTGAGTCCTGGCACTCCTGGCACCCAGTTCGGCCCGCTG
GCCGCCGCGGGCTTCCGGGTGGTCGCCCCGACCAGCGCGGATACGGGCGCAGCGACCATCCCGA
GGCCGTCGACGCGTACACCATCCTGCATCTGGTCGGCGATGTCGTCGGGCTGATCCGGGAGCTGGG
CGAGGAGAAGGCGTACGTGGTCGGGCACGACTGGGGCGCGCCGGTTCGCCTGGCACACCCGCGCTGC
TGCGGCCGACCTGGTGC GCGGGGTGGCCGGTCTGAGCGTGCCGCCCGCTTCAGGGGCTCGCGGC
CGCCGCTGCCC GCCATGGACCGGATGTTCCGGCGCCGCTTCTACTGGA ACTACTTCAACCGTCCCG
GGGTGCGGGACGCCGAGTTCGCCGAGGACACCCGACCGCGCTGCGCAAGTTCTTCTACTGGGCTT
CCGGGGACGCCCCCGCGCCGGGCGGAAGCAGCCGCTGGTCGACCCCGAGCGCGGCTGGCTCGCG
GACATGGCCGACCCGAGGTGCTGCCGGAGTGGTTCACCGAGGAGGATCTCGACGCGCTCACGGA
GAGCTTCTCCAGGGCTTACCCGGGGCCCTGAACTGGTACCGCAACCTCGACCGGAACTGGGAGCT
GACCGCCCCCTGGCACGGCGCCGTCGTCACCCCGCCGGCCCTGTACGTCTACGGCGACCGTGACGT
GGTCCCCGCCTTCCCCGGCACGCCGGAGCTCATCGAGCGGCTGCCCTCGCTGATGCCAACCTGTG
GCGGGAGCCGCTGAAGCTGGCCGGTGC G GACACTGGCCCCAGCAGGAGAGGCCCAAGGGCGGG
GGG

8.2.2.1.3 AS-Sequenz des CODEHOP-Fragments (284 AS)

PLVLLCHGFPESWHSWHRQFGPLAAAGFRVVAPDQRGYGRSDHPEAVDAYTILHLVGDVVGLIRELGE
EKAYVVGHDWGAPVAWHTALLRPDLVRGVAGLSVPPFRGSRPPLPAMDRMFGGRFYWNYFNRPV
ADAEFAEDTRTALRKFFYWASGDAPGAGRKQPLVDPERGWLADMADPEVLPEWFTTEEDLDALTESFS
QGFTGALNWYRNLDNRN WELTAPWHGAVVTPPALYVYGDRD VVPAFPGTPELIERLPSLMPNLWREPL
KLAGCGHWPQQERPKG

8.2.2.2 *Rhodococcus ruber* (LU760)

8.2.2.2.1 DNA-Sequenz des CODEHOP-Fragment (703 bp)

TGGTGCTGCTGTGCCATGGTTTTCCGGACACCCATGCAATCTGGCGGCTCCAGGTCCCGGCGCTGG
CGGCCAGGGGTTTTCCGGGTCATCGCGCCCGACATGCGTGGCTACGGCAAGACCGACGCGCCGCC
GCTGCCGGCGCCTACGCCGTCGAATTCCTGTGCGCCGACATCCTCGGCCTGATGGACCAGCTGGGC
CTGGAGAAGGCCTCGCTCGTGGGCCACGACTGGGGCGGGCTGGTCCGGCTGGCAGCTGTGCATGAA
TGCGCCGCAGCGCTTCGAGCGCTTTGTGCGCCCTGTCCACCGGACACCCGGCGGCGATCGCCAATGC
CGGCGTCACCCAGCAAATGCGCTTCTGGTACATCCTGGGTTTCCTGATGCCGGTGGTGGCCGAACA
CGCGATCCGCGCCAACGACTGGTTCTTCTGCGCCAGATGACGCGCCGTCCGGAACAGGTGGCGAT
CTGGCGGGCCGCGCTGGAGCCGGAAGGGCGCCTGACGGCGGCGCTGAACTACTACCGCGCCAACC
TCAAGCTGGGCCGGCCGCACCACTGGCGCCCGGTGCAAGTGCCGGTATGGGCGTCTGGAGCGAC
CGCGACCCGGCGCTGGGCGAGAAGCAGATGCTGGATTCCGGTCCACCACTGCCGCGGGGCTTCCGC
TACGAGCGCCTGAGCGGGGCGGGCACTGGACCCAGCAGGAGAGGCC

8.2.2.2.2 AS-Sequenz des CODEHOP-Fragments (215 AS)

THAIWRLQVPALAARGFRVIAPDMRGYGKTDAPPAAGAYAVEFLCADILGLMDQLGLXKASLVGHDW
GGLVGWQLCMNAPQRFERFVALSTGHPAAIANAGVTQQMRFWYILGFLMPVVAEHAIRANDWFFLRQ
MTRRPEQVAIWRAALEPEGRLTAALNYYRANLKLGRPHHWRPVEVPVMGVWSDRDPALGEKQMLDS
VHHCRAGRYERLSG

9 Literatur

1. Abdel-Mogib M, Albar HA, Batterjee SM. 2002. Chemistry of the genus *Plectranthus*. *Molecules* 7
2. Altenbuchner J, Cullum J. 1984. DNA amplification and an unstable arginine gene in *Streptomyces lividans* 66. *Mol Gen Genet* 195: 134-138
3. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-33402
4. Arahira M, Nong VH, Udaka K, Fukazawa C. 2000. Purification, molecular cloning and ethylene-inducible expression of a soluble-type epoxide hydrolase from soybean (*Glycine max* [L.] Merr.). *Eur J Biochem* 267: 2649-2657
5. Arand M, Hemmer H, Durk H, Baratti J, Archelas A, et al. 1999. Cloning and molecular characterization of a soluble epoxide hydrolase from *Aspergillus niger* that is related to mammalian microsomal epoxide hydrolase. *Biochem J* 34: 273-280
6. Arand M, Wagner H, Oesch F. 1996. Asp333, Asp495, and His523 form the catalytic triad of rat soluble epoxide hydrolase. *J Biol Chem* 271: 4223-4229
7. Archelas A, Furstoss R. 2001. Synthetic applications of epoxide hydrolases. *Curr Opin Chem Biol* 5: 112-119
8. Argiriadi MA, Morisseau C, Hammock BD, Christianson DW. 1999. Detoxification of environmental mutagens and carcinogens: structure, mechanism, and evolution of liver epoxide hydrolase. *Proc Natl Acad Sci U S A* 96: 10637-10642
9. Armstrong KM, Fairman R, Baldwin RL. 1993. The (i, i + 4) Phe-His interaction studied in an alanine-based alpha-helix. *J Mol Biol* 230: 284-291
10. Armstrong RN. 1999. Kinetic and chemical mechanism of epoxide hydrolase. *Drug Metab Rev* 31: 71-86
11. Baker D, Sali A. 2001. Protein structure prediction and structural genomics. *Science* 294: 93-96.
12. Barth S, Fischer M, Schmid RD, Pleiss J. 2003. The database of epoxide hydrolases and haloalkane dehalogenases: one structure, many functions. *submitted*
13. Barth S, Fischer M, Schmid RD, Pleiss J. 2003. Sequence and structure of epoxide hydrolases: a systematic analysis. *Proteins* accepted
14. Baxevanis AD. 2003. The Molecular Biology Database Collection: 2003 update. *Nucleic Acids Res* 31: 1-12
15. Beetham JK, Grant D, Arand M, Garbarino J, Kiyosue T, et al. 1995. Gene evolution of epoxide hydrolases and recommended nomenclature. *DNA Cell Biol* 14: 61-71
16. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL. 2002. GenBank. *Nucleic Acids Res* 30: 17-20
17. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, et al. 2002. The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* 58: 899-907
18. Bezalel L, Hadar Y, Cerniglia CE. 1997. Enzymatic mechanisms involved in phenantrene degradation by the white rot fungus *Pleurotus ostreatus*. *Appl Environ Microbiol* 63: 2495-2501
19. Bezalel L, Hadar Y, Fu PP, Freeman JP, Cerniglia CE. 1996. Metabolism of phenanthrene by the white rot fungus *Pleurotus ostreatus*. *Appl Environ Microbiol* 62: 2547-2553
20. Birnboim HC, Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* 7: 1513-1523
21. Bjedov I, Tenailon O, Gerard B, Souza V, Denamur E, et al. 2003. Stress-induced mutagenesis in bacteria. *Science* 300: 1404-1409

22. Blée E, Schuber F. 1992. Occurrence of fatty acid epoxide hydrolases in soybean (*Glycine max*). Purification and characterization of the soluble form. *Biochem J* 282: 711-714
23. Blée E, Schuber F. 1993. Biosynthesis of Cutin Monomers - Involvement of a Lipooxygenase Peroxygenase Pathway. *Plant Journal* 4: 113-123
24. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, et al. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* 31: 365-370
25. Bornscheuer UT, Kazlauskas RJ. 1999. *Hydrolases in organic synthesis*: Wiley-VCH
26. Choi WJ, Huh EC, Park HJ, Lee EY, Choi CY. 1998. Kinetic resolution for optically active epoxides by microbial enantioselective hydrolysis. *Biotechnol Tech* 12: 225-228
27. Chung CT, Niemela SL, Miller RH. 1989. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A* 86: 2172-2175
28. Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L. 1998. New goals for the U.S. Human Genome Project: 1998-2003. *Science* 282: 682-689
29. Cousin X, Hotelier T, Giles K, Toutant JP, Chatonnet A. 1998. aCHEdb: the database system for ESTHER, the alpha/beta fold family of proteins and the Cholinesterase gene server. *Nucleic Acids Res* 26: 226-228
30. Cronin A, Mowbray S, Durk H, Homburg S, Fleming I, et al. 2003. The N-terminal domain of mammalian soluble epoxide hydrolase is a phosphatase. *Proc Natl Acad Sci U S A* 100: 1552-1557
31. Damborsky J, Nyandoroh MG, Nemeč M, Holoubek I, Bull AT, Hardman DJ. 1997. Some biochemical properties and the classification of a range of bacterial haloalkane dehalogenases. *Biotechnol Appl Biochem* 26: 19-25
32. Dang. Hai Sy. The Nucleotide Database. in *Vorbereitung*
33. Dayhoff MO, Schwartz RM, Orcutt BC. 1978. *Atlas of protein sequence and structure*.: Natl Biomed Res Found
34. Debernard S, Morisseau C, Severson TF, Feng L, Wojtasek H, et al. 1998. Expression and characterization of the recombinant juvenile hormone epoxide hydrolase (JHEH) from *Manduca sexta*. *Insect Biochem Mol Biol* 28: 409-419
35. Demirjian DC, Shah PC, Moris-Varas F. 1999. Screening for novel enzymes. *Topics in Current Chemistry* 200: 1-29
36. D'Esposito M, Pilia G, Schlessinger D. 1994. BLOCK-based PCR markers to find gene family members in human and comparative genome analysis. *Hum Mol Genet* 3: 735-740
37. Dickerson RE. 1980. Evolution and gene transfer in purple photosynthetic bacteria. *Nature* 283: 210-212
38. Doderer K, Lutz-Wahl S, Hauer B, Schmid RD. 2003. Spectrophotometric assay for epoxide hydrolase activity toward any epoxide. *Anal Biochem* 321: 131-134
39. Eck RV, Dayhoff MO. 1966. *Atlas of Protein Sequence and Structure*.: National Biomedical Research Foundation
40. Edqvist J, Farbos I. 2003. A germination-specific epoxide hydrolase from *Euphorbia lagascae*. *Planta* 216: 403-412
41. Faber K, Mischitz M, Kroutil W. 1996. Microbial epoxide hydrolases. *Acta Chem Scand* 50: 249-258
42. Fahey JW, Zhang Y, Talalay P. 1997. Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens. *Proc Natl Acad Sci U S A* 94: 10367-10372
43. Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17: 368-376

44. Fischer M, Henke E, Schmid RD, Pleiss J. *in Vorbereitung*
45. Fischer M, Pleiss J. 2003. The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Res* 31: 319-21
46. Fischer M, Schmid RD, Pleiss J. 2003. DWARF: Data warehouse for proein families. *in Vorbereitung*
47. Freeman M, Baehler C, Spotts S. 1990. Automated laser-fluorescence sequencing. *Biotechnology (N Y)* 8: 147-148
48. Fretland AJ, Omiecinski CJ. 2000. Epoxide hydrolases: biochemistry and molecular biology. *Chem Biol Interact* 129: 41-59
49. Futujama DJ. 1986. *Evolutionary biology*: Sinauer Associates
50. Gibas C, Jambeck P. 2002. *Einführung in die praktische Bioinformatik*: O'Reilly
51. Gong PF, Xu JH, Tang YF, Wu HY. 2003. Improved Catalytic Performance of *Bacillus megaterium* Epoxide Hydrolase in a Medium Containing Tween-80. *Biotechnol Prog* 19: 652-654
52. Griffin MJ. 1999. Regulation of rat liver epoxide hydrolase by tightly bound phosphoinositides. *Proc Okla Acad Sci* 79: 1-6
53. Grogan G, Roberts SM, Willetts AJ. 1996. Novel aliphatic epoxide hydrolase activities from dematiaceous fungi. *FEMS Microbiol Lett* 141: 239-243
54. Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18: 2714-2723
55. Guo A, Durner J, Klessig DF. 1998. Characterization of a tobacco epoxide hydrolase gene induced during the resistance response to TMV. *Plant J* 15: 647-656
56. Haeggstrom JZ. 2000. Structure, function, and regulation of leukotriene A4 hydrolase. *Am J Respir Crit Care Med* 161: 25-31
57. Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* 166: 557-580
58. Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22: 160-174
59. Heikinheimo P, Goldman A, Jeffries C, Ollis DL. 1999. Of barn owls and bankers: a lush variety of alpha/beta hydrolases. *Structure Fold Des* 7: 141-146
60. Heinemann FS, Ozols J. 1984. The covalent structure of hepatic microsomal epoxide hydrolase. II. The complete amino acid sequence. *J Biol Chem* 259: 797-804
61. Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* 89: 10915-10919
62. Holmquist M. 2000. Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr Protein Pept Sci* 1: 209-235
63. Hosoya N, Hatayama A, Irie R, Sasaki H, Katsuki T. 1994. Rational design of Mn-Salen epoxidation catalysts: Preliminary results. *Tetrahedron* 50: 4311-4322
64. Jacobs MH, Van den Wijngaard AJ, Pentenga M, Janssen DB. 1991. Characterization of the epoxide hydrolase from an epichlorohydrin-degrading *Pseudomonas* sp. *Eur J Biochem* 202: 1217-1222
65. Janssen DB, Scheper A, Dijkhuizen L, Witholt B. 1985. Degradation of halogenated aliphatic compounds by *Xanthobacter autotrophicus* GJ10. *Appl Environ Microbiol* 49: 673-677
66. Kabsch W, Sander C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22: 2577-2637
67. Kamal A, Damayanthi Y, Rao MV. 1992. Stereoselective synthesis of (S)-propanol amines: lipase catalyzed opening of epoxides with 2-propylamine. *Tetrahedron: Asymmetry* 3: 1361-1364

68. Karlin S, Altschul SF. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc Natl Acad Sci U S A* 87: 2264-2268
69. Katsuki T, Martin VS. 1996. Asymmetric epoxidation of allylic alcohols: The Katsuki-Sharpless epoxidation reaction. *Org React (N.Y.)* 48: 1-299
70. Kazlauskas RJ, Weber HK. 1998. Improving hydrolases for organic synthesis. *Curr Opin Chem Biol* 2: 121-126
71. Keiser KC, Brandt KS, Silver GM, Wisnewski N. 2002. Cloning, partial purification and in vivo developmental profile of expression of the juvenile hormone epoxide hydrolase of *Ctenocephalides felis*. *Arch Insect Biochem Physiol* 50: 191-206
72. Khlebodarova TM, Gruntenko NE, Grenback LG, Sukhanova MZ, Mazurov MM, et al. 1996. A comparative analysis of juvenile hormone metabolizing enzymes in two species of *Drosophila* during development. *Insect Biochem Mol Biol* 26: 829-835
73. Kieser T. 1984. Factors affecting the isolation of CCC DNA from *Streptomyces lividans* and *Escherichia coli*. *Plasmid* 12: 19-36
74. Kiyosue T, Beetham JK, Pinot F, Hammock BD, Yamaguchi-Shinozaki K, Shinozaki K. 1994. Characterization of an *Arabidopsis* cDNA for a soluble epoxide hydrolase gene that is inducible by auxin and water stress. *Plant J* 6: 259-269
75. Kloosterman M, Elferink VHM, van Iersel J, Roskam JH, Meijer EM, et al. 1988. Lipases in the preparation of β -blockers. *TIBTECH* 6: 251-256
76. Knehr M, Thomas H, Arand M, Gebel T, Zeller HD, Oesch F. 1993. Isolation and characterization of a cDNA encoding rat liver cytosolic epoxide hydrolase and its functional expression in *Escherichia coli*. *J Biol Chem* 268: 17623-17627
77. Kohne DE. 1970. Evolution of higher-organism DNA. *Q Rev Biophys* 3: 327-375
78. Kroetz DL, Zeldin DC. 2002. Cytochrome P450 pathways of arachidonic acid metabolism. *Curr Opin Lipidol* 13: 273-283
79. Kronenburga NA, de Bont JA. 2001. Effects of detergents on specific activity and enantioselectivity of the epoxide hydrolase from *Rhodotorula glutinis*. *Enzyme Microb Technol* 28: 210-217
80. Kroutil W, Genzel Y, Pietzsch M, Syldatk C, Faber K. 1998. Purification and characterization of a highly selective epoxide hydrolase from *Nocardia* sp. EH1. *J Biotechnol* 61: 143-150
81. Kull F, Ohlson E, Lind B, Haeggstrom JZ. 2001. Saccharomyces cerevisiae leukotriene A4 hydrolase: formation of leukotriene B4 and identification of catalytic residues. *Biochemistry* 40: 12695-12703
82. Labrou NE, Mello LV, Clonis YD. 2001. Functional and structural roles of the glutathione-binding residues in maize (*Zea mays*) glutathione S-transferase I. *Biochem J* 358: 101-110
83. Léon J, Sánchez-Serrano JJ. 1999. Molecular biology of jasmonic acid biosynthesis in plants. *Plant Physiol Biochem* 37: 373-380
84. Li W-H, Graur D. 1991. *Fundamentals of molecular evolution.*: Sinauer Associates
85. Loewenthal R, Sancho J, Fersht AR. 1992. Histidine-aromatic interactions in barnase. Elevation of histidine pKa and contribution to protein stability. *J Mol Biol* 224: 759-770
86. Lu AY, Miwa GT. 1980. Molecular properties and biological functions of microsomal epoxide hydrase. *Annu Rev Pharmacol Toxicol* 20: 513-531
87. Lutje SHJ, Rink R, Kellogg RM, Janssen DB. 1998. Enantioselectivity of a recombinant epoxide hydrolase from *Agrobacterium radiobacter*. *Tetrahedron: Asymmetry* 9: 459-466
88. Madigan MT, Martinko JM, Parker J. 2009. *Brock biology of microorganisms*: Prentice-Hall Inc.

89. Martel P, Baptista A, Petersen SB. 1996. Protein Electrostatics. *Biotechnology Annual Reviews*: 315-372
90. Mechichi T, Stackebrandt E, Gad'on N, Fuchs G. 2002. Phylogenetic and metabolic diversity of bacteria degrading aromatic compounds under denitrifying conditions, and description of *Thauera phenylacetica* sp. nov., *Thauera aminoaromatica* sp. nov., and *Azoarcus buckelii* sp. nov. *Arch Microbiol* 178: 26-35
91. Michener CD, Sokal RR. 1957. A quantitative approach to a problem in classification. *Evolution* 11: 130-162
92. Misawa E, Chan Kwo Chion CK, Archer IV, Woodland MP, Zhou NY, et al. 1998. Characterisation of a catabolic epoxide hydrolase from a *Corynebacterium* sp. *Eur J Biochem* 253: 173-183
93. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Ehrlich H. 1985. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symp Quant Biol* 51: 263-273
94. Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247: 536-540
95. Nagata Y, Miyauchi K, Damborsky J, Manova K, Ansorgova A, Takagi M. 1997. Purification and characterization of a haloalkane dehalogenase of a new substrate class from a gamma-hexachlorocyclohexane-degrading bacterium, *Sphingomonas paucimobilis* UT26. *Appl Environ Microbiol* 63: 3707-3710
96. Nardini M, Ridder IS, Rozeboom HJ, Kalk KH, Rink R, et al. 1999. The x-ray structure of epoxide hydrolase from *Agrobacterium radiobacter* AD1. An enzyme to detoxify harmful epoxides. *J Biol Chem* 274: 14579-14586
97. Needleman SB, Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 48: 443-453
98. Newman JW, Morisseau C, Harris TR, Hammock BD. 2003. The soluble epoxide hydrolase encoded by EPXH2 is a bifunctional enzyme with novel lipid phosphate phosphatase activity. *Proc Natl Acad Sci U S A* 100: 1558-1563
99. Oesch F. 1972. Mammalian epoxide hydrases. Inducible enzymes catalysing the inactivation of carcinogenic and cytotoxic metabolites derived from aromatic olefinic compounds. *Xenobiotica* 3: 305-340
100. Oesch F, Bentley P. 1976. Antibodies against homogeneous epoxide hydratase provide evidence for a single enzyme hydrating styrene oxide and benz(a)pyrene 4,5-oxide. *Nature* 259: 53-55
101. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, et al. 1992. The alpha/beta hydrolase fold. *Protein Eng* 5: 197-211
102. Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM. 1997. CATH--a hierarchic classification of protein domain structures. *Structure* 5: 1093-1108
103. Orru RV, Archelas A, Furstoss R, Faber K. 1999. Epoxide hydrolases and their synthetic applications. *Adv Biochem Eng Biotechnol* 63: 145-167
104. Otyepka M, Damborsky J. 2002. Functionally relevant motions of haloalkane dehalogenases occur in the specificity-modulating cap domains. *Protein Sci* 11: 1206-1217
105. Page RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12: 357-358
106. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, et al. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp Phys Commun* 91: 1-14

107. Pearson WR, Lipman DJ. 1988. Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* 85: 2444-2448
108. Pedragosa-Moreau S, Archelas A, Furstoss R. 1994. Microbiological transformations--XXIX. Enantioselective hydrolysis of epoxides using microorganisms: a mechanistic study. *Bioorg Med Chem* 2: 609-616
109. Pedragosa-Moreau S, Archelas A, Furstoss R. 1995. Epoxydes énantiopurs: obtention par voie chimique ou par voie enzymatique. *Bull Soc Chim Fr* 132: 769-800
110. Pikkemaat MG, Linssen AB, Berendsen HJ, Janssen DB. 2002. Molecular dynamics simulations as a tool for improving protein stability. *Protein Eng* 15: 185-192
111. Pleiss J, Fischer M, Peiker M, Thiele C, Schmid RD. 2000. Lipase engineering database - Understanding and exploiting sequence-structure-function relationships. *Journal of Molecular Catalysis B-Enzymatic* 10: 491-508
112. Pleiss J, Fischer M, Schmid RD. 1998. Anatomy of lipase binding sites: the scissile fatty acid binding site. *Chem Phys Lipids* 93: 67-80
113. Prestwich GD, Wojtasek H, Lentz AJ, Rabinovich JM. 1996. Biochemistry of proteins that bind and metabolize juvenile hormones. *Arch Insect Biochem Physiol* 32: 407-419
114. Pries F, van den Wijngaard AJ, Bos R, Pentenga M, Janssen DB. 1994. The role of spontaneous cap domain mutations in haloalkane dehalogenase specificity and evolution. *J Biol Chem* 269: 17490-17494
115. Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, et al. 1987. A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238: 336-341
116. Rink R, Fennema M, Smids M, Dehmel U, Janssen DB. 1997. Primary structure and catalytic mechanism of the epoxide hydrolase from *Agrobacterium radiobacter* AD1. *J Biol Chem* 272: 14650-14657
117. Rink R, Janssen DB. 1998. Kinetic mechanism of the enantioselective conversion of styrene oxide by epoxide hydrolase from *Agrobacterium radiobacter* AD1. *Biochemistry* 37: 18119-18127
118. Rink R, Kingma J, Lutje Spelberg JH, Janssen DB. 2000. Tyrosine residues serve as proton donor in the catalytic mechanism of epoxide hydrolase from *Agrobacterium radiobacter*. *Biochemistry* 39: 5600-5613
119. Robertson HM. 1993. The mariner transposable element is widespread in insects. *Nature* 362: 241-245
120. Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res* 26: 1628-1635
121. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425
122. Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual, second edition.*: Cold Spring Harbor Laboratory Press
123. Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74: 5463-5467
124. Schanstra JP, Kingma J, Janssen DB. 1996. Specificity and kinetics of haloalkane dehalogenase. *J Biol Chem* 271: 14747-14753
125. Schladt L, Thomas H, Hartmann R, Oesch F. 1988. Human liver cytosolic epoxide hydrolases. *Eur J Biochem* 176: 715-723
126. Schmidt HA, Strimmer K, Vingron M, von Haeseler A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18: 502-504
127. Shuman S. 1994. Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J Biol Chem* 269: 32678-32684

128. Smith TF, Waterman MS. 1981. Identification of common molecular subsequences. *J Mol Biol* 147: 195-197
129. Stapleton A, Beetham JK, Pinot F, Garbarino JE, Rockhold DR, et al. 1994. Cloning and expression of soluble epoxide hydrolase from potato. *Plant J* 6: 251-258
130. Steinreiber A, Faber K. 2001. Microbial epoxide hydrolases for preparative biotransformations. *Curr Opin Biotechnol* 12: 552-558
131. Stoesser G, Baker W, van den Broek A, Garcia-Pastor M, Kanz C, et al. 2003. The EMBL Nucleotide Sequence Database: major new developments. *Nucleic Acids Res* 31: 17-22
132. Summerer S, Hanano A, Utsumi S, Arand M, Schuber F, Blée E. 2002. Stereochemical features of the hydrolysis of 9,10-epoxystearic acid catalysed by plant and mammalian epoxide hydrolases. *Biochem J* 366: 471-480
133. Suzuki Y, Marumo S. 1972. trans-to-cis 2,3-double bond isomerization of epoxyfarnesol and farnesol by fungus. *Tetrahedron Lett* 13: 5101-5104
134. Svawing J, de Bont JA. 1998. Microbial transformation of epoxides. *Enzyme Microb Technol* 22: 19-26
135. Tateno Y, Imanishi T, Miyazaki S, Fukami-Kobayashi K, Saitou N, et al. 2002. DNA Data Bank of Japan (DDBJ) for genome scale research in life science. *Nucleic Acids Res* 30: 27-30
136. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876-4882
137. Tzeng HF, Laughlin LT, Armstrong RN. 1998. Semifunctional site-specific mutants affecting the hydrolytic half-reaction of microsomal epoxide hydrolase. *Biochemistry* 37: 2905-2911
138. Van den Wijngaard AJ, Janssen DB, Witholt B. 1989. Degradation of epichlorohydrin and haloaldehydes by bacterial cultures isolated from freshwater sediment. *J Gen Microbiol* 135: 2199-2208
139. VanHook Harris S, Marin Thompson D, Linderman RJ, Tomalski MD, Roe RM. 1999. Cloning and expression of a novel juvenile hormone-metabolizing epoxide hydrolase during larval-pupal metamorphosis of the cabbage looper, *Trichoplusia ni*. *Insect Mol Biol* 8: 85-96
140. Vieira J, Messing J. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19: 259-268
141. Visser H, de Bont JA, Verdoes JC. 1999. Isolation and characterization of the epoxide hydrolase-encoding gene from *Xanthophyllomyces dendrorhous*. *Appl Environ Microbiol* 65: 5459-5463
142. Visser H, Vreugdenhil S, de Bont JA, Verdoes JC. 2000. Cloning and characterization of an epoxide hydrolase-encoding gene from *Rhodotorula glutinis*. *Appl Microbiol Biotechnol* 53: 415-419
143. Wehner R, Gehring W. 1995. *Zoologie*: Georg Thieme Verlag
144. Wisniewski JR, Rudnicka M, Kochman M. 1986. Juvenile hormone binding protein from the epidermis of *Galleria mellonella*. *Insect Biochem* 16: 843-9
145. Woese CR. 1987. Bacterial evolution. *Microbiol Rev* 51: 221-271
146. Woese CR, Fox GE. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* 74: 5088-5090
147. Wojtasek H, Prestwich GD. 1996. An insect juvenile hormone-specific epoxide hydrolase is related to vertebrate microsomal epoxide hydrolases. *Biochem Biophys Res Commun* 220: 323-9

-
148. Wu CH, Huang H, Arminski L, Castro-Alvear J, Chen Y, et al. 2002. The Protein Information Resource: an integrated public resource of functional annotation of proteins. *Nucleic Acids Res* 30: 35-37
 149. Zocher F, Enzelberger MM, Bornscheuer UT, Hauer B, Schmid RD. 1999. A colorimetric assay suitable for screening epoxide hydrolase activity. *Analytica Chimica Acta*: 345-351
 150. Zocher F, Enzelberger MM, Bornscheuer UT, Hauer B, Wohlleben W, Schmid RD. 2000. Epoxide hydrolase activity of *Streptomyces* strains. *J Biotechnol* 77: 287-292
 151. Zou J, Hallberg BM, Bergfors T, Oesch F, Arand M, et al. 2000. Structure of *Aspergillus niger* epoxide hydrolase at 1.8 Å resolution: implications for the structure and function of the mammalian microsomal class of epoxide hydrolases. *Structure Fold Des* 8: 111-122
 152. Zuckerkandl E, Pauling L. 1965. Molecules as documents of evolutionary history. *J Theor Biol* 8: 357-366

Sandra Barth
Dipl. Biologin (t.o.)
Seerosenstrasse 1
70563 Stuttgart

CURRICULUM VITAE

01.09.1972
Neuenbürg (Württ.)

ledig
deutsch

1979-1983
1983-1992

August 1992-
Juli 1994

WS 1994-
WS 1996

SS 1997-
SS 2000
19. Mai 2000

Dezember 1999-
Mai 2000

Seit Juni 2000

Geburtsdatum
Geburtsort

Familienstand
Nationalität

Schulbildung

Grundschule Feldrennach
Gymnasium Neuenbürg

Berufsausbildung

Chemisches Institut Dr. Flad
Abschluss: Chemisch technischer Assistent

Hochschulstudium

Studium des Chemieingenieurwesens
an der Universität Karlsruhe
Abschluss: Vordiplom

Studium der Technischen Biologie
an der Universität Stuttgart
Abschluss: Diplom

Diplomarbeit bei Prof. Dr. Rolf D. Schmid,
Institut für Technische Biochemie
Universität Stuttgart:

Thema:

Untersuchungen zur enzymatischen
Darstellung von optisch reinem L-
Phosphinothricin ausgehend von D,L-N-
Acetyl-Phosphinothricin

Promotion bei Prof. Dr. Rolf D. Schmid,
Institut für Technische Biochemie
Universität Stuttgart:

Thema:

Systematische Untersuchungen von
Epoxidhydrolasen: Erstellung einer familien-
spezifischen Datenbank, Entwicklung einer
strukturbezogenen Klassifizierung und
Amplifikation unbekannter Epoxidhydrolasen.