# Two novel glycyl radical decarboxylase systems from *Clostridium scatologenes* and *Tannerella forsythensis*

Dissertation

zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Biologie
der
Philipps-Universität Marburg

vorgelegt von
**Lihua Yu**
aus der V. R. China

Marburg/Lahn, Germany 2006

Die Untersuchungen zur vorliegenden Arbeit wurden von Oktober 2002 bis Oktober 2005 im Laboratorium für Mikrobiologie, Fachbereich Biologie, der Philipps Universität Marburg unter der Leitung von PD. Dr. T. Selmer durchgeführt.

Vom Fachbereich Biologie
der Philipps-Universität Marburg
als Dissertation am Jan. 16, 2006 angenommen.
Erstgutachter: PD Dr. T. Selmer
Zweitgutachter: Prof. Dr. W. Buckel
Tag der mündlichen Prüfung am Feb. 17, 2006.

# Contents

**Abbreviations**

**Zusammenfasung**

**Summary**

**1 Introduction ..... 5**

1.1 Organisms ..... 6

1.2 Glycyl radical enzymes ..... 7

1.2.1 Common features of glycyl radical enzymes ..... 9

1.3 Generation of the glycyl radical ..... 12

1.3.1 4-Hydroxyphenylacetate decarboxylase (Hpd) ..... 16

**2 Materials and Methods ..... 19**

2.1 Materials ..... 19

2.1.1 Chemicals and reagents ..... 19

2.1.2 Instruments, gases and columns ..... 19

2.1.3 Bacterial strains and plasmids ..... 20

2.1.4 Media and plates ..... 22

2.1.5 Oligonucleotides ..... 23

2.2 Microbiological Methods ..... 25

2.2.1 Growth of *E. coli* ..... 25

2.2.2 Storage of bacterial strains ..... 25

2.3 Molecular biology ..... 25

2.3.1 Isolation of DNA from *C. scatologenes* ..... 25

2.3.2 Determination of DNA concentration and purity ..... 26

2.3.3 Plasmid DNA isolation ..... 26

2.3.4 DNA gel eletrophoresis ····· 26

2.3.5 Purification of DNA from agarose gel ····· 27

2.3.6 DNA restriction and ligation ····· 27

2.3.7 Phenol-chloroform extraction ····· 28

2.3.8 Preparation of electrocompetent cell ····· 28

*2.3.9* *E. coli* transformation by electroporation ····· 28

2.3.10 Transformation of chemically competent cells ····· 29

2.3.11 PCR ····· 29

2.3.12 Southern blotting ····· 33

2.3.13 DNA sequencing ····· 35

2.3.14 Construction of the plasmids for sequencing and protein expression ····· 35

2.4 Biochemical methods ····· 41

2.4.1 Gene expression and protein purification ····· 41

2.4.2 Determination of protein concentrations ····· 42

2.4.3 Sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) ····· 43

2.4.4 Determination of relative molecular masses of the native enzymes ····· 44

2.4.5 Colorimetric determination of non-heme iron with Ferene ····· 45

2.4.6 Determination of acid labile sulfide ····· 45

2.4.7 Enzyme activity assays ····· 47

2.4.8 UV-visible measurements ····· 48

2.4.9 Reverse-phase HPLC separation of the aromatic compounds ····· 48

2.4.10 EPR Spectroscopy ····· 49

**3 Results ····· 50**

3.1 Sequencing of the *csd*-locus of *Clostridium scatologenes* ····· 50

3.2 Sequence analysis of the individual glycyl radical decarboxylase systems ····· 61

3.3 Cloning and expression of the *csd* and *tfd* genes ..... 65

3.4 Protein purification ..... 67

3.5 Physical and chemical characterisation of the recombinant proteins ..... 71

3.6 Functional characterisation of the individual systems *in vitro* ..... 74

3.6.1 Functionality and substrate specificity of the recombinant decarboxylases ..... 74

3.7 Kinetic studies ..... 77

3.7.1 Transient activation ..... 77

3.7.2 Different A/BC ratio activation ..... 78

3.7.3 Alternative chemical reductants for the AE ..... 80

3.7.4 Michaelis-Menten kinetics for CsdBC. ..... 80

3.8 EPR results ..... 83

3.8.1 Detection of the glycyl radical in HpdBC, CsdBC and TfdBC ..... 83

3.8.2 Initial characterisation of the metal centers in Hpd and Csd ..... 88

**4 Disscussion ..... 91**

4.1 Genetic arrangements of the decarboxylase loci ..... 91

4.2 Recombinant production of the GRE decarboxylases ..... 92

4.3 The biochemical characteristics of GRE decarboxylases ..... 93

4.4 Conclusions ..... 97

4.5 Outlook ..... 100

**5 References ..... 103**

## Abbreviations

| | |
|---|---|
| AHT | Anhydrotetracycline |
| Bss | Benzylsuccinate synthase |
| Csd | *Clostridium scatologenes* decarboxylase |
| GRE | Glycyl radical enzyme |
| GD | Glycerol Dehydratase |
| IAA | Indoleacetate |
| APA | 4-aminophenylacetate |
| 3,4-DHPA | 3,4-dihydroxyphenyactate |
| 4-HPA | 4-hydroxyphenylacetate |
| Hpd | 4-hydroxyphenylacetate decarboxylase from *C.difficile* |
| HPLC | High Performance Liquid Chromatography |
| IPTG | Isopropyl thio-β-D-galactoside |
| LB | Luria-Bertani medium |
| LBG medium | LB medium supplemented with 0.2% Glucose |
| Nrd | Ribonucleotide reductase |
| Pfl | Pyruvate Formate-Lyase |
| SAM | S-adenosylmethionine |
| Tfd | *Tannerella forsythensis* decarboxylase |

## Zusammenfasung

Die chemisch schwierige Decarboxylierung von 4-Hydroxyphenylacetat zu *p*-Kresol wird durch das Enzym 4-Hydroxyphenylacetat-Decarboxylase (4-Hpd) katalysiert. Dieses Enzym wurde gereinigt und als Prototyp einer neuen Gruppe innerhalb der Glycylradikalfamilie (GREs) charakterisiert. Frühere Studien haben gezeigt, dass dieses System in *C. difficile* Eigenschaften aufweist, die es von den gut untersuchten Systemen Pyruvat Format Lyase (Pfl) und anaerobe Ribonucleotid Reduktase (Nrd) unterscheiden. In dieser Arbeit wurden ähnliche Gene aus *Clostridium scatologenes* (Csd) und *Tannerella forsythensis* (Tfd) kloniert und in *Escherichia coli* exprimiert. Die rekombinanten Enzyme wurden gereinigt und vorläufig charakterisiert.

Die rekombinanten Decarboxylasen konnten als Hetereokotamere (HpdBC und CsdBC) oder Heterotetramerer (TfdBC) gereinigt werden, waren aus großen (B) und kleinen (C) Untereinheiten in äquimolaren Mengen zusammen gesetzt, und enthielten im Gegensatz zu Pfl und Nrd vier Eisen- und vier Schwefel-Atome pro Heterodimer. Während das Csd System 4-Hydroxyphenylacetat-Decarboxylase Aktivität zeigte und sowohl von CsdA als auch von HpdA aktiviert wurde, war das Tfd System unter allen getesteten Versuchsbedingungen inaktiv, zeigte aber eine teilweise Aktivierung zur Glycyl-Radikal-Form.

Wurden die große Untereinheiten der einzelnen Decarboxylasen genetisch mit den kleinen Untereinheiten kombiniert, konnten in einigen Fällen lösliche Proteine gereinigt werden. Auch hier betrug das molare Verhältnis der beiden Untereinheiten 1:1 und es konnten Eisen und Schwefel nachgewiesen werden. Allerdings waren die nativen Komplexe dieser Proteine deutlich kleiner und konnten nicht in die Glycyl-Radikal-Form überführt werden.

Die rekombinanten Aktivatoren (CsdA bzw. TfdA) waren Monomere und enthielten 7-8 Eisen- und 6-7 Schwefel-Atome pro Monomer, das auf einen zweiten, zusätzlich zu dem aus Pfl und Nrd bekannten, [4Fe-4S]-Kluster schließen ließ. Der im EPR detektierte, katalytisch entscheidende $[4Fe\text{-}4S]^{+}$ Kluster wurde in CsdA und HpdA nachgewiesen, unterschied sich aber von den Signalen des Pfl- Activator sowie des Nrd- Activator; im Gegensatz zu letzteren beeinflusste die Substratbindung das EPR-Signal nicht signifikant.

Der Aktivierungsprozess von CsdBC sowie HpdBC mit den jeweiligen Aktivatoren war transient; einem steilen Anstieg an spezifischer Aktivität in etwa 10 Minuten folgte ein langsamerer Inaktivierungsprozess mit einer Halbwertszeit von etwa 30 Minuten. Das hierbei das Glycylradikal verschwindet, konnte durch sauerstoff-induzierte Spaltung den aktiven Decarboxylase mittels SDS-PAGE sowie in EPR Messungen gezeigt werden. Ob diese

Inaktivierung durch ein Elektron aus dem zusätzlichen [4Fe-4S]-Kluster des Aktivators verursacht wird oder durch den [4Fe4S]-Kluster der Decarboxylasen vermittelt wird, bleibt Gegenstand aktueller Untersuchungen.

## Summary

The chemically difficult decarboxylation of 4-hydroxyphenylacetate to *p*-cresol is catalysed by the enzyme 4-hydroxyphenylacetate decarboxylase (Hpd). The enzyme has been purified and characterised as the prototype of a novel class of glycyl radical enzymes (GREs). Previous studies have shown that the Hpd system from *Clostridium difficile* shows distinct properties, which distinguishes this novel GRE system from the well-characterised pyruvate formate lyase (Pfl) and anaerobic ribonucleotide reductase (Nrd) system. In this work, the similar genes from *Clostridium scatologenes* (Csd) and from *Tannerella forsythensis* (Tfd) were cloned and expressed in *Escherichia coli*. The individual recombinant proteins were produced in *Escherichia coli*, purified and initially characterized.

The recombinant proteins of CsdBC (hetero-octamer) and TfdBC (hetero-tetramer) were composed of large and small subunits in a 1 to 1 molar ratio and contained 4 irons and 4 sulfurs per heterodimer as that of the HpdBC (hetero-octamer), which distinguishes these systems from Pfl and Nrd. While the Csd system exhibited 4-hydroxyphenylacetate decarboxylase activity like Hpd and was activated by either HpdA or CsdA, the Tfd system was inactive under any conditions tested, but the glycyl radical formation was observed by EPR.

When the glycyl radical subunits of the individual decarboxylases were genetically combined with the small subunits of the other systems, soluble recombinant proteins were formed and purified for some of these combinations. For all combinations yielding soluble enzymes, the molar ratio of the glycyl radical and small subunits was 1:1 and the presence of an iron-sulfur centre per hetero-dimer was evident, while the oligomeric state differed from the wild-type complexes, exhibiting a lower order of complexity. All the purified hybrid decarboxylases were inactive under current assay conditions.

The recombinant activating enzymes (CsdA and TfdA) were monomers according to size exclusion chromatography and contained 7-8 mol iron and 6-7 mol acid labile sulfur per mol of enzyme, indicating at least one additional metal centre, as compared to the Pfl and Nrd activators. The catalytically essential $[4Fe\text{-}4S]^{+}$ cluster in CsdA and HpdA was detected by EPR, which was different to that of Pfl-AE or Nrd-AE, there was no significant signal changes after addition of SAM.

The activation of CsdBC or HpdBC by its cognate activating enzyme was transient, yielding maximum specific activities within 10 min followed by a slow inactivation with $t_{1/2} \sim 30$ min, which was accompanied by a quenching of the glycyl radical. This radical quenching process

was monitored by using the oxygen-induced cleavage of decarboxylase on SDS-PAGE, and also by EPR. Combining these results, it was proposed that the second iron-sulfur centre (I-cluster, which located within a ~ 60 amino acids long insert between the SAM cluster and the first glycin-rich motif) could be responsible for the ‘inactivase’ activity of HpdA or CsdA. Alternatively, the unique metal centre in the GRE decarboxylase itself may provide the radical quenching properties of the systems.

# 1 Introduction

The production of various phenols or indols from aromatic amino acids is known for a long time [1]. Scatole (3-methylindole) and *p*-cresol (4-methylphenol) are common bacterial metabolic end products in avian and mammalian feces [2-9]. These compounds are formed by fermentation of the proteinogenic amino acids tyrosine and trypthophan by direct decarboxylation of the corresponding arylacetates, respectively [1].

The cresol formation has been studied most intensively in *Clostridium difficile* [10, 11] and has been suggested to provide an important virulence factor for the human pathogen. *p*-Cresol is formed by the enzyme 4 - hydroxyphenylacetate decarboxylase (Hpd, EC 4.1.1.82).

Though several hundreds of glycyl radical enzyme (GRE) systems were found in the genomes of strict and facultative anaerobes, only one putative arylacetate decarboxylase gene locus was identified in the unfinished genome of *Tannerella forsythensis*. However, no cresol formation has been reported for this organism [12]. In turn, *Clostridium scatologenes* was known to produce *p*-cresol [1] and, moreover, has been named according to its capability to form 3-methylindole (skatole) from indole-3-acetate. Preliminary data strongly suggested that indole-3-acetate decarboxylase is a second arylacetate decarboxylase in this organism [11].

.

## Organisms

Most members of genus *Clostridium* are Gram-positive, spore-forming rods. These strict anoxic, motile bacteria are widely distributed in anoxic niches and play an important role in the anaerobic degradation of organic matter. These organisms exhibit a versatile capability to degradate organic molecules and play an imortant role in the spoilage of food. Some of the clostridia secrete powerful exotoxins that are responsible for fatal diseases including tetanus (*C. tetani*), botulism (*C. botulinum*), and gas gangrene (*C. perfringens* and others).

*Clostridium difficile* causes gastrointestinal infections in humans [13-17] and is one of the most important causes of diarrhetic complications in long term antibiotic treatments. The organism grows on various media and has been shown one of the most versatile clostridia concerning its ability to ferment amino acids [18]. The pathogenesis of the organism is mediated by two large toxins, which primarily act as glucosyl-transferases. The glucosylation of *ras* and *rho* proteins causes cytoskeleton disassembly and is responsible for the cytotoxic and tissue damaging effects. The fermentative formation of *p*-cresol may account for an ongoing suppression of the gut microbiota, which has been previously altered by antibiotic treatment, and may therefore account for an important virulence factor in *C. difficile*-associated disease.

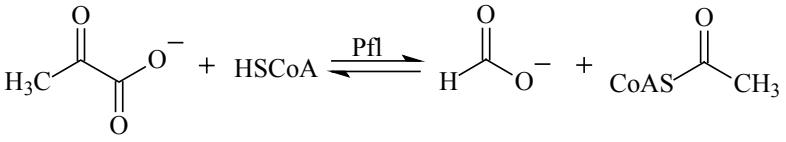
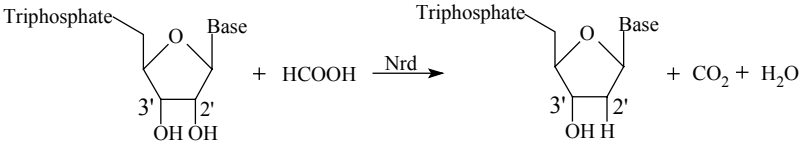
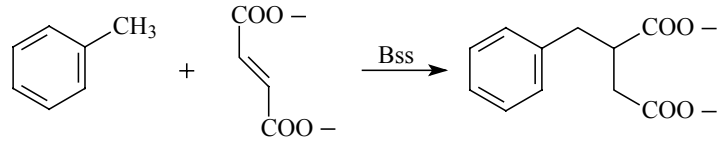
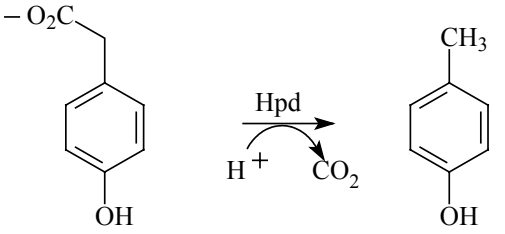
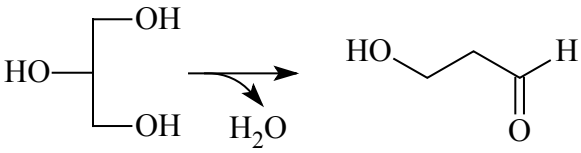
*Clostridium scatologenes* was first described by Weinberg and Ginsbourg in 1927 [19]. The organism has been isolated from soil [20], contaminated food, and feces of infants undersized at birth. The organism is also capable to produce *p*-cresol and has been named according to its ability to form scatole from tryptophane.

*Tannerella forsythensis* forms Gram-negative, non-motile fusiform cells with pointed ends. The organism grows very slow anaerobically on blood agar plates, poor in broth media and requires *N*-acetylmuramic acid as growth factor [21]. Its principal habitats are periodontal pockets that show progressive tissue loss. The organism has been considered as one of the important species associated with periodontal diseases.

## Glycyl radical enzymes

During the last decades, it has become obvious that protein-based radicals are involved in several biologically important reactions. Many of these enzymes are activated by one-electron oxidation of protein side chains such as tyrosine, tryptophan or cysteine residues. However, in particular the glycyl radical enzymes (GREs) seem to be more prevalent in nature than previous anticipated, and the recent genomic analyses have revealed that this kind of enzyme may be widespread amongst obligate and facultative anaerobic microbes, including Bacteria, Archaea and Eukarya [22]. They are involved in a variety of reactions (Table 1.1), for example in the anoxic degradation of pyruvate, the activation of recalcitrant molecules like toluene, anaerobic ribonucleotide reduction, vitamine $B_{12}$-independent glycerol dehydratation and 4-hydroxyphenylacetate decarboxylation. Moreover, a number of novel systems have been identified in the genomes of anaerobes, whose functions remain to be established.

**Table 1.1 Individual reactions catalysed by GREs.**

| Enzyme/ Organism | References | Reaction catalysed |
|---|---|---|
| Pfl/ *E. coli* | [23] |  |
| Nrd/ *E. coli* | [24] |  |
| Bss/ *T. aromatica* | [25] |  |
| Hpd/ *C. difficile* | [11] |  |
| GD/ *C. butyricum* | [26] |  |

**Common features of glycyl radical enzymes**

When synthesised, the GREs are catalytically inactive and must undergo activation through the process of hydrogen atom abstraction from specific glycine residue located close to the C-terminus of the polypeptide chain. This process is catalysed by an iron-sulfur containing activating enzyme (AE) that utilises S-adenosylmethionine (SAM) and a one-electron donor, usually dihydroflavodoxin, for the generation of a 5'-deoxyadenosyl radical. A unique [4Fe-4S] $^{+1/+2}$ cluster in the AE is directly involved in this reaction. The transiently formed 5'-deoxyadenosyl radical is proposed to abstract subsequently a hydrogen atom from a strictly conservered glycine residue in the GRE.

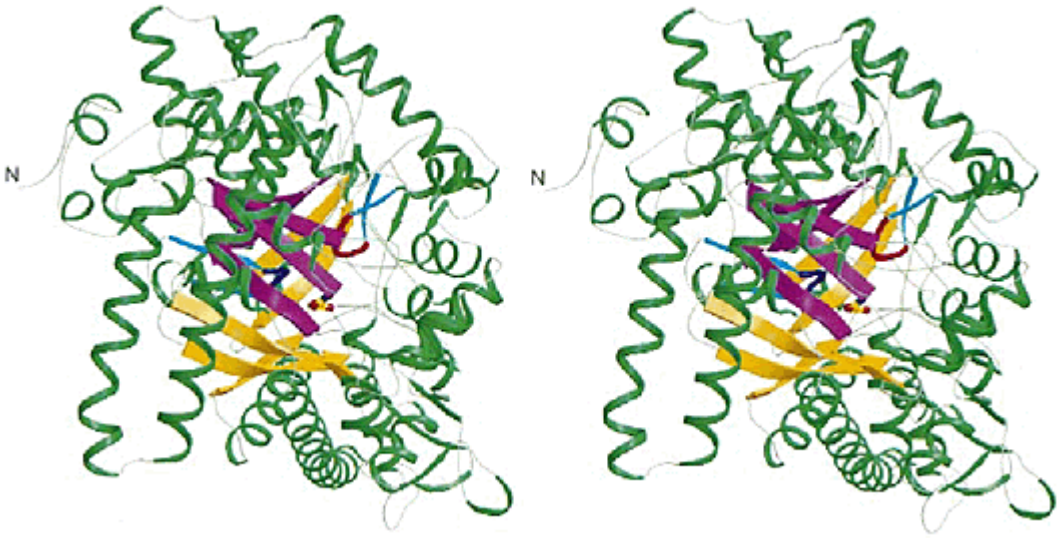
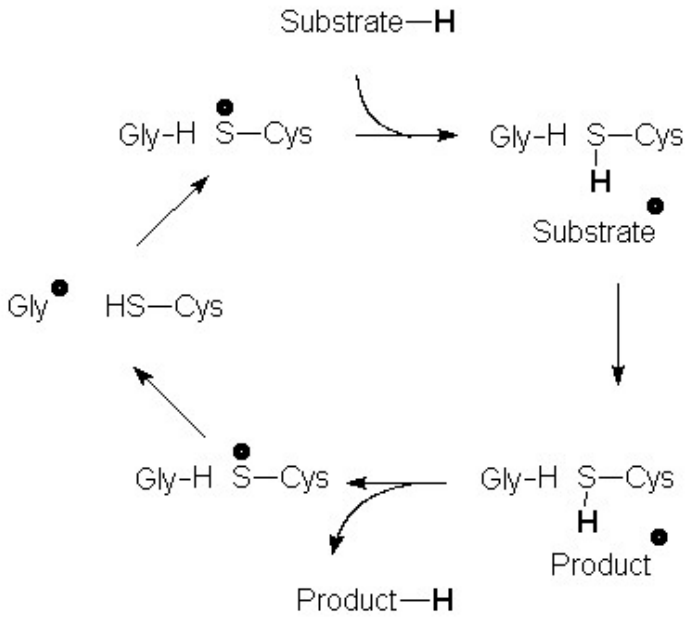



**Figure 1.1 Ribbon diagram of one monomer of Pfl in stereo representation.** The barrel is assembled in an antiparallel manner from two parallel five-stranded β-sheets (shown in purple and orange). The Gly 734 loop (shown in red) opposes the Cys 418−Cys 419 loop (shown in blue) in the center of the barrel. The active site is occupied by the substrate analog oxamate [27].

All GREs are in general large enzymes of about 900 amino acids that carry the strictly conserved glycine in a highly indicative sequence, the glycyl radical finger print motif RVxG [FWY]$x_{6-8}$[FL]$x_4$Q$x_2$[IV]$x_2$R. DNA sequence analysis revealed that the genes encoding the GREs and their cognate AEs are located adjacently on the chromosomes, but do not essentially form transcriptional units. From the crystal structures of Pfl, Nrd and Gdh, it is apparent that these enzymes form homodimers and that the individual polypeptides exhibit a

very similar three-dimensional fold. The unique GRE fold consists of a core 10-stranded β-barrel motif that is assembled in an antiparallel manner from two parallel 5-stranded β-sheets. This β-barrel core is entirely surrounded by α-helices (figure 1.1). All three GRE structures share the presence of two flexible loops protruding into the central core of β-sheets [27-31]. While one of these loops carries the glycyl radical site, the second loop carries the active site cysteinyl residue.

The reaction cycle of all known GREs is thought to be initiated by transfering the radical to the thiol group of the conserved cysteine located in the middle of the glycyl radical subunit amino acid sequence, which is in close neighbourhood of the glycyl radical site in the tertiary structure. The resulting highly reactive thiyl radical then initiates a radical reaction path with the substrates (scheme 1.1) [27, 28, 32, 33].




**Scheme 1.1 General scheme of the GRE mechanistic cycle** ([33] Selmer, T. *et al*. 2005, for detaila see the text)**.**

The active site thiyl radical then abstracts a hydrogen atom from the substrate, yielding an enzyme-bound substrate radical, which undergoes the reaction to yield an intermediate product radical. Finally, the product radical re-abstracts the hydrogen atom from the thiol-group of the enzyme, rendering it back to the active, radical-containing state. The only deviation from this general scheme is known for Pfl, where the initial radical transfer is

relayed on the thiol of another conserved cysteine that forms a covalent bond with the substrate to activate it to a radical intermediate.

The characteristic glycyl radical EPR signal has a remarkably constant appearance (a doublet with a hyperfine coupling of 1.4-1.5 mT centred around $g_{iso}$ = 2.0035 at X-band) in Pfl [23, 34], Nrd [35, 36], Hpd (this work), Bss [37] [38] and methylpentylsuccinate synthases [39]. Combined with the sensitivity of detection (down to concentrations of 0.2 μM) the fingerprint EPR signal allows convenient studies on the kinetics and extent of activation [40] as well as identification of GREs in whole cells and cell-free extracts [38, 39]. The glycyl radical is stabilized by delocalisation of the free electron over the adjacent peptide bonds in the protein backbone (captodative effect). In the presence of oxygen, the glycyl radical becomes highly unstable, and the enzyme is irreversibly inactivated by cleavage of the polypeptide chain at the site of the radical [23].

Temperature and microwave power dependence of the EPR signal pointed out that glycyl radicals occur in a magnetically isolated form. In addition, the limited *g*-anisotropy in high frequency EPR spectroscopy [41] provided evidence for a similar local environment. A rhombic *g*-tensor was ressolved for Pfl, Nrd and Bss, which was in agreement with EPR studies on irradiated *N*-acetylglycine crystals. However, certain differences have been observed. Whereas the 'residual' hydrogen atom at the α-carbon atom of the glycyl radical can be exchanged in $D_2O$ in Pfl [34] and Bss [37, 38], but exchange in Nrd was not detected [35].

Cleavage of SAM [42] and oxidation of the $[4Fe\text{-}4S]^+$ cubane [40] occurs stoichiometrically with formation of the glycyl radical. Although no X-ray crystallographic structural information is available yet for GRE-AEs, the direct coordination of SAM to the [4Fe-4S] cubane has been elucidated by Mössbauer [43] and ENDOR spectroscopies [44, 45]. The mode of binding is in agreement with X-ray crystallographic data on SAM-binding in coproporphyrinogen III decarboxylase [46] and biotin synthase [47].

While the inactive precoursor of GREs are entirely stable under air, exposure of the active GREs to dioxygen results in their rapid and irreversible inactivation, which is accompanied with specific scission of the polypeptide chain in the position of the glycyl radical, yielding a characteristic α, α' profile on SDS polyacrylamide gels [23, 25, 48, 49]. The exposure to oxygen results in loss of the glycyl radical signal in EPR [23, 50]. Although the exact mechanism of oxygen-induced cleavage of GREs is not fully understood, the cleavage reaction appears to be central to all GREs and certainly accounts for the restriction of these enzymes to strictly anoxic environments.

**Generation of the glycyl radical**

All GREs must undergo a post-tranlational activation to generate the radical form. The introduction of the radical is achieved by an iron-sulfur protein, which is predominantly specific for individual GREs [51-54]. These activating enzymes (AEs) are members of the large SAM radical enzyme superfamily [55] and are characterised by a unique C-$x_3$-C-$x_2$-C iron-sulfur cluster binding motif and the dependence on SAM as co-substrate. The diverse range of functions in this superfamily is remarkable and goes far beyond the GRE-AEs.

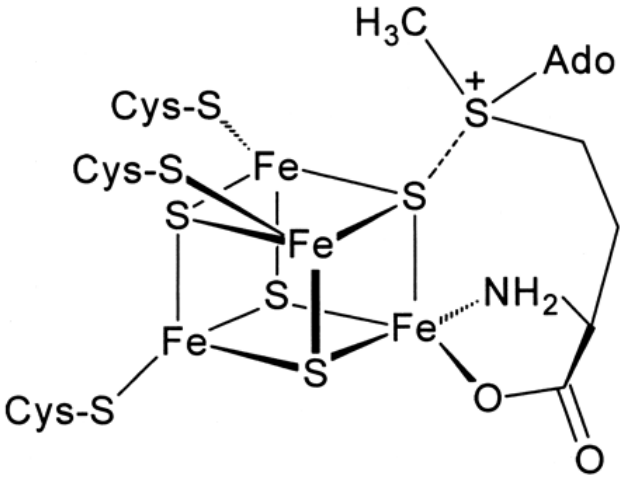
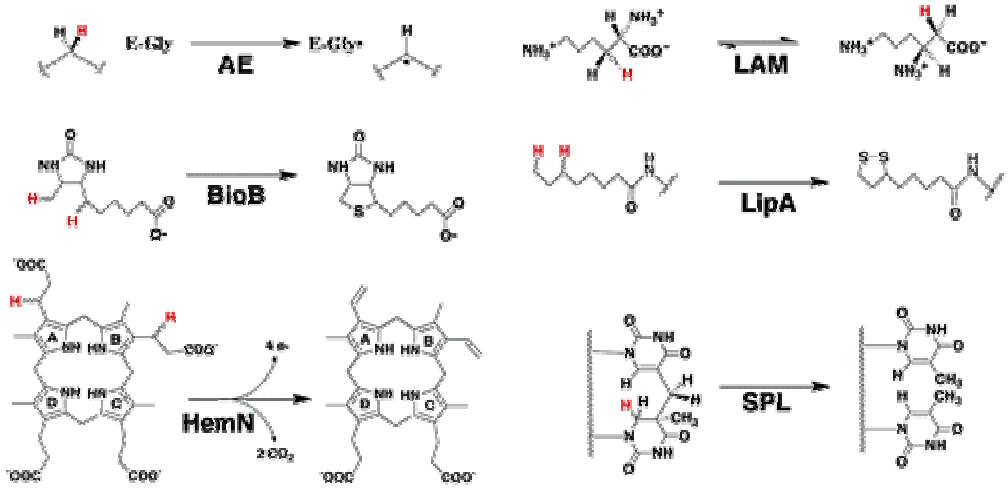



**Figure 1.2 Proposed mode of interaction of SAM with the [4Fe-4S] cluster of PFL-AE.** The model is based on electron-nuclear double resonance and Mössbauer studies ([44, 45, 56]).

In spite of the various reactions carried out by SAM radical enzymes, mechanistic details of Fe-S cluster-mediated SAM cleavage are about to emerge from spectroscopic studies that aim to understand the nature of the interaction of SAM with the $[4Fe\text{-}4S]^{2+/1+}$ cluster [44, 45, 56, 57]. These experiments suggested that there is a close interaction of SAM with the $[4Fe\text{-}4S]^{2+/1+}$ clusters in lysine-2, 3-aminomutase, Pfl-AE, biotin synthase and others. A non-cysteinyl-ligated unique Fe site in the $[4Fe\text{-}4S]^{2+}$ cluster of PFL-AE has been identified by Mössbauer spectroscopy [56]. Moreover, the Mössbauer parameters of this unique Fe can be dramatically affected by the addition of SAM to PFL-AE, suggesting that SAM must be interacting with the unique Fe [56]. Furthermore, electron-nuclear double-resonance (ENDOR) spectroscopy has shown that SAM coordinates to the unique Fe in the $[4Fe\text{-}4S]^{2+/1+}$ cluster [44, 45]. These studies led to a model in which the amino acid moiety of SAM coordinates the unique Fe, whereas the sulfonium ion interacts with one of the $\mu_3$-bridging sulfides of the [4Fe-4S]

cluster (Fig. 1.2). Similarly, a combination of EPR, resonance Raman, and Mössbauer spectroscopies has provided evidence for the interaction of SAM with the unique Fe site of the $[4Fe\text{-}4S]^{2+}$ cluster in BioB [58].

Reversible production of the 5'-deoxyadenosyl radical was formerly regarded as a unique property of adenosylcobalamin; however, evidence was obtained indicating that SAM could also serve as a source of this radical [55, 59]. The role of SAM as a free radical initiator has been suggested to play a central role in GRE-AEs, lysine 2,3-aminomutase, biotin and lipoate synthase, MoaA, HemN and many others. This radical is thought to propagate catalysis by hydrogen abstraction from either substrate molecules or protein partners (scheme 1.2).




**Scheme 1.2 Reactions catalysed by representative members of the radical-SAM superfamily.** AE, activating enzymes; LAM, lysine 2,3-aminomutase; BioB, biotin synthase; LipA, lipoate synthase; HemN, oxygen-independent coproporphyrinogen III synthase; SPL, spore photoproduct lyase (cited from [60] Walsby, C. J. *et al.* (2005)).

All radical-SAM enzymes characterized to date contain iron-sulfur clusters, although the detailed properties of the clusters vary from enzyme to enzyme. As outlined above, PFL-AE was isolated with a $[3Fe\text{-}4S]^{+}$ cluster as its primary cluster form [61, 62], and this cluster is converted to the $[4Fe\text{-}4S]^{+}$ cluster upon reduction. Lysine 2,3-aminomutase exhibits similar (though not identical) cluster chemistry, with intact $[4Fe\text{-}4S]^{2+/+}$ clusters present in the isolated protein and $[3Fe\text{-}4S]^{1+}$ clusters observed upon oxidation [63]. Biotin synthase, however, contains $[2Fe\text{-}2S]^{2+}$ clusters in the isolated protein and $[4Fe\text{-}4S]^{+}$ clusters upon

reduction [64, 65] and has not been shown to contain [3Fe-4S] $^{+}$ clusters under any conditions. More recently, it has been found that, under appropriate reconstitution conditions, biotin synthase contains both [2Fe-2S] $^{2+}$ and [4Fe-4S] $^{2+}$ clusters [66-69]. Likewise, only [2Fe-2S] and [4Fe-4S] clusters have been observed in lipoate synthase [70-72], and only [4Fe-4S] clusters have been identified in HemN [73]. These oxygen and reductant dependent cluster interconversions initially hampered the characterisation of these enzymes and are just about to be overcome recently. But this is not the case in Hpd-AE, where only [4Fe-4S] clusters have been identified.

A hyperthetical mechanism for the Fe-S/AdoMet-dependent formation of the glycyl radical must consider the notion that the S-C bond dissociation energy in AdoMet is more than 252kJ/mol and homolytic cleavage would reqiure labilization of that bond by a one-electron reduction. This is in agreement with the need for a reducing agent, either an enzymatical system such as NADPH: flavodoxin-reductase: flavodoxin *in vivo* or a chemical such as dithionite or photoreduced deazaflavin *in vitro* [74-76].

It has been postulated that the reduction of AdoMet is catalysed by the characteristic [4Fe-4S]$^{+}$ cubane cluster only in the presence of the dedicated substrate of the 5'-deoxyadensyl radical-mediated reaction. The reduced cubane might have the potential to inject one electron into AdoMet bound in its close proximity. The resulting unstable sulfuranyl radical breaks down into methionine and the 5'-deoxyadenosyl radical, 5'dA°, which subsequently abstracts a hydrogen atom (that has been shown to be stereospecific abstracted from the *proS* position in Pfl) of a specific glycine residue in the polypeptide chain in GREs [77] [78]. The reduced [4Fe-4S] $^{+}$ cluster coordinates SAM and this complex is stable in the absence of the GRE. Interestingly, the existence of a tight complex between AdoMet and the AE was demonstrated by a filter-binding assay with Nrd [74] [Fig. 1.4].

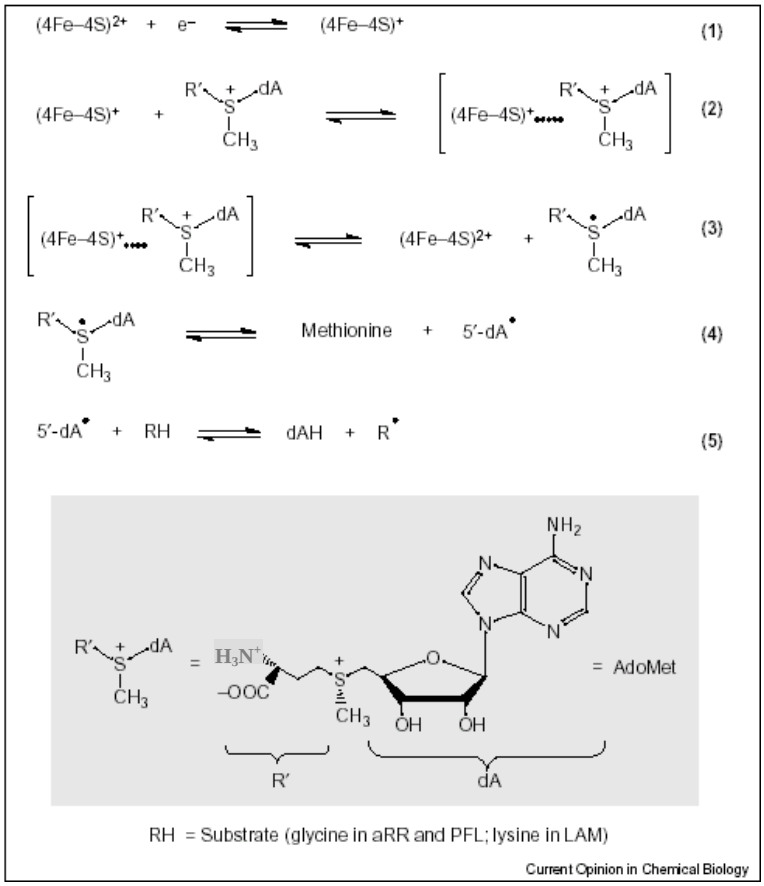



**Figure 1.4 A hyperthetical mechanism for the Fe-S/AdoMet-dependent formation of free radicals.** The structure of the AdoMet is shown in below in the panel [74].

The SAM cluster anchors the methionine of SAM, whereas the adenosyl group is bound by a glycine-rich SGG-sequence motif [55]. This motif is located directly adjacent to the cysteines coordinating the SAM cluster in Pfl-AE and Nrd-AE, whereas other GRE-AEs contain a 31 to 64 amino acid insertion between these functional motifs (Figure 3.13). Within this insert, 4 to 8 cysteinyl residues are located, which form up to two ferredoxin-like $Cx_{2\text{-}4}Cx_2Cx_{7\text{-}34}C$ motifs. This might allow the binding of one or two additional FeS clusters as recently found in recombinant Hpd-AE [12].

**4-Hydroxyphenylacetate decarboxylase (Hpd)**

The chemically difficult decarboxylation of 4-hydroxyphenylacetate (4-HPA) to *p*-cresol (Figure 1.7) is catalysed by the enzyme 4-hydroxyphenylacetate decarboxylase (Hpd, E.C. 4.1.1.18) [10]. The HPA decarboxylase from *C. difficile* is the first arylacetate decarboxylase purified and characterized in detail, and is described as the prototype of a novel class of GRE [11]. Consistent with other GREs, the protein was irreversibly inactivated by molecular oxygen [11], which has been recognised before it was identified as a member of the glycyl radical family [10, 79].

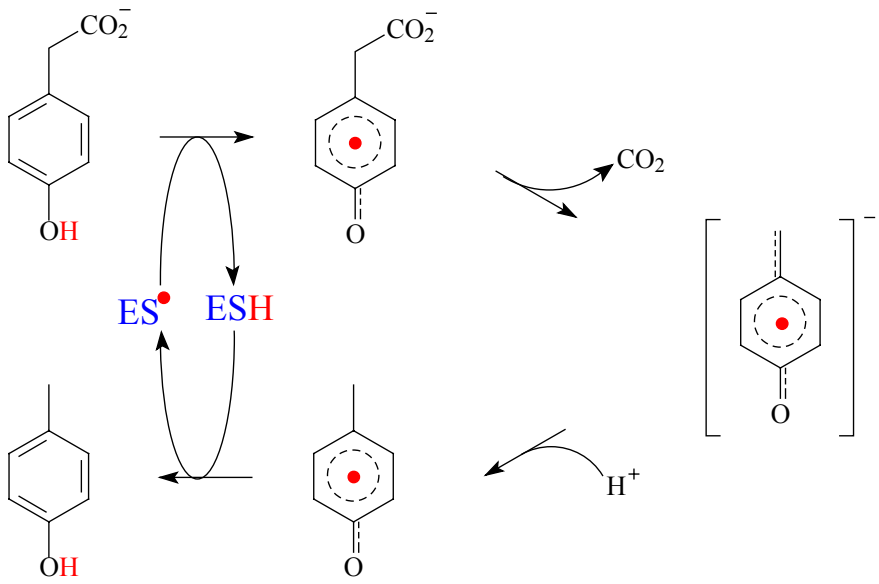



**Figure 1.4 Proposed mechanism of 4-hydroxyphenylacete decarboxylase:** the catalytic cycle is initiated by the abstraction of a hydrogen atom from the substrate by an emzyme thiyl radical. The radical anion decarboxylates spontaneously to release $CO_2$ and forms the product related radical anion, on which the negative charge remaining at the aromatic nucleus is highly stablized by the resonance structure. Protonation of this gives rise to the formation of a neutral radical, which then re-abstracts a hydrogen atom from the enzyme to yield the product *p*-cresol, regenerating the enzyme bound thiyl radical. The oxidation of the phenolic substrate to a radical can be regarded as 'Umpolung' of the electron rich aromatic system to an electronic-poorer radical, which can stablize the evolving negative charge [11].

The Hpd system of *C. difficile* consists of the decarboxylase, which is composed of two subunits, and of its cognate AE. The large subunit of the decarboxylase (HpdB, 100 kDa, 902 amino acids) carries the glycyl radical essential for catalytic activity, while the functional role of the small subunit (HpdC, 9.5 kDa, 85 amino acids) remains to be established [12]. In contrast with the decarboxylase purified from *C. difficile*, which was an almost inactive homo-dimeric protein ($\beta_2$), the recombinant enzyme produced in *E.coli* was a hetero-octameric ($\beta_4\gamma_4$), catalytically competent complex, which was activated using endogenous AE from *C. difficile* or recombinant AE. The loss of the small subunits upon purification of the wild type protein from *C. difficile* has been regarded to cause severe loss of activity. Indeed, the specific activity of purified endogenous decarboxylase was almost 2 orders of magnitude lower than the activity observed with recombinant enzyme after *in vitro* activation (up to 20 U/mg). However, the activation process did not affect the oligomeric state of the recombinant decarboxylase *per se* and preliminary data suggested that the remarkable difference in complex stability between the endogenous and the recombinant enzymes is controlled by a reversible serine-phosphorylation of HpdB (P. Andrei & T. Selmer, unpublished results).

Recombinant HpdBC contained 4 irons and 4 acid-labile sulfurs per glycyl radical subunit [12]. UV/Vis and EPR spectroscopy strongly suggest the presence of [4Fe-4S] $^{2+/+}$-type cubanes, providing a second enzyme-bound co-factor in addition to the glycyl radical, which is not observed in Pfl, Nrd and Gdh, but in Bss. Benzylsuccinate synthase is the second GRE that contains additional small subunits. However, neither the function of this additional prosthetic group for decarboxylase function nor its location in either of the subunits is known. The lack of iron in the initial enzyme preparation and the presence of four strictly conserved cysteinyl residues in the small decarboxylase subunits suggested metal-binding by HpdC, but the recombinant production of iron-free HpdC in *E. coli* contradicts this view [12].

## Aims of this work

Previous studies have shown that the Hpd system from *C. difficile* shows distinct properties, which distinguishes this novel GRE system from the well-characterised Pfl and Nrd system. Among these properties are the presence of small subunits and a metal cofactor in the decarboxylase. Moreover, initial characterisation revealed evidence for a complex, serine-phosphorylation-dependent activity control of the enzyme. The decarboxylase AE also differs in its metal content from the other systems as it obviously contains at least one iron-sulfur centre in addition to the SAM cubane.

The restriction of the unique features to a single system hampered a conclusive analysis of the Hpd system. Therefore, this work aimed to increase our knowledge by the cloning and functional reconstition of related enzymatic system from two other organisms, *C. scatologenes* and *T. forsythensis*.

The recombinant enzymes thus obtained were characterised in order to obtain biochemical data, which may explain the functional relevance of some of the recently discovered features of 4-hydrophenylacetate decarboxylase. In particular the functional role of the small subunit and the activation process of the inactive, radical-free precursor proteins to the active radical form of the recombinant decarboxylases were studied.

# Materials and Methods

## Materials

### Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich (Deisenhofen, Germany), Lancaster (Mühlheim, Germany), Fluka (Buchs, Germany) or Merck (Darmstadt, Germany) and of the highest quality available.
The enzymes used for the molecular biology experiments were from Roche (Mannheim, Germany), MBI Fermentas GmbH (St. Leon-Rot, Germany) or Amersham (Freiburg, Germany).

### Instruments, gases and columns

Anoxic experiments were done in a glovebox (Coy Laboratories, Ann Arbor MI, USA) providing an atmosphere of $N_2/H_2$ (95%/5%). FPLC system and UV/Vis photometer (Ultrascopec 400) were from Amersham Biosciences (Freiburg, Germany). *Strep*-Tactin MacroPrep column and gravity flow *Strep*-Tactin sepharose columns were from IBA GmbH (Göttingen). $N_2$ (99.996%), and $N_2/H_2$ (95:5 v/v) were purchased from Messer-Griesheim (Düsseldolf, Germany).

## Bacterial strains and plasmids

Table 2.1 Bacterial strains

| Strain | Genotype | Reference/Source |
|---|---|---|
| Tuner™ (DE3) pLysS | F⁻ *ompT hsdS$_B$* ($r_B^-$ $m_B^-$) *gal dcm lacY1* (DE3) *pLys*S (Cam$^r$) | Novagen |
| BL21™ (DE3) Codon Plus RIL | *F⁻ ompT hsdS* ($r_B^-$ $m_B^-$) *dcm*$^+$ *Tet*$^r$ *gal (DE3) endA Hte [argU ileY leuW Cam*$^r$*]* | Fermentas |
| Rossetta™ (DE3) pLysS | *F⁻ ompT hsdS$_B$* ($r_B^-$ $m_B^-$) *gal dcm (DE3) pLysS/RARE [argU argW ileX glyT leuW [proL] (Cam*$^r$*)* | Novagen |
| DH5α | *F⁻ ϕ80dlacZ ΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17* ($r_k^-$, $m_k^+$) *phoA supE44 λ⁻ thi-1 gyrA96 relA1* | Invitrogen |
| GM2159 | *F⁻ dam⁻13 Tn9 recF143 thr-1 ara14 leuB6 proA2 lacY1 supE44 galK2 hisG4 rpsL31 xyl-5 mtl-1 argE3 thi-1 tsx-33* | Marinus M.G. *et al*., 1983 |

Table 2.2 Plasmids

| Plasmid | Characteristic | Reference/Source |
|---|---|---|
| pBluescript II SK$^+$ | Ap$^r$, high copy number cloning vector. | Stratagene |
| pASK-IBA7 | Ap$^r$, cloning and expression vector, providing *tet*A promoter, and N-terminal *Strep*-tag II. | Institut für Bioanalytik, Göttingen |
| pPR-IBA2 | Ap$^r$, cloning and expression vector, providing T7 promoter, and N-terminal *Strep*-tag II. | Institut für Bioanalytik, Göttingen |
| pCR2.1 | Ap$^r$ and Kan$^r$, cloning vector. | Novagen |

Table 2.3 Molecular biology kits

| Kit | Company |
|---|---|
| Extensor Hi-Fidelity PCR Enzyme Mix | Abgene (Hamburg, Germany) |
| QIAquick PCR purification kit | Qiagen (Hilden, Germany) |
| QIAquick gel extraction kit | Qiagen (Hilden, Germany) |
| QIAprep Spin Miniprep kit | Qiagen (Hilden, Germany) |
| T4 DNA ligase | Amersham Pharmacia Biotech (Freiburg) |
| Fast-Link DNA Ligase | BIOenzyme (Oldendorf, Germany) |
| TA cloning kit | Invitrogen (Karlsruhe, Germany) |
| Easystart PCR kit | M$\beta$P (UK) |
| TOPO Walker kit | Invitrogen (Karlsruhe, Germany) |
| PfuUltra$^{TM}$ Hotstart DNA Polymerase | Stratagene |
| Subcloning Efficiency$^{TM}$ DH5$\alpha^{TM}$ Chemically Competent *E. coli* cells | Invitrogen (Karlsruhe, Germany) |

## Media and plates

All the media were autoclaved at 121°C and 1 bar (15 psi) for 20 min.

### LB medium

| | |
|---|---|
| Trypton | 10.0 g |
| Yeast extract | 5.0 g |
| NaCl | 10.0 g |
| Distilled water | 0.8 L |

The medium was adjusted to pH 7.0, filled up to 1 L and autoclaved.

### LBG medium

| | |
|---|---|
| Trypton | 10.0 g |
| Yeast extract | 5.0 g |
| NaCl | 10.0 g |
| Glucose | 2.0 g |
| Distilled water | 0.8 L |

The medium was adjusted to pH 7.0, filled up to 1 L and autoclaved.

### Solid media

Unless otherwise stated, LB (Luria-Bertani Medium) agar plates (1.5%) supplemented with the required antibiotic(s) were used to isolate colonies.

For blue-white selection of individual colonies on agar-plates, X-Gal (40 µl of 40 µg/µl x-Gal in DMF) was evenly spread on the surface of the plates just prior to use.

## Oligonucleotides

All the primers used were synthetisied by MWG Biotech (Ebersberg).

Table 2.4 Cloning primers for the *T. forsythensis* genes

| Name | Nucleotide sequence (5'-3') | Cloning site |
|---|---|---|
| DecCterm2 | TAGGGAGA**GGATCC**TTCTCCCTATATCTTG | BamHI |
| ActCterm2 | TATGGCAT**GGATCC**AATTGATTATCGAACAATTGCTTATGTT | BamHI |
| Dec-SacII | ATTCATTAAGGTA**CCGCGG**TTATGGCTAAGAGTTTCAAAG | SacII |
| Act-SacII | AACACAAACGGTA**CCGCGG**TAATGCTGATGGAGAATAAAG | SacII |

Table 2.5 Sequencing primers for *T. forsythensis* genes

| Name | Nucleotide Sequence (5'-3') |
|---|---|
| sDec700-1 | GTTGACCATGCGTAAAGGC |
| sDec700-2 | CCATATCCGGACTTTCTCCC |
| sDec700-3 | GAGTATTGGATTCCGGGAGG |
| sDec700-4 | ACAATATTCTGCTCGGCTGG |
| asDec800-1 | ATTGATGACTTCACGCCCC |
| asDec800-2 | GATTGCCGTCTTGATTACGC |
| asDec800-3 | ACGGTCTTTTGCCAGTATGC |
| asDec800-4 | CTGCGAATTTTGTGACATGG |
| M13-forward(s) | GTAAAACGACGGCCAGT |
| M13-reverse(as) | CCTTTGTCGATACTGGTAC |

Table 2.6 Primers for cloning and sequencing of the 4-hydroxyphenylacetate decarboxylase genes from *Clostridium scatologenes*

| Name | Nucleotide Sequence (5'-3') |
|---|---|
| Primer 1 | CGYGTTGCYGGRTTYACWCAATATTGG |
| Primer 2 | ACACCATTTACAMMTCARTGGRCAWCC |
| Primer 3 | TCAAGCTTCTTGTTGTGAGG |
| Primer 4 | TGAATCAATGGAAAGAGTACATGG |
| Primer 5 | CCTGCAAGTCTTCCATCAGC |
| Primer 6 | CCAATAAAATGAACACCTGTACCAC. |
| Primer 7 | GCACCTCCTGGAATCATTGTTAC |
| Primer 8 | GATGGTGAAGGAAGTGAAGC |
| Primer 9 | CTTTGTGCACTAACAAAGGG |
| Primer 10 | TCAAGCTTCTTGTTGTGAGG |
| Primer 11 | CTGGTTCTACAACCTGGTCC |
| Primer 12 | GACCAGGTTGTAGAACCAGTGTA |
| Primer 13 | AGGA**CCGCGG**TTTAATGAACGTTAAAGAAACTAAACTTGAAGATG |
| Primer 14 | ATATC**GGATCC**CTATTCTGCTTTATAACTAGGAC |
| Primer 15 | A**CCGCGG**TAATGATATGAAGGAAAAAGGTTTAATATTTGATATACAAAGC |
| Primer 16 | TACCTC**GGATCC**AGCTCTATTAAAATATAAGATTAAAAAGG |
| M13-20 | GTAAAACGACGGCCAGTG |
| M13-rev | GGAAACAGCTATGACCATGA |

Table 2.7 Mutagenic primers for the generation of hybrid decarboxylases

| Name | Nucleotide Sequence (5'-3') |
|---|---|
| Cscat-BC-BbsI-s | GCTACAGCTGATG**GAAGAC**TTGCAGGAAC |
| Cscat-B-ClaI_rev | GC**ATCGAT**ATTACCTCCATTCATTTTTTAGTATACCTTATTATTTATC |
| IadC_s(ClaI) | AAAAA**ATCGAT**GGAGGTAATAGCAATGCGCC |
| IadC_as(BamHI) | ATATC**GGATCC**CTATTCTGCTTTATAACTAGGAC |
| Cdiff_StuCla_s | ATGCTAGAATTC**AGGCCT**GTCTTGATGCACC |
| Cdiff_StuCla_as | GCTTTCT**ATCGAT**TACACCCCTTCATACTCTGTTC |
| Cdiff_ClaBamHI_s | AGAACA**ATCGAT**AAGGGGTGTAAAATATGAGAAAGCATAG |
| Cdiff_ClaBamHI_as | ACCTCGAG**GGATCC**TTTTGAC |

## Microbiological Methods

### Growth of *E. coli*

Recombinant *E. coli* cells were cultivated on LB–agar plates supplemented with required antibiotics at 37 °C. Isolated colonies were used for inoculation of overnight cultures. These precultures were then used to inoculate 1.5 L of LBG-medium supplemented with appropriate antibiotics for the production of recombinant proteins pre-equilibrated at the necessary temperature (table 3.2). The expression of plasmid-encoded genes was induced by anhydrotetracycline (AHT) or IPTG.

### Storage of bacterial strains

Recombinant strains and bacterial cultures were stored at –80 °C in appropriate media supplemented with 8 % [v/v] sterile glycerol.

## Molecular biology

### Isolation of DNA from *C. scatologenes*

The genomic DNA of *C. scatologenes* was prepared from the aerobically harvested cells, which were grown anaerobically in CBM for 16 h at 37 °C. The DNA was extracted by the follwing procedure:

**Solutions:**

| | |
|---|---|
| Tris-sucrose buffer | 50 mM Tris/HCl pH 8.0, 25 % (w/v) Sucrose; |
| Tris-EDTA buffer | 50 mM Tris/HCl pH 8.0, 25 mM $Na_2EDTA$ pH 8.0; |
| SDS buffer | 1 % (w/v) SDS in Tris/EDTA Buffer; |
| TE buffer | 10 mM Tris/HCl pH 8.0, 1 mM EDTA. |

The cells were resuspended in 3 ml Tris-sucrose buffer. 1 ml Tris-sucrose buffer containing 100 mg lysozyme (Serva) was added and the sample was gently shaked for 90 min at 37 °C. Tris-EDTA buffer (4 ml) was then added and the mixture was incubated on ice for additional 15 min. 10 ml SDS buffer, 10 µl of 20 mg/ml RNAse and 20 mg Proteinase K (Roche) were added and the solution was incubated at 37 °C for 3 h.

Phenol (15 ml) (Roti-Phenol for DNA/RNA isolation, saturated with TE buffer, Roth) was added and the mixture was gently mixed 2-3 times. The aqueous and organic phases were separated by centrifugation at 5000 g at 6 °C for 20 min. The aqueous phase containing the nucleic acids (upper layer) was transfered carefully to a new Falcon tube and the phenol extraction was repeated twice with equal volumes of phenol. The nucleic acids in the aqueous phase were washed with an equal volume chloroform/isoamyl alcohol (24 + 1, v + v) and subsequently dialysed overnight in 5 L TE buffer at 4 °C. Genomic DNA was stored at 4 °C.

**Determination of DNA concentration and purity**

DNA concentrations were calculated from the absorption at 260 nm using an absorbance of 1.0 for 50 µg/ml of double-stranded DNA.

$$\text{DNA concentration } (\mu g/ml) = \Delta E_{260} \times 50 \times \text{dilution}$$

**Plasmid DNA isolation**

Plasmid DNA was isolated by using either QIAprep Spin Miniprep column according to the suppliers instructions or alkaline lysis method.
5 ml of LB or standard I medium with the required antibiotic(s) was inoculated with bacterial colonies and incubated at 37 °C room overnight with shaking. The overnight culture was then harvested by centrifugation. The pellet was re-suspended in 100 µl 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl, pH 8.0, then lysed by 200 µl 1 % SDS in 200 mM NaOH, and neutralized with 150 µl 3 M potassium acetate (pH 4.8). Precipitates were removed by centrifugation and plasmid DNA was precipitated with 2 volumes of isopropanol. The DNA was then washed with 1ml 70 % ethanol, dried and re-dissolved in either 10 mM Tris/HCl, pH 8.0, 1 mM EDTA (TE buffer) or sterile water.

**DNA gel eletrophoresis**

The required amount of agarose was suspended in 1x TAE buffer, boiled in a microwave and cooled to about 55 °C prior pouring into a mould. After 15-30 min at room temperature, the slot-forming template (comb) was removed gently and the gel was placed in the electrophoresis device and run at 80-100 Volt for 60 to 90 min. The location of DNA was

visualised with UV light at a wavelength of 256 nm (analytical gels) or 340 nm (preparative gels) after staining with ethidiumbromide solution.

**50x TAE buffer**

| | |
|---|---|
| 40 mM Tris | 242 g |
| 20 mM glacial acetic acid | 57.1 ml |
| 0.5 mM EDTA (pH 8.0) | 100 ml 0.5 M pH 8.0 |
| | Fill to 1 L with $H_2O$ |

**Loading buffer**

| | |
|---|---|
| 0.25 % (w/v) | bromphenol blue |
| 0.25 % (w/v) | xylene cyanol FF |
| 0.25 % (w/v) | orange G |
| in 40 % (w/v) | sucrose |

Unless otherwise stated, *PstI*-digested DNA of the bacteriophage λ was used as a standard to estimate the relative size of DNA fragments in agarose gels.

**Purification of DNA from agarose gel**

DNA bands of the desired size were cut out with sterile blades from gels under illumination at 340 nm. The extraction of DNA from the gel slice was performed using the QIAquick gel extraction kit following the manufactures instructions.

**DNA restriction and ligation**

Unless otherwise stated, restriction digests were performed at 5- to 20- fold over-digestion (e.g. 1 µg of DNA was digested with 10 U endonucleases for 1 h to achieve 10-fold over-digestion). The digested fragments were analysed on or prepared from agarose gels.
DNA fragments were re-ligated with T4 ligase using 1 unit (cohensive ends) or 5 units (blunt ends) of enzyme per µg of DNA at 16°C overnight.

**Phenol-chloroform extraction**

For the phenol-chloroform extraction, the procedures were carried out adding the same volume of phenol: chloroform solution to overnight digests or ligation mixtures and centrifuged at 13,000 rpm for 3 min. The upper layer was transferred to a new Eppendorf tube and 10 µl 3 M sodium acetate, 1 µl glycogen (20 mg/ml), 300 µl cold ethanol were added. After incubation on dry ice for 10min, DNA was pelleted at 13,000 rpm for 15 min at 4 °C. The pellet was then washed with 500 µl cold 80 % ethanol. Residual Ethanol was allowed to evaporate prior dissolving the DNA in 40 µl sterile water.

**Preparation of electrocompetent cell**

Growing overnight cultures on LB media, supplemented with the required antibiotic(s), were used to inoculate 500 ml fresh medium. At $OD_{578nm}$ = 0.35-0.5 the cultures were chilled on ice for 30 min, and centrifuged at 6,000 rpm for 10 min in a SLA-3000 rotor at 4 °C. The cell pellets were washed twice with 500 ml and 250 ml ice-cold sterilized $H_2O$, respectively. Then the cells were resuspended in 10 ml ice cold, sterilized 10 % glycerol and centrifuged at 5000 rpm for 10 min at 4 °C in Falcon tubes. The final pellets were resuspended in 1 ml ice cold, sterilized 10 % glycerol solution, and aliquoted into 50 µl portions in sterile Eppendorf-tubes and stored at –80 °C.

***E. coli* transformation by electroporation**

Transformation of electrocompetent *E.coli* cells was carried out in 1 mm disposable, pre-chilled PEQlab electroporation cuvettes (Biotechnological GmbH, England). An appropriate number of competent cells were thawn on ice. Plasmid (about 50 ng of DNA) or 2-5 µl ligation mixtures were added. Electroporation was carried out at 1.8 V, 200 Ω, and 25 µF. 250 µl LB medium was added to the cuvettes after the pulse. The cells were either plated directly on LB-agar plates containing the corresponding antibiotics for selection (for purified plasmid DNA) or grown at 37 °C for 1h prior plating in order to increase plating efficiency (for ligation mixtures).

**Transformation of chemically competent cells**

Chemically competent cells were used for the generation of expression strains for production of recombinant proteins and for the amplification of plasmid DNA. The transformation of Rosetta (DE3) single™ chemical competent cell and the subcloning efficiency DH5α™ chemically competent *E. coli* cells was performed according to the manufacturers instructions.

**PCR**

**Normal PCR reaction**

In this work, the normal *Taq* DNA polymerase and a various proofreading DNA polymerases were used for the amplification of target DNA. The composition of individual PCR reactions and the temperature profiles for DNA amplification were indicated below:

***Taq* polymerase or Extensor Hi-fidelity polymerase**

**Composition of a standard mixture of 50 µl final volume**

| | |
|---|---|
| 10x PCR reaction buffer | 5 µl |
| 5 mM dNTPs solution | 1 µl |
| 25 pmol/µl sense primer | 0.4 µl |
| 25 pmol/µl anti-sense primer | 0.4 µl |
| Template | varies |
| Polymerase | 0.25 µl |
| Sterile water | fill to 50 µl |

**EasyStart PCR Mix-in-a-Tube**

The EasyStart PCR Mix-in-a-tube system yielded optimum reaction conditions for Taq polymerase and some other enzymes. Most components were premixed and only primer, template and polymerase were added just prior starting PCR reaction.

**Components for the upper layer yielding 100 µl final PCR volume**

| | |
|---|---|
| 25 pmol/µl sense primer | 1.25 µl |

| | |
|---|---|
| 25 pmol/µl anti-sense primer | 1.25 µl |
| Template | varies |
| Polymerase | 0.50 µl |
| Sterile water | fill to 50 µl |

**PfuUltra™ Hotstart DNA Polymerase**

**50 µl final volume**

| | |
|---|---|
| 10x PfuUltra™ HF reaction buffer | 5 µl |
| 5 mM dNTPs solution | 2 µl |
| 25 pmol/µl sense primer | 0.4 µl |
| 25 pmol/µl anti-sense primer | 0.4 µl |
| Template | varies |
| PfuUltra™ Hotstart DNA Polymerase | 1.0 µl |
| Sterile water | fill to 50 µl |

**PCR programs used for *Taq* polymerase and Extensor Hi-Fidelity polymerase**

| | | | |
|---|---|---|---|
| Initial denaturation: | 94 (95) °C | 2 min (5 min) | |
| Denaturation: | 94 (95) °C | 15 s (45 s) | } 29 to 34 cycles |
| Annealing: | determined by primers | 30 s | |
| Extension: | 68 (72) °C | x min | |

(for the normal *Taq* polymerase, 1 min/1 kb; for the Extensor Hi-Fidelity polymerase, according to the manufacturer's manual)

Final extension: 68 (72) °C

(conditions for *Taq* polymerase are given in brackets)

**PCR program used for PfuUltra™ Hotstart DNA Polymerase**

| | | | |
|---|---|---|---|
| Initial denaturation: | 95 °C | 1 min | |
| Denaturation: | 95 °C | 1 min | } 30 cycles |
| Annealing: | determined by primers | 1 min | |
| Extension: | 72 °C | x min | |

(1 min for targets less than 1 kb, 1 min /kb for targets bigger than 1 kb)

Final extension: 72 °C 10 min

## Colony PCR

Positive colonies were identified by colony PCR. Therefore, individual colonies were picked from the agar, suspended in PCR mix using XYZ primers.

**Colony PCR (25 µl final volume)**

| | |
|---|---|
| 2.5 µl | 10x PCR reaction buffer |
| 0.4 µl | sense primer (25 pmol/µl) |
| 0.4 µl | anti-sense primer (25 pmol/µl) |
| 1.0 µl | 25 mM $MgCl_2$ |
| 1.0 µl | 5 mM dNTPs solution |
| 0.2 µl | Taq DNA polymerase (Amersham Pharmacia Biotech) |
| 19.5 µl | sterile $H_2O$ to 25 µl final volume |

**PCR program used**

| | | | |
|---|---|---|---|
| Initial denaturation: | 95 °C | 5 min | |
| Denaturation: | 95 °C | 45 s | } 30 cycles |
| Annealing: | determined by primers | 30 s | |
| Extension: | 72 °C | x min | |
| (for the normal *Taq* polymerase, 1 min for 1 kb) | | | |
| Final extension: | 72 °C | 10 min | |

## Inverse PCR

A major limitation of the PCR reaction is that it enables only the amplification of DNA situated between two convergent primers. Hence, the amplification of unknown sequences from genomic DNA usually requires at least partial knowledge of the target gene DNA sequence, which is usually obtained for unknown genes using either highly conserved sequence motifs in evolutionary related genes or partially known amino acid sequences for deduction of degenerated primer pairs. While the utilization of oligonucleotide primers that hybridize to opposite strands, results in an exponential increase in the number of copies of the region between these primers, DNA sequences that lie immediately outside the primers are inaccessible. In contrast, inverse PCR (iPCR) permits the amplification of the regions of unknown sequences flanking any DNA segment of known sequence, either upstream or

downstream or both. The primer pairs used in these experiments are homologous to the ends of the known sequence, but are chosen in divergent orientation (e.g. facing outside of the known sequence).

Due to the enormous size and heterogeneity of genomic DNA, divergent primers will not yield defined PCR products with this template, while the amplification on smaller, circular DNA molecules will yield products. These small circular fragments are readily generated from genomic DNA by cleavage and re-ligation of the genome by restriction enzymes. The fragment size distribution in individual digests is readily accessible by Southern blot analysis. Suitable fragments are religated under conditions that favour the formation of monomeric circles (e.g. ligation in diluted solution). The circular ligation products are suitable templates for the amplification of DNA sequences flanking the known region and must contain the restriction site used for the generation of circular DNA. Typical conditions for endonuclease digest, re-ligations and iPCR experiments are given below.

**Restriction digestion of genomic DNA**

| | |
|---|---|
| Genomic DNA (800 ng/µl) | 6 µl |
| 10× restriction buffer | 5 µl |
| Restriction enzyme (10 U/µl) | 2 µl |

Sterile $H_2O$ fill to 50 µl, 37 °C, 3 h

**Circularization**

2.8 ng/µl final concentration of size restricted DNA fragments

| | |
|---|---|
| 20 U | T4 ligase |
| 50 µl | 10x ligation buffer |

Fill to 500 µl with sterile $H_2O$, 16 °C, overnight

The T4 ligase was removed prior PCR amplification by phenol-chloroform extraction and precipitated DNA was re-dissoled in $H_2O$.

**iPCR (50 µl final volume)**

| | |
|---|---|
| 10x PCR reaction buffer | 5 µl |
| 5 mM dNTPs solution | 1 µl |
| 0.25 µM primer 1 | 0.4 µl |
| 0.25 µM primer 2 | 0.4 µl |

| | |
|---|---|
| Template (X ng) | 5-20 µl |
| Polymerase | 0.25 µl |
| Sterile water | fill to 50 µl |

The PCR products were purified by gel electrophoresis and directly custom sequenced using the PCR primers.

**Southern blotting**

In this work, we used DIG-labelled probes, and the procedures were described as the following:

**20X SSC stock (for 1 liter)**

175.3 g NaCl (3 M)
88.2 g Na-Citrate (0.3 M)
pH to 7.0 with HCl

**Prehybridization buffer**

5x SSC
0.1 % milk powder
1 % N-laurylsarkosine
0.02 % sodium dodecyl sulfate (SDS)

**Hybridization buffer**

15 µl DIG-labeled probes in 30ml pre-hybridization buffer

**Buffers used for the washing of the membrane and color detection**

Buffer 1: 0.1 M Maleic acid
0.15 M NaCl
Adjust pH to 7.5 with NaOH;

Buffer 2: 1 % milk powder in buffer1;

Buffer 3: 0.1 M Tris/HCl

0.1 M NaCl

50 mM $MgCl_2$

Adjust pH to 9.5 with HCl;

Buffer A: Antibodyconjugate (Anti-digoxigenin-AP, Fab fragments, 0.75 U/µl, Boeringer Manheim), 1:5000 in buffer 2;

Solution for color detection (NBT/BCIP):

45 µl NBT solution (Nitroblue-tetrazolium-salt, Roth, 75 mg/ml in 70 % DMF)

35 µl BCIP solution (5-brom-4-chlor-3-indolyphosphate, toluidine salt, Roth, 50 mg/ml in 100 % DMF)

In 10 ml buffer 3;

**Probe preparation**

The hexanucleotide mixture (10x), the dNTP labeling mixture (10x) and the Klenow enzyme (labeling grade) were purchased from Roche. The probe preparation procedures were as the following: dilute DNA template (PCR products) in sterile $H_2O$ to a total volume of 15 µl (10 ng-3 µg DNA), then heat-denature the DNA template in a boiling water bath for 10 min, and quickly chill it on ice for about 10 min. Add 2 µl hexanucleotide mixture, 2 µl dNTP labeling mixture to the tube (on ice), and 1 µl Klenow enzyme to a final concentration of 100 U/ml, mix. Incubate the reaction tube at 37 °C overnight. To stop the reaction, add 2 µl EDTA to the reaction tube. The final 15 µl labelling probe was then directly used for hybridization.

**Southern blot analysis**

Genomic DNA of *Clostridium scatologenes* was digested using various restriction enzymes (see figure 3.4, 3.6). Following the fractionation on a 0.8% agarose gel, acid depurination (0.25 M HCl), alkaline denaturation (0.5 M NaOH, 1.5 M NaCl), and neutralisation (1 M Tris-HCl, pH7.5, 1.5 M NaCl) steps were performed during the blotting procedure onto a nylon membrane (Qiagen, Hilden, Germany). Then, the membrane was placed on Whatman 3 MM paper wetted with 2 x SSC and UV cross-linked to the membrane. The membrane was now ready for hybridization.

DIG-labelled probes used to detect the target DNA sequences on the blot were the PCR products labelled with DIG-11-dUTP according to the manufacturer's instructions (Roche), which are shortly described above.

After pre-hybridization of the membrane with 40 ml pre-hybridization buffer for 1h at 68 °C, the hybridization was performed overnight at 68 °C in hybridization buffer. Membrane was then washed twice for 5min at room temperature with 2 x SSC plus 0.1 % SDS, twice 15min at room temperature with 0.5 x SSC, 0.1 % SDS. Then the digoxigenin-labelled probes were detected with alkaline phosphatase-α-digoxygenin-antibody conjugates (Roche) according to the manufacturer's instruction. Simply, the membrane was washed with buffer1, buffer2 and bufferA for 1 min, 30 min and 30 min, respectively, then 3 times in buffer1 for 10 min and 5 min in buffer3. The membrane was then put into a plastic bowl supplemented with 10 ml color detection solution and the reaction was performed in a dark place till a clear signal observed.

**DNA sequencing**

Custom sequencing of DNA molecules (PCR products and vectors) were performed by MWG-Biotech, Ebersberg. Standard vector-derived primers and internal primers were used for the complete sequencing of DNA templates using dye-terminator chemistry. In order to exclude DNA polymerase-derived mutations, three different clones from three different PCRs were double-stranded sequenced.

**Construction of the plasmids for sequencing and protein expression**

**Construction of plasmids for sequencing**

**Using pBluescript II SK$^+$ vector**

Genomic DNA from *T. forsythensis* strain ATCC 43037 was purchased from ATCC and used as template for amplification of the *tfdBC* and *tfdA* genes by PCR. The primer pairs were DecCterm2 + Dec-SacII and ActCterm2 + Act-SacII (Table 2.4) for *tfdBC* and *tfdA*, respectively. These primer pairs were designed in order to introduce suitable cloning sites (*SacII and BamHI*) up- and downstream of the coding sequences in order to allow a directed, *in frame* cloning of the genes in pASK-IBA7 vectors (see below). The PCR products were purified using the QIAquick PCR purification kit, and digested with *SacII* and *BamHI*, purified by gel electrophoresis and ligated with T4-ligase (Amersham Pharmacia Biotech) into the multicloning site of pBluescript II SK$^+$. Three clones of the three individual PCRs were double-stranded sequenced for each construct.

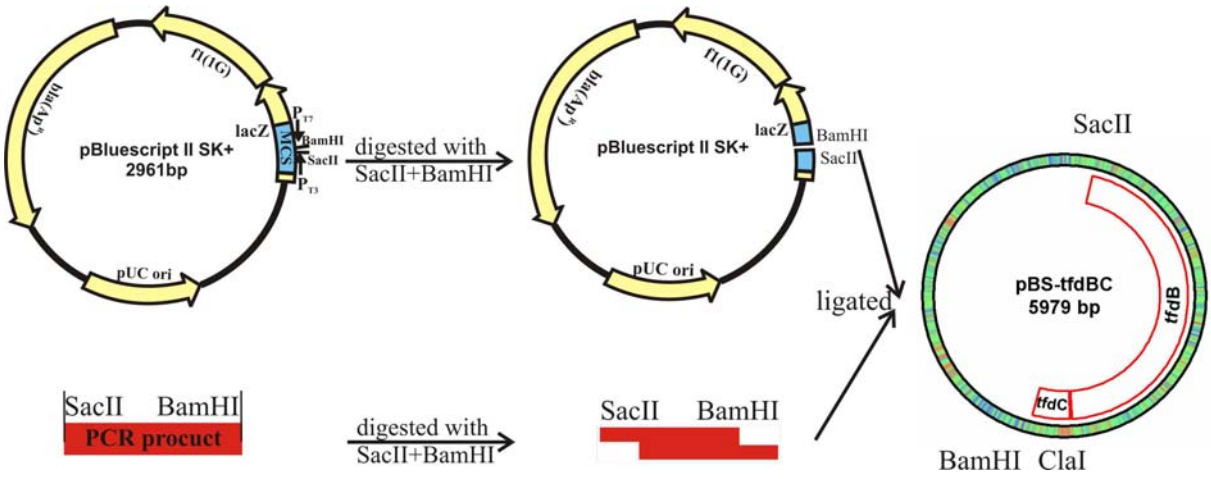



**Figure2.1 The strategy used to clone *tfdBC* or *tfdA* into pBluescript II SK + for sequencing.** The final plasmids are named pBS- *tfdBC and* pBS- *tfdA*, respectively.

**Using TA-cloning kit**

For the *C. scatologenes* system, the pCR2.1 vector (provided in the TA cloning kit) was used instead of pBluescript II $SK^+$. The TA Cloning® vector, pCR2.1, contains the lacZ-alpha complementation fragment for blue/white screening, ampicillin and kanamycin resistance genes for selection, and a versatile polylinker. Using genomic DNA from *C. scatologenes* strain DSM 757 as template to amplify *csdBC* and *csdA* genes by PCR techniques. The primers used are shown in Table 2.6. The PCR products were purified and directly ligated into the pCR2.1 vector using the TA cloning kit. Three clones of three individual PCRs were double-stranded sequenced for each gene.

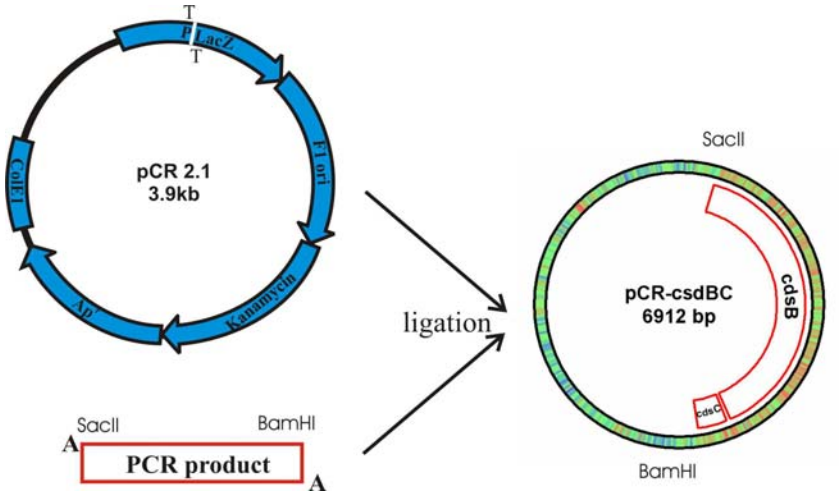



**Figure2.2 Strategy used to clone the PCR fragments of *csdBC/A* into pCR2.1 vector for sequencing.** The final plasmids are named pCR-*csdBC and* pCR-*csdA*, respectively.

## Construction of expression plasmids

### pASK-IBA7-derived expression plasmids

Genes inserted *in frame* with the N-terminal Strep-tag II in the pASK-IBA7 vector are expressed under control of the tetracycline promotor/operator. The strength of the tetA promoter is comparable to that of the lac-UV5 promoter, and it is fully induced by the addition of anhydrotetracycline at a concentration that is not yet antibiotically effective [80]. The genes were excised from the sequencing vectors with *SacII* and *BamHI*, purified by gel electrophoresis and ligated into the *SacII*/*BamHI* sites of the expression vector. The expression plasmids were named pASK7-*tfdBC*, pASK7-*tfdA*, pASK7-*csdBC* and pASK7-*csdA*. The strategy of the subcloning steps is shown in figure 2.3.

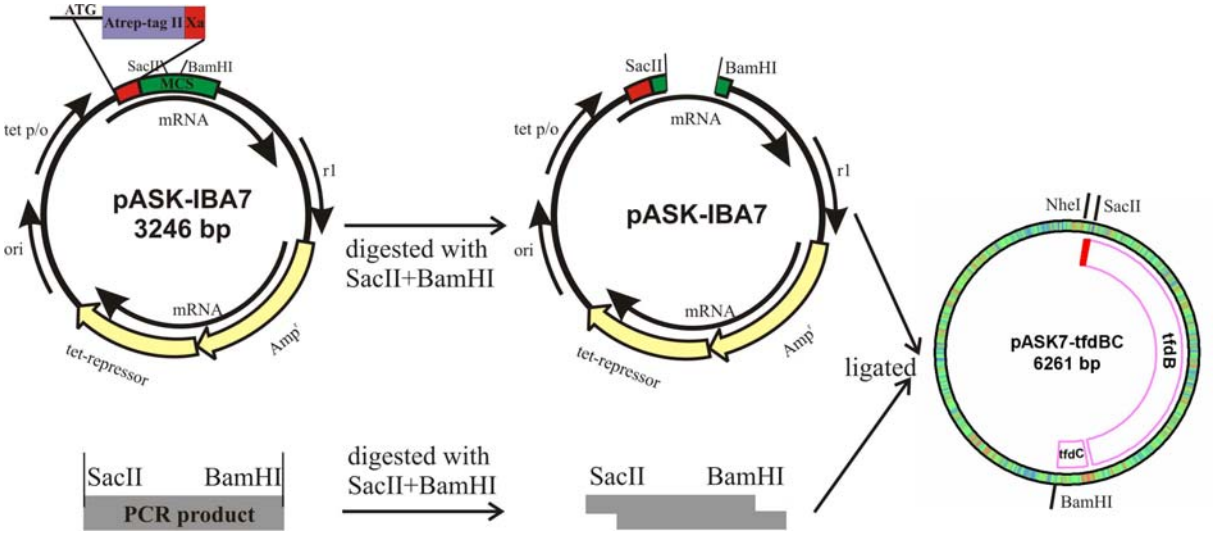



**Figure 2.3 Subcloning strategies for the generation of expression plasmids in pASK-IBA7.** Details are given in text.

### Cloning of *tfdBC* and *tfdA* into pPR-IBA2 vector

The pPR-IBA2 vector encodes the T7 expression system, which uses the T7 promoter and T7 RNA polymerase for high-level transcription of the gene of interest. The expression of the target genes is accomplished in *E. coli* BL21 cells, which contain a chromosomal copy of the T7 RNA polymerase gene. The latter is under control of the lacUV5 promoter, which can be induced by IPTG. For the subcloning of *tfdBC* and *tfdA* into pPR-IBA2, the expression plasmids pASK7-*tfdBC* and pASK7-*tfdA* were digested with *NheI* and *BamHI*, the desired fragments were subsequently ligated into pPR-IBA2. The resulting plasmids were designated pPR2- *tfdBC* and pPR2-*tfdA*.

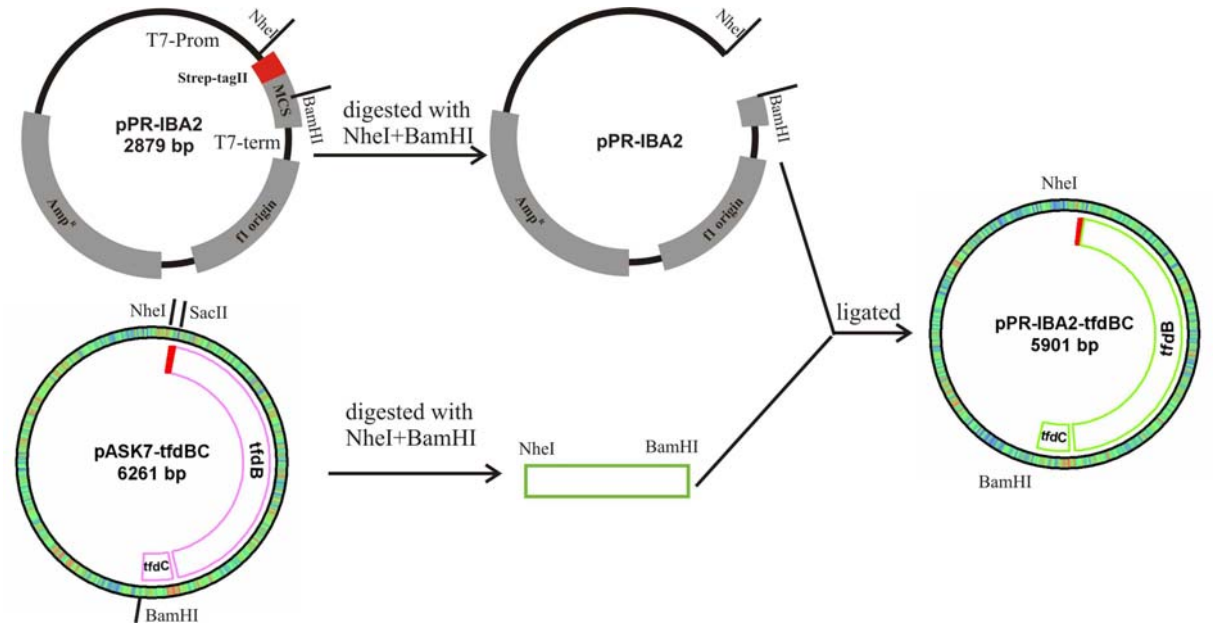



**Figure 2.4 Subcloning of *tfd* genes into pPR-IBA2.** The details are given in text.

**Expression plasmids for hybrid decarboxylases**

In order to study the function of the small subunits, plasmids encoding hybrid decarboxylases were generated, which allowed the co-expression of individual small subunit genes (*hpdC*, *csdC* and *tfdC*) in all nine possible combinations with the glycyl radical subunit genes (*hpdB*, *csdB* and *tfdB*). While a natural unique *ClaI*-cleavage is located between the stop codon of *tfdB* and the ribosomal binding site of the *tfdC* gene, neither the *hpdBC* genes nor the *csdBC* genes contained *ClaI*-sites. Hence this endonuclease cleavage site was generated between the stop codons of the glycyl radical subunit genes and the RBS of the small subunit genes of the latter systems, and subsequently used to generate the desired genetic combinations of subunits by exchange of *ClaI*/*BamHI*-fragments between the parental expression plasmids.

The *ClaI* site was introduced in the pASK7-*hpdBC* and the pASK7-*csdBC* expression plasmids in a two-step procedure. First, the small subunit gene was eliminated and a *ClaI*-site was generated downstream of the glycyl radical subunit gene. Therefore, mutagenic antisense primers with the *ClaI*-site directly followed by a *BamHI*-site were designed (Cscat-B-ClaI_rev, Cdiff_StuCla_as). These primers were used in combination with internal sense primers from the glycyl radical subunit genes, which were located upstream of unique *StuI*-(*hpdB*) and *BbsI*-sites (*csdB*). The resulting PCR products obtained with the expression plasmids (pASK7-*hpdBC* or pASK7-*csdBC*) as templates were cut with the appropriate

enzymes and used to replace the small subunit genes in the parental plasmids by a *ClaI*-cleavage site. Note that the elimination of the small subunit genes also generated plasmids for the expression of the glycyl radical subunits alone. Subsequently, a second mutagenic PCR was performed in order to generate the *ClaI*-site in the desired position upstream of the small subunit gene RBS. These sense primers (IadC_s (ClaI), Cdiff_ClaBamH_s) were used in combination with a vector-derived antisense, which was located dorwnstream of the *BamHI*-site. The *ClaI*/*BamHI*-cut PCR products were subsequently ligated in the previously generated vectors, and the wild-type co-expression of *hpdBC* and *csdBC* (designated pASK7-*hpdB(ClaI)C* and pASK7-*csdB(ClaI)C*, respectively) was restored. Subsequently, the *ClaI*/*BamHI*-fragments carrying the small subunit genes were exchanged between individually altered wild type plasmids to generate hybrid decarboxylase expression plasmids pASK7-*hpdB*/*tfdC*, pASK7-*hpdB*/*csdC*, pASK7-*tfdB*/*csdC*, pASK7-*tfdB*/*hpdC*, pASK7-*csdB*/*tfdC*, pASK7-*csdB*/*hpdC*. (See figure 2.5)

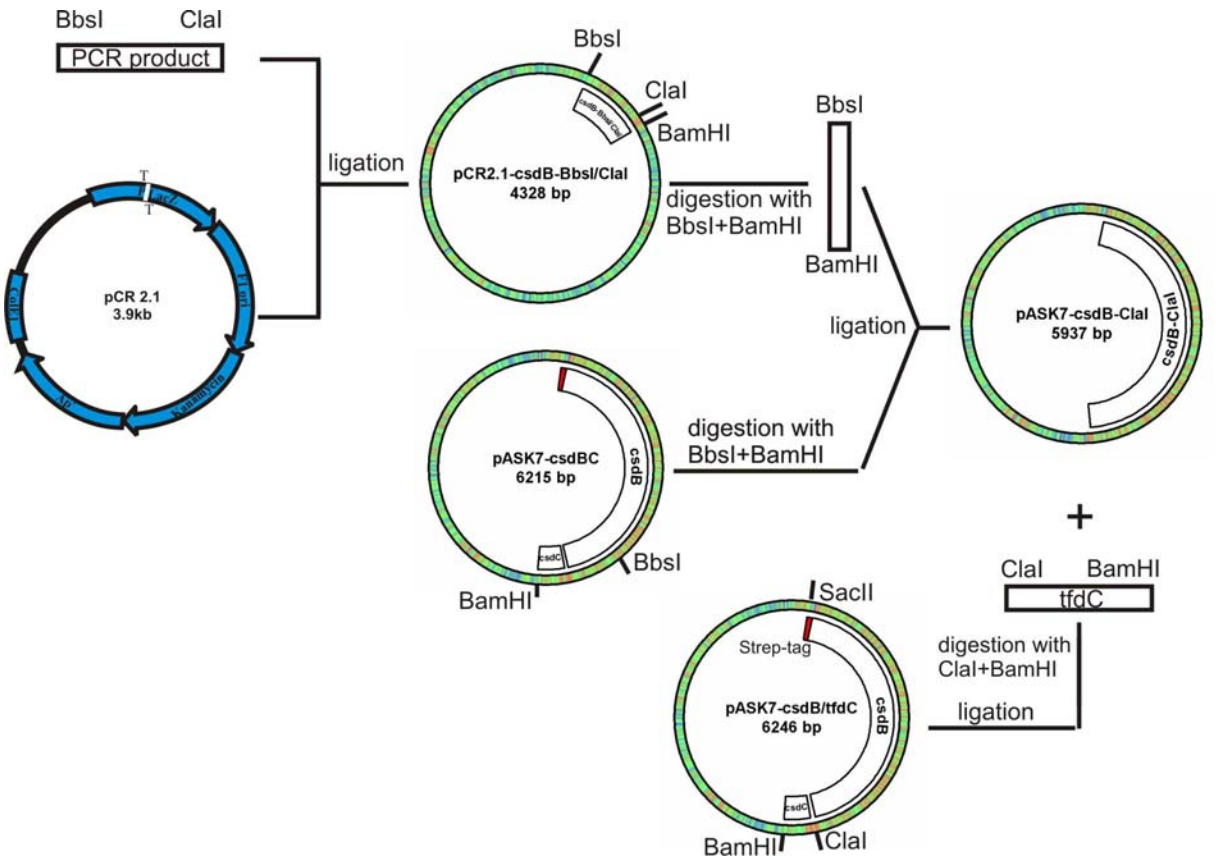



**Figure 2.5 Cloning strategies to generate the *ClaI*/*BamHI* fragments and final expression plasmids for the production of hybrid decarboxylases.**

## Biochemical methods

### Gene expression and protein purification

Recombinant proteins were produced in *E. coli* BL21 (DE3) Codon Plus-RIL™ or Rosetta (DE3) pLysS™ cells, which were transformed with the individual expression plasmids and kept under selection with 50 µg/ml carbenicillin and 34 µg/ml chloramphenicol during growth. Unless otherwise stated, all cultures were grown aerobically and aerated by constant stirring. Initially, 200 ml of LBG medium was inoculated with freshly isolated colonies from LB agar plates and grown over night at 37°C. This pre-culture was used to inoculate 1.5 L of medium. The production parameters are summerized in table 3.2. The cells were grown to $OD_{578nm}$ of 0.8-1.0, then 27 ml ethanol was added and the culture was allowed to grow another 30 min prior induction, typically harvested 3 h post-induction, washed with TBS and frozen at –80 °C.

#### Protein purification

Unless otherwise stated, all protein purification steps were performed in an anoxic chamber (Coy Laboratoroes, Ann Arbor, MI, USA) with a $N_2/H_2$ (95 %/5 %) atmosphere at 15-20 °C.
The decarboxylases were typically purified from ~3.0-5.0 g of wet packed cells. The cells were initially suspended in 10 to 20 ml anoxic buffer BC (100 mM Tris/HCl, pH 7.5, 150 mM NaCl, 5 mM $(NH_4)_2SO_4$, 5 mM $MgCl_2$, 5 mM DTT) supplemented with 0.2 mM $Fe(NH_4)_2(SO_4)_2$, 5 mM L-Cysteine-HCl, 5 mM ATP, and 0.02 g/ml DNase/RNase. The cell suspension was sonicated at 50 W, for 4 x 5 min for *E. coli* Bl21 ™ (DE3) Codon Plus-RIL cells or 2 x 5 min for Rossetta ™ (DE3) pLysS cells with a Branson sonifier (Branson Ultrasonics, Danbury, CT, USA) under cooling with ice-$H_2O$ in Rossetta flasks. Cell debris and membranes were removed by centrifugation 60 min at 100,000 x g at 18 °C. The clear supernatant was then loaded on 5 ml Streptactin macroprep™ columns. The columns were washed with 5 column bed volumes (CBV) buffer BC and bound target proteins were eluted with 3 CBV 2-2.5 mM Desthiobiotin in buffer BC.
The activating enzymes were purified accordingly in buffer A (100 mM Tris/HCl, pH 8.0, 150 mM NaCl, 5 mM $(NH_4)_2SO_4$, 5 mM $MgCl_2$, 50 mM arginine, 50 mM glutamate acid, 100 mM NaOH, 5 mM DTT). However, significant losses of recombinant AEs during the washing of the columns was observed and the washing volume was therefore reduced to 3 CBV in AE purifications.

The resulting preparations were concentrated in Vivaspin™ concentrators of appropriate cut-offs (100 kDa for decarboxylases, 10 kDa for activating enzymes) to final concentrations > 5 mg/ml and stored in suitable aliquots in gas tight sealed HPLC sample flasks at –80 °C.

**Determination of protein concentrations**

Relative protein concentrations of crude samples were determined using Bradford reagent with bovine serum albumin as standard [81]. The absolute concentration of pure enzymes was calculated from the absorbance at 280 nm in 4 M guanidinium hydrochloride in 0.1 % trifluoroacetic acid. Under these conditions, the iron sulfur clusters of the proteins were destroyed and did not interfere with the measurements. The molar absorbance coefficients and the equivalence values for 0.1 % [w/v] solutions are summarized in table 2.8.

**Table 2.8 The molar absorbtion coefficients and the absorbance values for 0.1% [w/v] solutions of the individual proteins**

| Enzyme | Extinction coefficient at 280 nm (units of $M^{-1}$ $cm^{-1}$) | Abs 0.1% (=1 g/l): |
|---|---|---|
| HpdA | 50070 | 1.306 |
| HpdBC | 135410 | 1.195 |
| CsdA | 56470 | 1.565 |
| CsdBC | 150060 | 1.329 |
| TfdA | 55190 | 1.413 |
| TfdBC | 162580 | 1.422 |
| CsdB-TfdC | 144370 | 1.278 |
| CsdB-HpdC | 144940 | 1.288 |
| HpdB-CsdC | 140530 | 1.237 |
| HpdB-TfdC | 134840 | 1.187 |
| TfdB-CsdC | 168270 | 1.472 |
| TfdB-HpdC | 163150 | 1.432 |

Therefore, the flow through of the spin concentrator and the concentrated enzyme were kept separately. To 0.5 ml guanidine HCl solution, 5 µl of the flow through were added, mixed and the $E_{280\,(ft)}$ was read. Then, 5µl of the purified protein solution was added and the value $E_{280\,(pr)}$ was measured. The final protein concentration was calculated according to $C_{pr}$= (($OD_{280}$

$_{(pr)}$*510/5)-(2*$OD_{280\ (ft)}$*505/5))/ε. The molar extinction coefficient ε was determined aided by the Biology Workbench (http://workbench.sdsc.edu/).

**Sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE)**

The polyacrylamide was cast as a separating gel topped by a stacking gel and secured in a BioRad electrophoresis apparatus, Mini Protein II, (Heidelberg, Germany). The solution for one gel was prepared as in table 2.9:

**Table 2.9 Protocols for SDS-PAGE gel preparation:**

| | Stacking gel | Separating gel | | | |
|---|---|---|---|---|---|
| | 6 % | 8 % | 10 % | 12 % | 15 % |
| 1 M Tris-Cl, pH 8.8 | 0 µl | 2250 µl | 3000 µl | 3000 µl | 3000 µl |
| 1 M Tris-Cl, pH 6.8 | 338 µl | 0 µl | 0 µl | 0 µl | 0 µl |
| $H_2O$ | 1950 µl | 1940 µl | 1796 µl | 1195 µl | 450 µl |
| Acrylamid/Bis-acrylamide 30 % ig/ 37 :1 | 565 µl | 1600 µl | 2407 µl | 2960 µl | 3660 µl |
| 10 % (w/v) SDS | 29 µl | 60 µl | 85 µl | 85 µl | 85 µl |
| 5 % TEMED | 29 µl | 60 µl | 57 µl | 69 µl | 85 µl |
| 10 % APS | 55 µl | 90 µl | 77 µl | 92 µl | 115 µl |

**Electrophoresis running buffer**

| | |
|---|---|
| 0.3 % (w/v) | Tris |
| 1.44 % (w/v) | Glycin |
| 0.1 % (w/v) | SDS |

**Protein loading buffer**

| | |
|---|---|
| 125 mM | Tris, pH 6.8 |
| 4 % (w/v) | Glycerol |
| 10 % (w/v) | Glycerin |
| 0.1 %(w/v) | Bromophenol blue |
| 10 mM | DTT (added just before use) |

**Staining solution for Coomasie Blue staining**

| | |
|---|---|
| 1.2 g | Coomassie R 250 |

| | |
|---|---|
| 200 ml | acetic acid |
| 1 L | Methanol |
| Fill to 2 L with $H_2O$ | |

**Destaining solution for Coomasie Blue staining**

| | |
|---|---|
| 50 ml | Ethanol |
| 71 ml | acetic acid |
| Add to 1 L with $H_2O$ | |

The molecular mass markers used in this study were the prestained ready-to-use (Bio-Rad) (the marker size is 250, 150, 100, 75, 50, 37, 25, 20, 15 and 10 kDa) or the LMW (Amersham Bioscience Europe GmbH, Freiburg), which was a mixture of phosphorylase b from rabbit muscle (97 kD); albumin from bovine serum (66 kD); albumin from chicken egg white (45 kD); carbonic anhydrase from bovine erytrocyte (30 kD); trypsin inhibitor from soybean (20.1 kD); α-lactalbumin from bovine milk (14.4 kD);
The protein samples were mixed with protein loading buffer (+ 10 mM DTT) and incubated at 95 °C for 5 min prior loading to the gel. The electrophoresis was performed at a constant voltage of 200 V till the tracking dye reached the bottom of the gel. Then the gel was disassembled and the proteins were stained with Coomasie Brilliant Blue.

**Determination of relative molecular masses of the native enzymes**

The native molecular masses for all proteins were determined by gel filtration on a Superdex 200 HR 10/30 prepacked column, equilibrated with 100 mM Tris-HCl, pH7.5, 150 mM NaCl. Ribonuclease A from bovine pancreas (13.7 kDa), chymotrypsinogen A from bovine pancreas (27 kDa), ovalbumin from hen egg (43 kDa), albumin from bovine serum (67 kDa), aldolase from rabbit muscle (158 kDa), ferritin from horse spleen (440 kDa) and thyroglobulin from bovine thyroid (669 kDa) were used as molecular mass standards (Amersham Biosciences, Germany). The native molecular masses were calculated from plots of $\log[M_w]$ vs the fractional exclusion coefficient $K_{av}$ of the individual samples and standards.

**Colorimetric determination of non-heme iron with Ferene**

| **Reagents:** | | |
|---|---|---|
| | 1.0 % (w/v) | HCl |
| | 7.5 % (w/v) | Ammonium acetate solution |
| | 2.5 % (w/v) | SDS |
| | 4.0 % (w/v) | Ascorbic acid (freshly prepared) |
| | 1.5 % (w/v) | Ferene (iron chelator) |

*Non-heme* iron from enzymes was quantified with Ferene (3-(2-pyridyl)-5,6-bis (5-sulfo-2-furyl)-1,2,4-triazine, disodium salt trihydrate) according to [83]. For iron standards preparation freshly prepared solutions of 0.2 mM $(NH_4)_2Fe(SO_4)_2$ x $6H_2O$ (Mohr's Salt) was used (0 to 20 µM $[Fe^{2+}]$ final). The $E_{593nm}$ was measured and the iron content was determined from absorbance plots vs. iron concentrations.

Diluted two protein solutions ($V_1$, $2 \times V_1$ µl) and 6 iron standards (10, 20, 40, 60, 80, 100 µl) and 100 µl with demineralized water were mixed in 1.5 ml Eppendorf tubes together with two blanks. 1 % HCl (100 µl) were added to each tube, mixed and incubated at 80 °C for 10 min. The vessels were kept closed till they were cooled down. Finally, 500 µl 7.5 % ammonium acetate, 100 µl 4 % ascorbic acid, 100 µl 2.5 % SDS, and 100 µl iron chelator were sequentially added with vortexing after each additon. The vessels were centrifuged at 13000 x g for 10min before measuring the absorbance of the solutions at 593 nm against water. The iron content of the samples was calculated from the calibration curve of the absorbance values vs. $[Fe^{2+}]$ of the standard solutions.

**Determination of acid labile sulfide**

| Reagents: | | |
|---|---|---|
| | 1.0 % (w/v) | Zinc acetate (freshly prepared) |
| | 7.0 % (w/v) | Sodium hydroxide |
| | 0.1 % (w/v) | N,N'-dimethyl-*p*-phenylene-diamine (DMPD) in 5 M HCl |
| | 4.0 % (w/v) | Ascorbic acid (freshly prepared) |
| | 10 mM | $FeCl_3$ in 1 M HCl |
| | 2 mM | Sulfide standard in 10 mM NaCl |

The acid labile sulfide was determined with N, N'-dimethyl-p-phenylenediamine (DMPD) [82]. Sulfide stock solutions (2 mM) were prepared from dried crystals of $Na_2S$ x 9 $H_2O$ in 10 mM NaOH under a $N_2$ inert atmosphere. Standards containing between 0 and 45 µM [$S^{2-}$] were prepared by appropriate dilutions of the stock solution. Standard plots of the absorbance vs. the sulfide concentration have been used to calculate the sulfide content of the protein.

Two protein samples ($V_1$ and 2 × $V_1$), two blanks, 5 sulfide standards (5, 10, 15, 20 and 25 µl) and two protein samples (v1) with sulfide (standard) additions of 10 and 20 µl were put in 1.5 ml Eppendorf tubes, and made up to 200 µl with distilled water. 600 µl Zinc acetate (1 %) was then added, and the sample was carefully mixed. After addition of 50 µl NaOH (7 % [w/v]), the vessels were closed and inverted to mix. The mixtures were then incubated for 15 min at room temperature, and centrifuged at low speed for several seconds. The following steps were done individually for all samples to minimize the loss of sulfide. DMPD solution (150 µL) was carefully addedd and mixed stirring the solution with the tip. When zinc hydroxide and sulfide precipitates were completely dissolved, $FeCl_3$ solution (150 µL) was rapidly added. The cups were immediately closed and vortexed vigorously for 30 sec. The samples were repeatedly mixed over a 20 min time period and finally centrifuged for 10min before the absorbance at 670 nm was measured against water.

**Iodimetric determination of sulfide standards**

The gravimetric preparation of a sulfide standards starting with solid $Na_2S$ x $9H_2O$ is inaccurate because of the hydroscopic nature of the compound, i.e. the sulfide concentration in the standard may be lower than anticipated, leading to the overestimation of the sulfide in the protein sample.

An accurate amount of iodine ($I_2$) is partly reduced with a known volume of sulfide standard solution. The remaining amount of $I_2$ is then determined by titration with sodium thiosulfate, and the sulfide concentration of the $Na_2S$ solution can be calculated by substraction.

In a 100 ml Erlenmeyer flask, 25 ml $H_2O$ and 5 ml iodine solution (0.5g iodine dissolved in 50 mL 300 mM KI) and 1 ml 1M sulfuric acid were mixed on a white sheet of paper using a magnetic stirrer. Sodium thiosulfate solution was carefully added from a burette filled with 5.6 g sodium thiosulfate dissolved in 500 mL water until the solution turns almost colorless (it should still be pale yellow). Then, 0.5 mL indicator solution (0.35 g soluble starch in 70 mL

water) was added. Thiosulfate solution was carfully dropped in until blue colour disappeared and the volume of sodium thiosulfate consumed was read from the burret. The titration is then repeated with a sample containing 25 mL of the sulfide standard. The determinations were repeated at least 3-times and average values were used to calculate the exact concentration of the sulfide standards.

Note the stoichiometry: $I_2 + 2\ S_2O_3^{2-} = 2I^- + S_4O_6^{2-}$

$I_2 + S^{2-} = S + 2I^-$

**Enzyme activity assays**

The enzyme activities were determined for the purified enzymes in bufferA (100 mM Tris-HCl pH7.5, 40 mM NaCl, 5 mM $MgCl_2$, 5 mM $(NH_4)_2SO_4$, 5 mM cysteine, 25 mM substrate, dithionite, and 5 mM DTT) supplemented with titanium citrate plus sodium sulfite or dithionite. The purified proteins of the Hpd-system were used as the positive controls. Assays omitting the SAM in the pre-activation were used as negative controls.

**Kinetic experiments and substrate specificity**

Substrate specificity and kinetic parameters of the recombinant enzymes were determined in time-resolved stop-time assays. In order to obtain comparable results, the enzyme used for individual determinations at various substrate concentrations or for different substrates derived from one activation mixture. Preliminary substrate specificity test employed HPA, 3,4-dihydroxyphenylacetate (3,4-DHPA), 4-aminophenylactetate (APA) and indole-3-acetate (IAA). Since only HPA and 3,4-DHPA were suitable substrates, the kinetic characterisation was restricted to these compounds.

Therefore, 0.2 mg/ml *BC* complex and 0.27 mg/ml activator were prereduced in bufferA supplemented with 2.5 mM Titanium-(III)-citrate and 2.5 mM $Na_2SO_3$ for 30 min at 30 °C. Then, SAM (0.3 mM final concentration) was added and incubated for 5 min. The decarboxylation reaction was started by mixing 20 µl of the activation mixture with 1ml anaerobic assay buffer (100 mM Tris-HCl pH7.5, 40 mM NaCl, 5 mM $MgCl_2$, 5 mM $(NH_4)_2SO_4$, cysteine, 1 mM dithionite, and 5 mM DTT) supplemented with the desired concentrations of various substrates. The reaction was allowed to preceed at 30 °C. At different time points, 100 µl aliquots were withdrawn, stopped with 100 µl 10% $HClO_4$ and neutralized with 100 µl 10% potassium carbonate containing 300 mM phenol as internal

standard. The samples were centrifuged at 13,000 rpm for 15 min to remove the denatured protein and potassium perchlorate, and analysed by reverse-phase HPLC.

**Detection of glycyl radical formation by oxygen-induced cleavage of the polypeptide**

The activation and assay (0.2 mg/ml BC complex, 0.27 mg/ml activator) were performed under strictly anaerobic conditions directly in the glovebox at 30 °C. The final activation volume was 200 µl. The activation solution mixture was prereduced for 30 min at 30 °C. Just prior addition of SAM, 10 µl of the non-acivated solution was withdrawn and mixed with 10 µl protein aerated loading buffer for SDS-PAGE. The activation was started by adding SAM (250µM). At various time points, 10 µl of the activated sample were withdrawn for subsequent detection of oxygen-induced cleavage by SDS-PAGE (8 % gel). Note that additional samples were analysed for activity in standard assay buffer.

**UV-visible measurements**

Purified CsdA, CsdBC, TfdA and TfdBC proteins were analysed with an HP 8453 UV-visible diode array spectrophotometer (USA). The samples were prepared under anoxic conditions and measured in rubber-stoppered quartz cuvettes. After the measurements of the as-purified proteins, a 10-molar excess of sodium dithionite was added and the sample was incubated for 1 h at the glove-box temperature before the measurement. Finally, the stopper was removed and the sample was oxidized with air for 1 h at room temperature with shaking.

**Reverse-phase HPLC separation of the aromatic compounds**

The analysis of the formation of *p*-cresol, 4-methylcatechol and skatole was performed using a LiChroCART$^{TM}$ 250 x 4 mm HPLC cartrisge (Merck, Darmstadt) filled with LiChrospher$^{TM}$ 100 RP-18 (5 µm) and operated at a flow rate of 1.2 ml·min$^{-1}$ at 50 °C. The eluent contained 0.1 % (v/v) trifluoroacetic acid in acetonitrile/water. The acetonitrile concentration was 30 %, 20 % and 50 % for the elution of *p*-cresol / 4-methylcatechol, 4-toluidine and skatole, respectively.

## EPR Spectroscopy

EPR spectra at X-band were obtained with a Bruker ESP-300E EPR spectrometer equiped with an ER-4116 dual mode cavity and an Oxford Instruments ESR-900 helium-flow cryostat with an ITC4 temperature controller.

### Sample preparation for EPR Spectroscopy

Various samples were analyzed by EPR spectroscopy. Suitable activity tests were performed in order to confirm functionality of enzymes. In an anaerobic glove box (atmosphere of 95 % $N_2$ and 5 % $H_2$), 300 µl of the reaction samples were transferred to EPR tubes. Anaerobic tubing was placed over the end of each EPR tube and closed. The EPR tubes were removed from the glove box and immediately frozen in liquid nitrogen prior analysis by EPR spectroscopy.

For the Fe-S cluster detection, the low temperature experiments were performed in liquid helium. All samples were prepared in parallel for the Hpd and Csd systems. Unless otherwise stated, EPR experiments were carried out in 100 mM Tris/HCl, pH 7.5, 25 mM NaCl, 5 mM $(NH_4)_2SO_4$, 5 mM $MgCl_2$, 5 mM DTT using 25 µM decarboxylases and 100 µM activator. When required, the samples were prereduced with 2 mM sodium dithionite for 30 min at 30 °C and subsequently activated with 0.25 mM SAM at 30 °C for 5 min prior to freezing in liquid nitrogen.

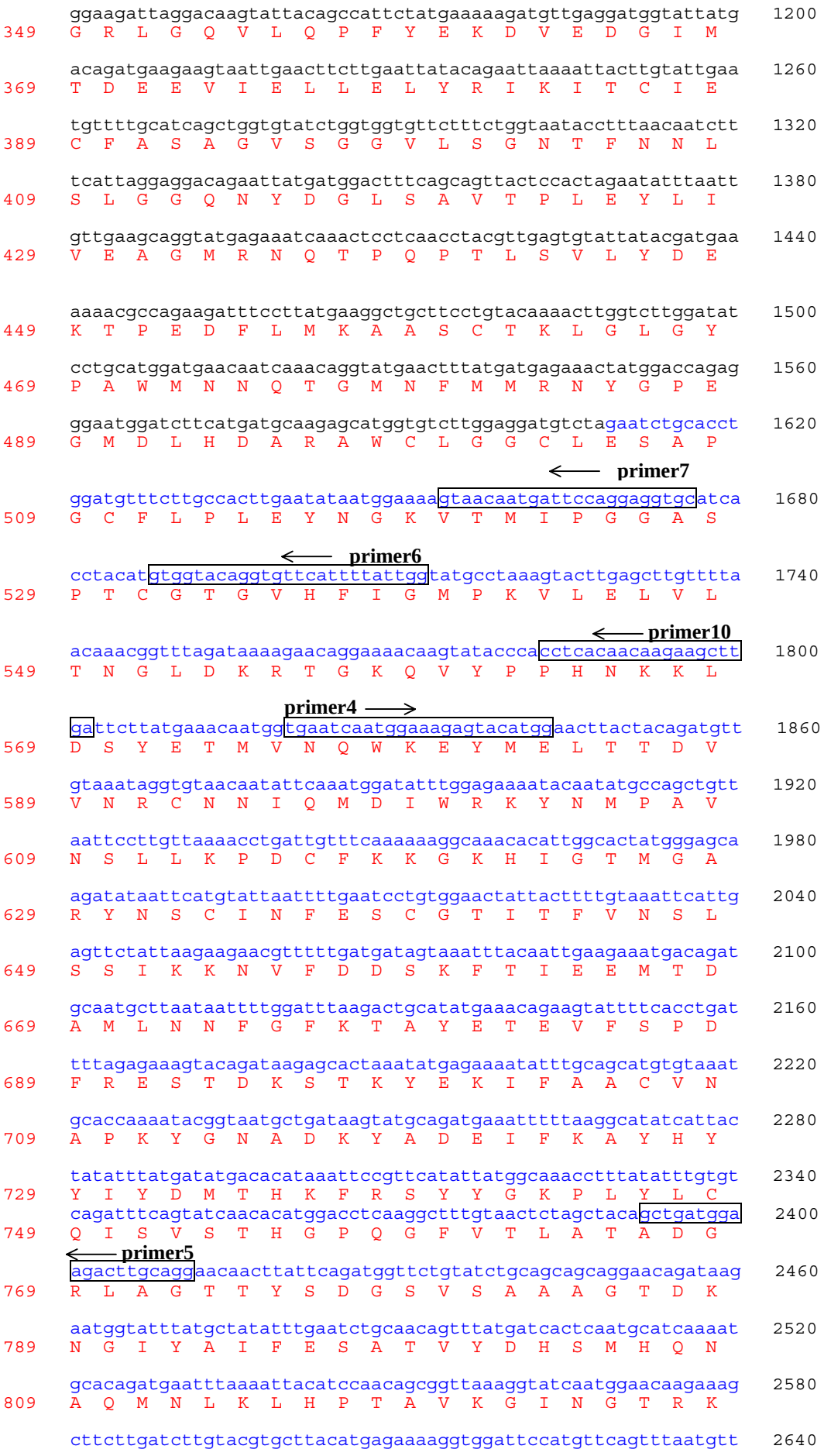
Additional X-band EPR spectra for the glycyl radical signal were recorded in liquid nitrogen. For deuterium-exchange experiments, a 20× buffer (2 M Tris/HCl, pH 7.5, 100 mM $MgCl_2$, 100 mM $(NH_4)_2SO_4$, 800 mM NaCl) was prepared anoxically in 99.9 % $D_2O$. The enzymes were 3 times rebuffered in this solution by repetetive concentration on a 10kD filter. Note that titanium(III)-citrate/sodium sulfite replaced sodium dithionite in these experiments. Experimental details and deviations from the general scheme given here are indicated in the figure legends.

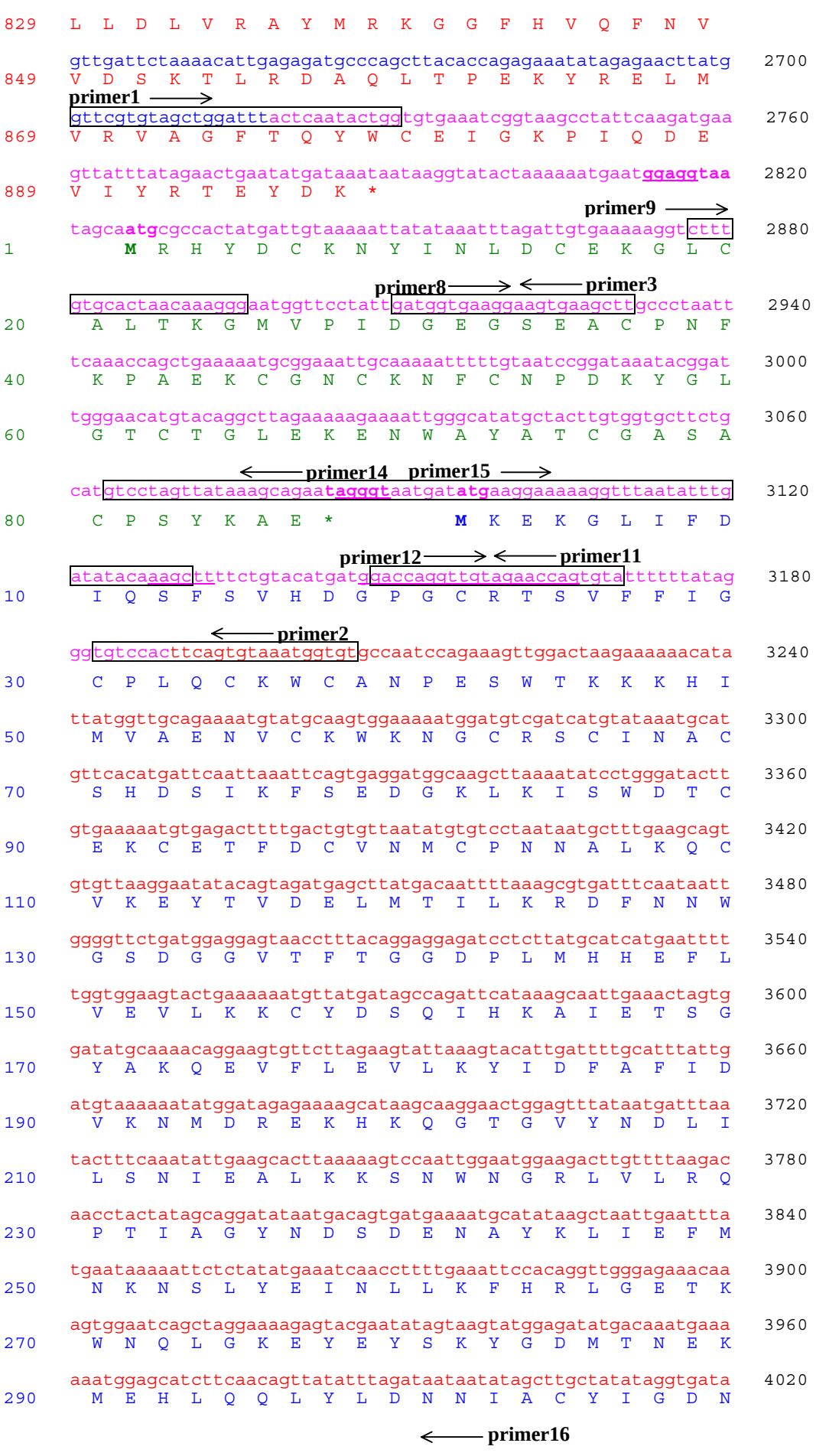
## Results

### Sequencing of the *csd*-locus of *Clostridium scatologenes*

While the genomic DNA sequences encoding the *hpd* locus of *Clostridium difficile* and the *tfd* locus of *Tannerella forsythensis* were known [12], no genetic information about the 4-hydroxyphenylacetate decarboxylase (*csd*) locus from *Clostridium scatologenes* was available at the start of this work. In this work, the Csd genes were sequenced by PCR based cloning strategies and the sequence is shown in figure 3.1.

```
      gctttgggaacacttagagtgttaaagggtgaagaagaagcaagagaatatgttgaataa     60

                                  primer13 ⟶
      tatgttttggaaaatattaaaaaggaggtatttttaatgaacgttaaagaaactaaacttt    120
1                                           M  N  V  K  E  T  K  L

      gaagatgttttaaaaagtcgtggtattgacatgaaagatgcttacaatatttcagaagca    180
9     E  D  V  L  K  S  R  G  I  D  M  K  D  A  Y  N  I  S  E  A

      gatattccagaagcaaaggaaagtactcaaaagcttatggatatctattatactttaaag    240
29    D  I  P  E  A  K  E  S  T  Q  K  L  M  D  I  Y  Y  T  L  K

      gttactgcagatatggaagctgcatattggtataacagaacttggtgggaaaatgatggt    300
49    V  T  A  D  M  E  A  A  Y  W  Y  N  R  T  W  W  E  N  D  G

      gaagtaattgaagtaagaagagctaaagctgtagcagcttctttatcacacatgacacct    360
69    E  V  I  E  V  R  R  A  K  A  V  A  A  S  L  S  H  M  T  P

      accattttaccttatgagaagcttgttatgaacaagacaaaaaatgttagaggtgcattt    420
89    T  I  L  P  Y  E  K  L  V  M  N  K  T  K  N  V  R  G  A  F

      ccttttccttgggtttgtgcaagtttttttaatgctcaagcagaagctttaatgaacgaa    480
109   P  F  P  W  V  C  A  S  F  F  N  A  Q  A  E  A  L  M  N  E

      gttgatgctccagcagagaatgaagctgactcagtaagtgttgttggtgctggtggaggt    540
129   V  D  A  P  A  E  N  E  A  D  S  V  S  V  V  G  A  G  G  G

      aatgttactgaaagttatggaaatgttatttctattgctaaaaagtttggtatgagaaaa    600
149   N  V  T  E  S  Y  G  N  V  I  S  I  A  K  K  F  G  M  R  K

      gaagaaattcctgttttagttaaaacctcaaagccttgggaaggaatttcagtagaagaa    660
169   E  E  I  P  V  L  V  K  T  S  K  P  W  E  G  I  S  V  E  E

      ttaagtaataaatactcaaagatgacaccaggatatgatcagtttaaaaatatcatggaa    720
189   L  S  N  K  Y  S  K  M  T  P  G  Y  D  Q  F  K  N  I  M  E

      agtgttatctgtatgtttgactcttttgcgattccacaggggcgtgaagtaattaactat    780
209   S  V  I  C  M  F  D  S  F  A  I  P  Q  G  R  E  V  I  N  Y

      tatatgcctttgcagtatggttttgatggaattatcaaattatgtgatgaaaaaattgct    840
229   Y  M  P  L  Q  Y  G  F  D  G  I  I  K  L  C  D  E  K  I  A

      gaagtcatgggtgaagctggtgacgatggtgattttggtatgagcagaggttattactat    900
249   E  V  M  G  E  A  G  D  D  G  D  F  G  M  S  R  G  Y  Y  Y

      gcagcaatgaaggaaattactaagggtttaagtgcatggtgtgaaaattattcaaagaga    960
269   A  A  M  K  E  I  T  K  G  L  S  A  W  C  E  N  Y  S  K  R

      gcaaaatatttagcttccattgaaacagattcagaaattaaagccaactatgaaaaaatt   1020
289   A  K  Y  L  A  S  I  E  T  D  S  E  I  K  A  N  Y  E  K  I

      gaagaagttatgggaaacattgctcataagaaacctgcaaacttctgggaagcaattcag   1080
309   E  E  V  M  G  N  I  A  H  K  K  P  A  N  F  W  E  A  I  Q

      atgacactttgctgtcattttggtgttgttaatgaagatcctcaatctggtctttcaatt   1140
329   M  T  L  C  C  H  F  G  V  V  N  E  D  P  Q  S  G  L  S  I
```


ggaagattaggacaagtattacagccattctatgaaaaagatgttgaggatggtattatg 1200
349 G R L G Q V L Q P F Y E K D V E D G I M
acagatgaagaagtaattgaacttcttgaattatacagaattaaaattacttgtattgaa 1260
369 T D E E V I E L L E L Y R I K I T C I E
tgttttgcatcagctggtgtatctggtggtgttctttctggtaatacctttaacaatctt 1320
389 C F A S A G V S G G V L S G N T F N N L
tcattaggaggacagaattatgatggactttcagcagttactccactagaatatttaatt 1380
409 S L G G Q N Y D G L S A V T P L E Y L I
gttgaagcaggtatgagaaatcaaactcctcaacctacgttgagtgtattatacgatgaa 1440
429 V E A G M R N Q T P Q P T L S V L Y D E
aaaacgccagaagatttccttatgaaggctgcttcctgtacaaaacttggtcttggatat 1500
449 K T P E D F L M K A A S C T K L G L G Y
cctgcatggatgaacaatcaaacaggtatgaactttatgatgagaaactatggaccagag 1560
469 P A W M N N Q T G M N F M M R N Y G P E
ggaatggatcttcatgatgcaagagcatggtgtcttggaggatgtctagaatctgcacct 1620
489 G M D L H D A R A W C L G G C L E S A P
primer7
ggatgtttcttgccacttgaatataatggaaaagtaacaatgattccaggaggtgcatca 1680
509 G C F L P L E Y N G K V T M I P G G A S
primer6
cctacatgtggtacaggtgttcatttttattggtatgcctaaagtacttgagcttgtttta 1740
529 P T C G T G V H F I G M P K V L E L V L
primer10
acaaacggtttagataaaagaacaggaaaacaagtataccacctcacaacaagaagctt 1800
549 T N G L D K R T G K Q V Y P P H N K K L
primer4
gattcttatgaaacaatggtgaatcaatggaaagagtacatggaacttactacagatgtt 1860
569 D S Y E T M V N Q W K E Y M E L T T D V
gtaaataggtgtaacaatattcaaatggatatttggagaaaatacaatatgccagctgtt 1920
589 V N R C N N I Q M D I W R K Y N M P A V
aattccttgttaaaacctgattgtttcaaaaaaggcaaacacattggcactatgggagca 1980
609 N S L L K P D C F K K G K H I G T M G A
agatataattcatgtattaattttgaatcctgtggaactattacttttgtaaattcattg 2040
629 R Y N S C I N F E S C G T I T F V N S L
agttctattaagaagaacgtttttgatgatagtaaatttacaattgaagaaatgacagat 2100
649 S S I K K N V F D D S K F T I E E M T D
gcaatgcttaataatttttggatttaagactgcatatgaaacagaagtattttcacctgat 2160
669 A M L N N F G F K T A Y E T E V F S P D
tttagagaaagtacagataagagcactaaatatgagaaaatatttgcagcatgtgtaaat 2220
689 F R E S T D K S T K Y E K I F A A C V N
gcaccaaaatacggtaatgctgataagtatgcagatgaaatttttaaggcatatcattac 2280
709 A P K Y G N A D K Y A D E I F K A Y H Y
tatatttatgatatgacacataaattccgttcatattatggcaaacctttatatttgtgt 2340
729 Y I Y D M T H K F R S Y Y G K P L Y L C
cagatttcagtatcaacacatggacctcaaggctttgtaactctagctacagctgatgga 2400
749 Q I S V S T H G P Q G F V T L A T A D G
primer5
agacttgcaggaacaacttattcagatggttctgtatctgcagcagcaggaacagataag 2460
769 R L A G T T Y S D G S V S A A A G T D K
aatggtatttatgctatatttgaatctgcaacagtttatgatcactcaatgcatcaaaat 2520
789 N G I Y A I F E S A T V Y D H S M H Q N
gcacagatgaatttaaaattacatccaacagcggttaaaggtatcaatggaacaagaaag 2580
809 A Q M N L K L H P T A V K G I N G T R K
cttcttgatcttgtacgtgcttacatgagaaaaggtggattccatgttcagtttaatgtt 2640


primer1
primer9
primer8
primer3
primer14
primer15
primer12
primer11
primer2
primer16

```
      atacgccttttaatcttatattttaatagagctgctattgaggta                4080
310      T   P   F   *
```

**Figure 3.1 The final sequence of the *csd*-locus from *C. scatologenes*.** The nucleotide numbers are given at the right, the amino acid numbering of the corresponding ORFs is shown left. The DNA sequences from different sequencing approaches are written in different colours. The positions of individual primers are indicated and the corresponding target sequences are boxed. Further details are given in the text.

The extraordinary high sequence identity and the unique genetic arrangement of the individual genes in the *hpd* and *tfd* loci suggested that related systems in *C. scatologenes* may exhibit a similar genetic arrangement. Based on this assumption, a pair of degenerated primers was deduced from the known genomic sequences of the *hpd* and *tfd* loci. A forward primer (primer1) was deduced, which covered the sequence encoding the glycyl radical motif at the C-terminal ends of the large subunits, while the reverse primer (primer2) was deduced based on the SAM radical cluster motif located close to the 5'-end of the AE genes (figure 3.2).

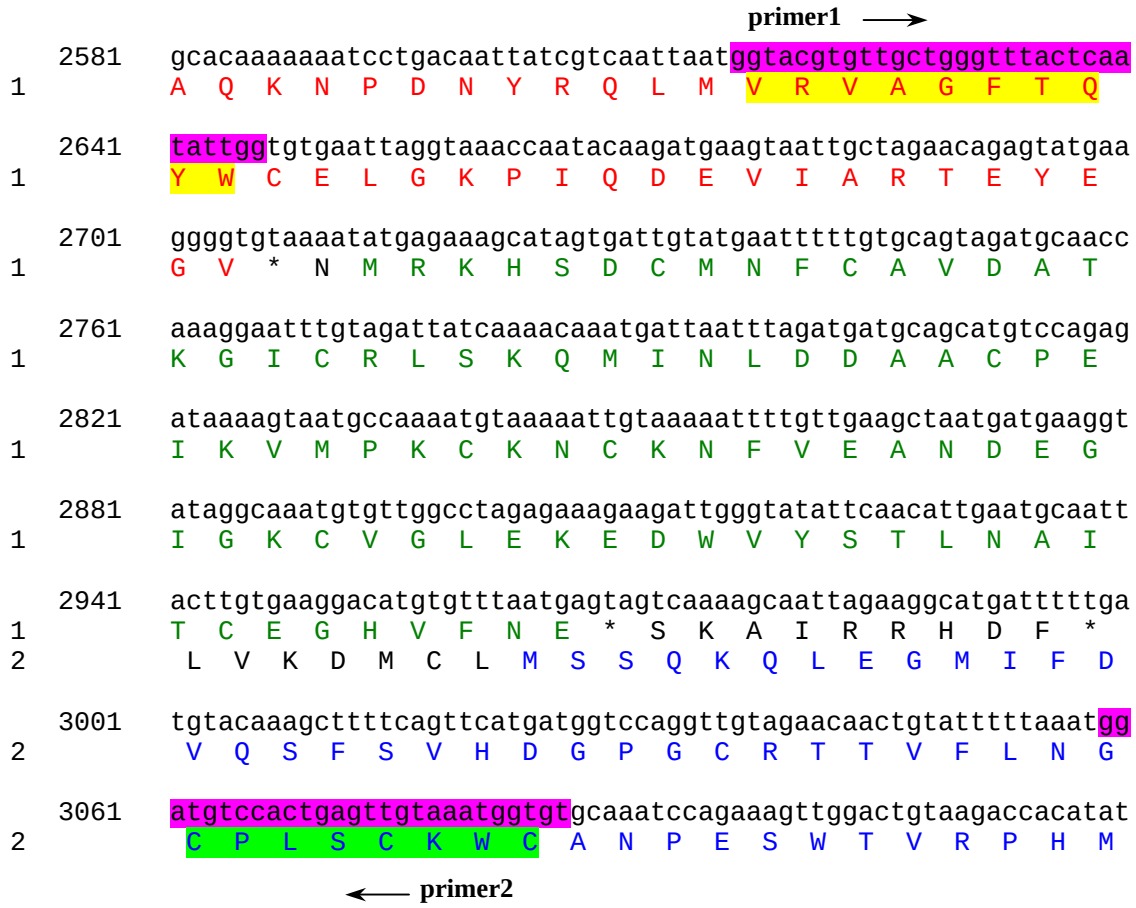



**Figure 3.2 Partial DNA sequence of the *hpd*-locus from *Clostridium difficile*.** The numbering is in aggreement with the genome-encoded sequence of strain 1296$^{T}$. The numbers preceding the amino acid sequences refer to the relative frame with the start-**atg** of hpdB as reference. The 3'-end of the *hpdB* gene (translated amino acid sequence in red) with the highly conserved glycyl radical motif (yellow) and the 5'-end of the *hpdA* gene (translated amino acid sequence in blue) with the SAM radical motif (light green) are shown. The degenerated primer sequences deduced are marked pink. Note that the insertion of the *hpdC* gene between these structural elements (green letters) significantly increases the size of the intergenic region as compared to other GRE loci.

The unique insertion of a small subunit gene in GRE systems between these strictly conserved elements was predicted to yield significant larger PCR products from the proposed decarboxylase system in *C. scatologenes* (>500 bp) than expected for other GRE systems (< 300 bp), which were likely to be amplified, too. A 500 bp PCR product (Figure3.3) thus obtained was subsequently ligated into a pCR2.1-vector using the TA-cloning kit. White-blue screening of the obtained colonies enabled the identification of insert-containing clones from which the plasmid DNA was purified. The insert was sequenced using the vector-derived M13 sense and antisense primers. It covered the 3'-region of a glycyl radical enzyme gene and the 5'-region of its cognate AE gene. More important, an ORF encoding an 86 amino acid protein similar to the small *hpdC* and *tfdC* genes was found (figure 3.1, pink).

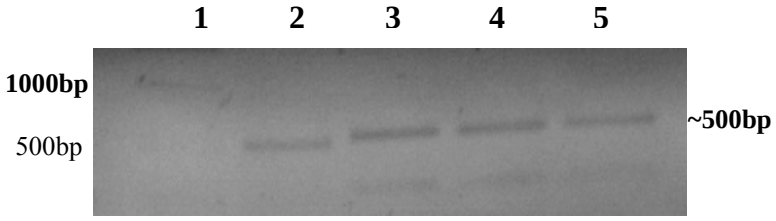



**Figure 3.3 Agarose gel electrophoresis of PCR products amplified with the primer pair 1/ 2 from genomic DNAs.** Lane 1: 500bp-ladder from Roche. Lane 2: positive control with genomic DNA from *C. difficile*. Lane 3-5: Amplified products with genomic DNA from *C. scatologenes* as template. Note that also faint bands below 300 bp are visible.

The ~ 500 bp PCR products were subsequently used as template to synthesize a digoxigenin-labelled probes for a screening of endonuclease digests of genomic DNA from *C. scatologenes* in Southern blots. As shown in Figure 3.4, the digests with *EcoRI*, *XbaI*, *HindIII* and *EcoRV* yielded products of suitable sizes. From preparative agarose gels, size restricted fragments of ~ 6 kb (*XbaI*), ~ 3.6 kb (*EcoRV*) and ~ 4.9 kb (*HindIII*) were recovered from the individual digests and subsequently ligated into pBluescript vectors cut with an appropriate enzyme. The circular vectors thus obtained were not used to transform *E. coli* cells but severed as templates in four individual PCR reactions. The primers in these PCRs were the possible combinations of two vector derived M13-primers with

specific primers deduced from the insert sequence. The results showed that only from the *XbaI* cut fragements, a PCR product was obtained with primer 3/M13-rev.

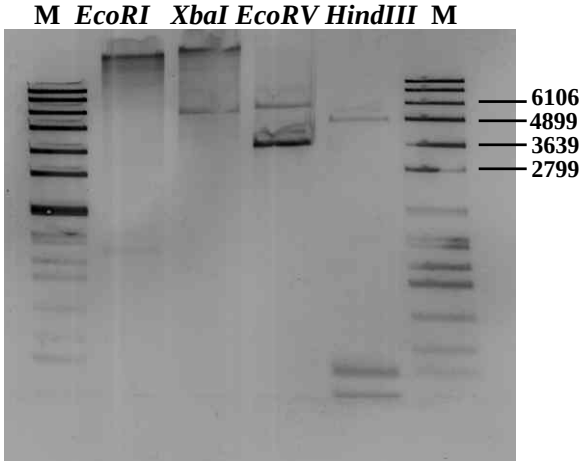



**Figure 3.4 Southern blot analysis of restriction digests of *C. scatologenes* genomic DNA using a primer pair 1/2 derived probe (probe 1).** The genomic DNA (~3.2µg) was transferred onto nylon membrane and hybridized with the probe. M: DIG-labelled DNA molecular weight marker III (Roche Applied Science).

An about 1.6 kb long PCR product was subsequently ligated in pCR2.1 and sequenced (Figure 3.5). The PCR product contained a 1,200 bp long 5'-extention in addition to the already know partial sequence of the *csd* locus and encoded the C-terminal half of the *csdB* gene (Figure 3.1, blue).

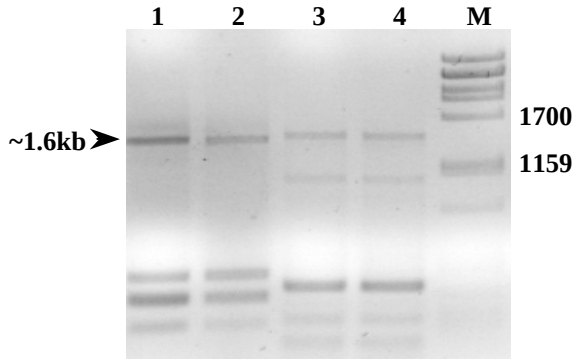



Figure 3.5 Aagarose gel electrophoresis of PCR products obtained with size restricted *XbaI*-fragments ligated in pBluescriptII SK$^+$. Lane 1 and Lane 2: (M13-20/primer 3), Lane

3 and Lane 4: (M13-as/primer 8), M: *PstI*-cut Lambda-DNA. Note that only the 1.6 kb fragment was amplified by all four primer pairs, which was cut out from the gel and sequenced.

Based on the novel sequence, new primers were deduced, which matched the 5'-region of this extension (primer4/primer5), and subsequently used for the synthesis of a new digoxygenin-labelled probe (probe2). This probe was used in another approach of Southern blotting with a different subset of restriction endonucleases (Figure 3.6).

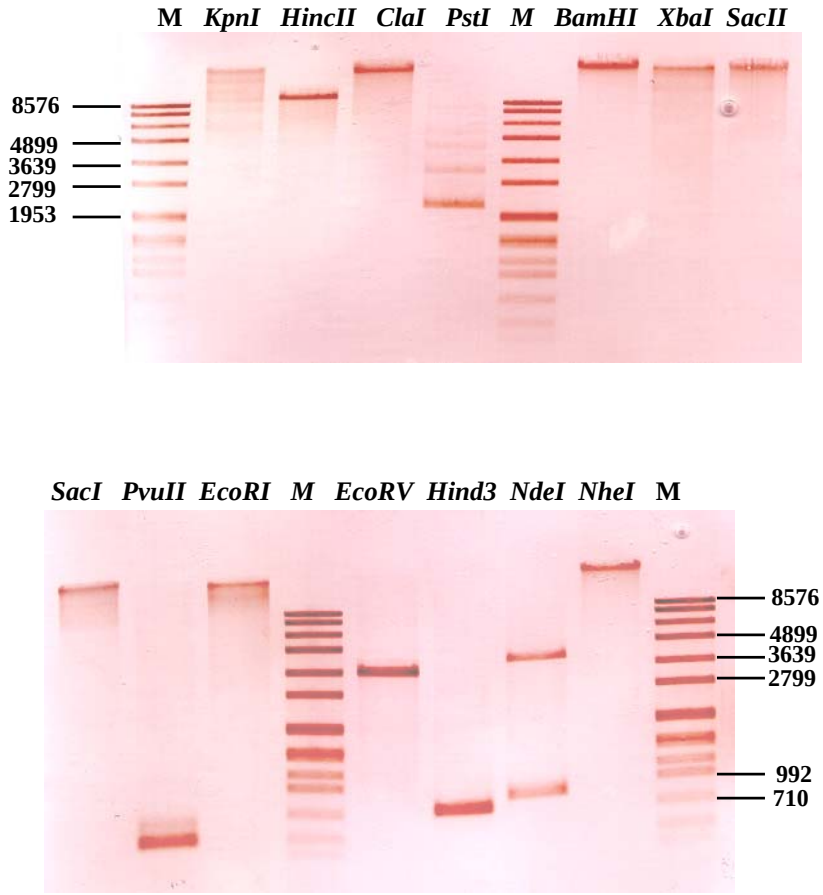



**Figure 3.6 Southern blot analysis of endonuclease digested genomic DNA from *C. scatologenes*.** A specific probe (probe2) was synthesized with the primer pair 4/5 from genomic DNA of *C. scatologenes*. M: DIG-labelled DNA molecular weight marker III (Roche Applied Science).

As shown in Figure 3.6, several of these digest yielded fragments of appropriate sizes, which were subsequently used in two different experimental approaches in order to obtain the missing parts of the *csd* locus.

The digest of genomic DNA with *PstI* generated fragments with a 3'-overhang and was therefore suitable to be used in the TOPO Walker kit (Invitrogen). The fragments were first dephosphorylated and then subjected to primer extension. The products thus obtained were subsequently linked to the LinkAmp primer 1 provided by the TOPO Walker kit and subsequently used as a template in series of individual PCRs with this primer in combination with insert-specific primer: primer6, primer7, primer8 and primer9. As indicated in Figure 3.7, a 1.7 kb fragment was obtained from gene specific primers 6 and 7, which was subsequently cloned into pCR2.1 and sequenced. This approach yielded a further extension of the known sequence of the *csd* locus towards the start codon of the *csdB* gene by about 1200 bp (Figure 3.1, black).

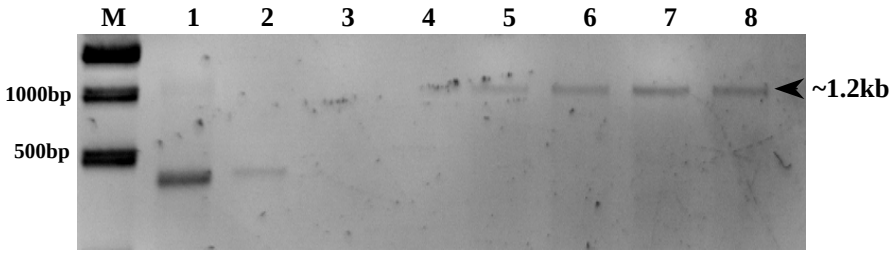



**Figure 3.7 TOPO® Walker amplified PCR products. M: 500bp-ladder DNA marker from Roche.** Lane 1-4: amplifications by the gene specific primer 8/ 9 at different annealing temperature (50°C and 52°C). Lane 5-8 were from the amplifications by using primer 6/primer 7 and different extension times (1 min and 2 min). The ~1.2kb amplified fragment was then cut out and sequenced.

The nucleotide sequences of the 5'-end of the *csdB* gene and of the majority of the *csdA* gene derived from inverse PCRs with *NdeI* or *HindIII* digests, respectively. Therefore, the restriction enzymes were removed by chloroform-phenol extraction and individual fragments were then self-ligated in diluted solutions in order to allow preferential circularisation rather than recombination. The circularised DNA was subsequently used as template for PCR amplification with two divergent primers, which were deduced to locate

in the close vicinity of *NdeI* or *HindIII* cleavage sites in the already known *csd* sequence (Figure 3.1).

The iPCR product (Figure 3.8) obtained for the *csdB* part of the locus was about 2 kb in size, subsequently ligated into pCR2.1 and sequenced. The majority of its sequence covered already known parts of the *csdB* gene. However, it also contained an about 95 bp long 5'-extension which contained the start codon of the *csdB* gene and a small part of the 5'-UTR of the *csd* locus (Figure 3.1, brown).

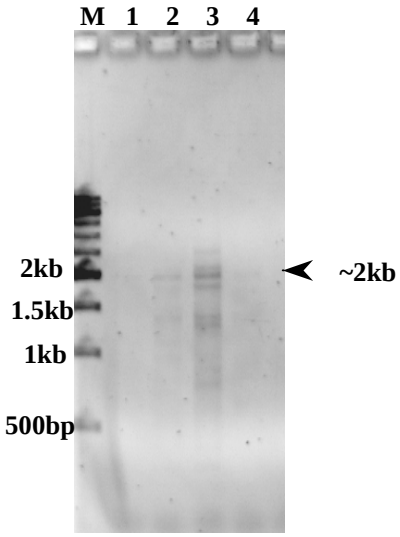



**Figure 3.8 Inverse PCR (iPCR) of the 5'-end of *csdB*.** The template for the iPCR was circularized *NdeI*-cut genomic DNA from *C. scatologenes* with the primer pair 10/4. Circularised *NdeI*-cut genomic DNA (at 1 ng/µL) was amplified with the primer pair 10 and 4. After ethanol/chloroform extraction, 5 µl (lane 1- lane 4) and 20 µl (lane 5- lane 8) were used as template for the iPCR amplification, respectively. The fragment at ~2kb was cut out from gel and sequenced. More details are given in text.

IPCR amplification of circularised *HindIII* cut DNA with the primer pair 11/12 produced a ~ 1.4 kb fragment (Figure 3.9), which contained the full sequence of the *csdA* gene plus a small part of the 3'-UTR region (Figure 3.1, red).

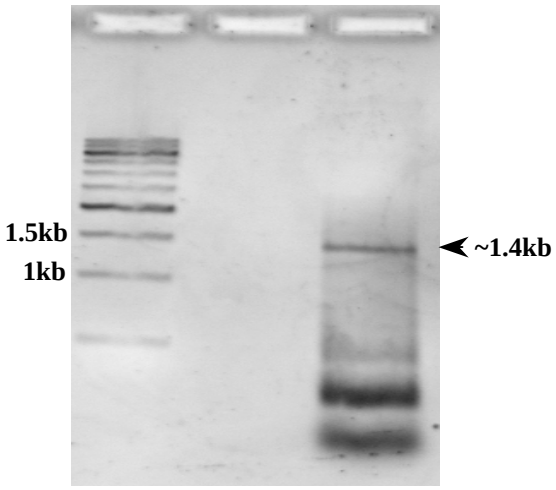



**Figure 3.9 Gel electrophoresis of the iPCR product for the activating enzyme gene.** The fragment used was circularized *HindIII*-cut genomic DNA of *C. scatologenes*. The individual reaction conditions were as described in figure 3.8. The ~1.4kb PCR product was cut from the gel and sequenced.

**Sequence analysis of the individual glycyl radical decarboxylase systems**

At the start of this work, the *hpd* system from *C. difficile* was the only known glycyl radical decarboxylase, although a very similar system was found in the unfinished genome of *T. forsythensis* [11]. The sequencing of the *csd* locus from *C. scatologenes* increased the number of these systems and might allow the definition of some unique features of this novel sub-class of GREs.

The genetic arrangements of the individual genes within the *hpd* locus of *C. difficile* and of the *csd* locus of *C. scatologenes* were very similar, while the *tfd* locus of *T. forsythensis* differred significantly. All three genes encoding both decarboxylase subunits and the AE were directly neighboured, and most likely form a transcriptional unit in the *csd* and *hpd* systems. However, the *tfd* locus differs from *C. difficile* and *C. scatologenes* in that one gene (*tfdD*) was located between *tfd*C and *tfd*A, which encodes a 424 amino acid long polypeptide, which is predicted to belong to the major facilitator protein superfamily [83, 84]. Since these proteins usually act as secondary carriers, which use electrochemical gradients to thrive active transport of small molecules across the cytoplasma membrane, TfdD might act in substrate uptake or product secretion. Interestingly, the *hpd* locus of *C. difficile* lacks an orthologue system but contained a putative membrane protein downstream of the AE gene (*hpdD*). The latter was predicted to contain 8 trans-membrane-helixes (TMHs) and showed similarities to the poorly characterized abortive infection proteins [85] and CAAX prenyl endopeptidases [86]. Remarkably, this protein family also contains determinants of a signal transducing pathway involved in the plantaricin A induced bacteriocin production in *Lactobacillus plantarum* [87], suggesting a similar function of HpdD in the regulation of *p*-cresol formation by *C. difficile*.

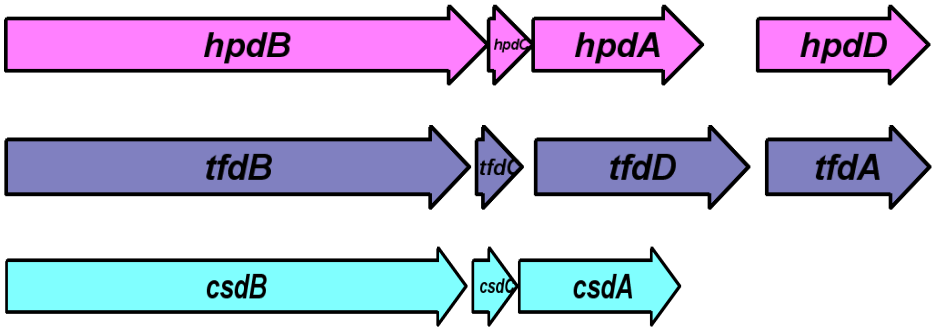



**Figure 3.10 Comparison of the genetic arrangements among *hpd*, *tfd* and *csd* systems.**

Table 3.1: Statistical analysis of the amino acid sequences

| | Amino acids: | Mass: | IEP: |
|---|---|---|---|
| Activators: | | | |
| HpdA | 316 | 35.902 kDa | 4.86 |
| CsdA | 319 | 36.722 kDa | 6.27 |
| TfdA | 317 | 36.673 kDa | 5.59 |
| Large subunits: | | | |
| HpdB | 902 | 101.276 kDa | 5.22 |
| CsdB | 897 | 100.637 kDa | 5.24 |
| TfdB | 903 | 102.012 kDa | 5.36 |
| Small subunits: | | | |
| HpdC | 85 | 9.504 kDa | 5.40 |
| CsdC | 86 | 9.396 kDa | 7.65 (6.69) [1)] |
| TfdC | 87 | 9.936 kDa | 9.67 (8.33) [1)] |

[1)]The number in brackets refers to the theoretical values considering the contribution cysteinyl thiols.

Irrespectively of these differences in the genetic arrangment, the individually encoded polypetides were very similar (table 3.1). As shown in figure 3.11, in particular the glycyl radical subunits of all three systems were extremly similar. In a multiple sequence alignment, more than 50 % of the amino acid positions were identical in all three sequences and this identical value reaches almost 60 % in pairwise alignments. This fact complicates a meaningful functional annotation of individual amino acids and the more distantly related GRE enzymes Bss, Gdh, Pfl were introduced in order to identify the most relevant amino acids. The typical GRE fingerprint motif is readily recognised. Moreover, the putative active site cysteinyl residue is also readily assigned (C507 in HpdB, C502 in CsdB, C508 in TfdB). However, it should be noted that all three arylacetate decarboxylases share seven strictly conserved cysteinyl residues. Noteworthy, one of these cysteines (C620 in HpdB) was also strictly conserved in Bss, Gdh and several GREs of unknown function derived from genome projects.

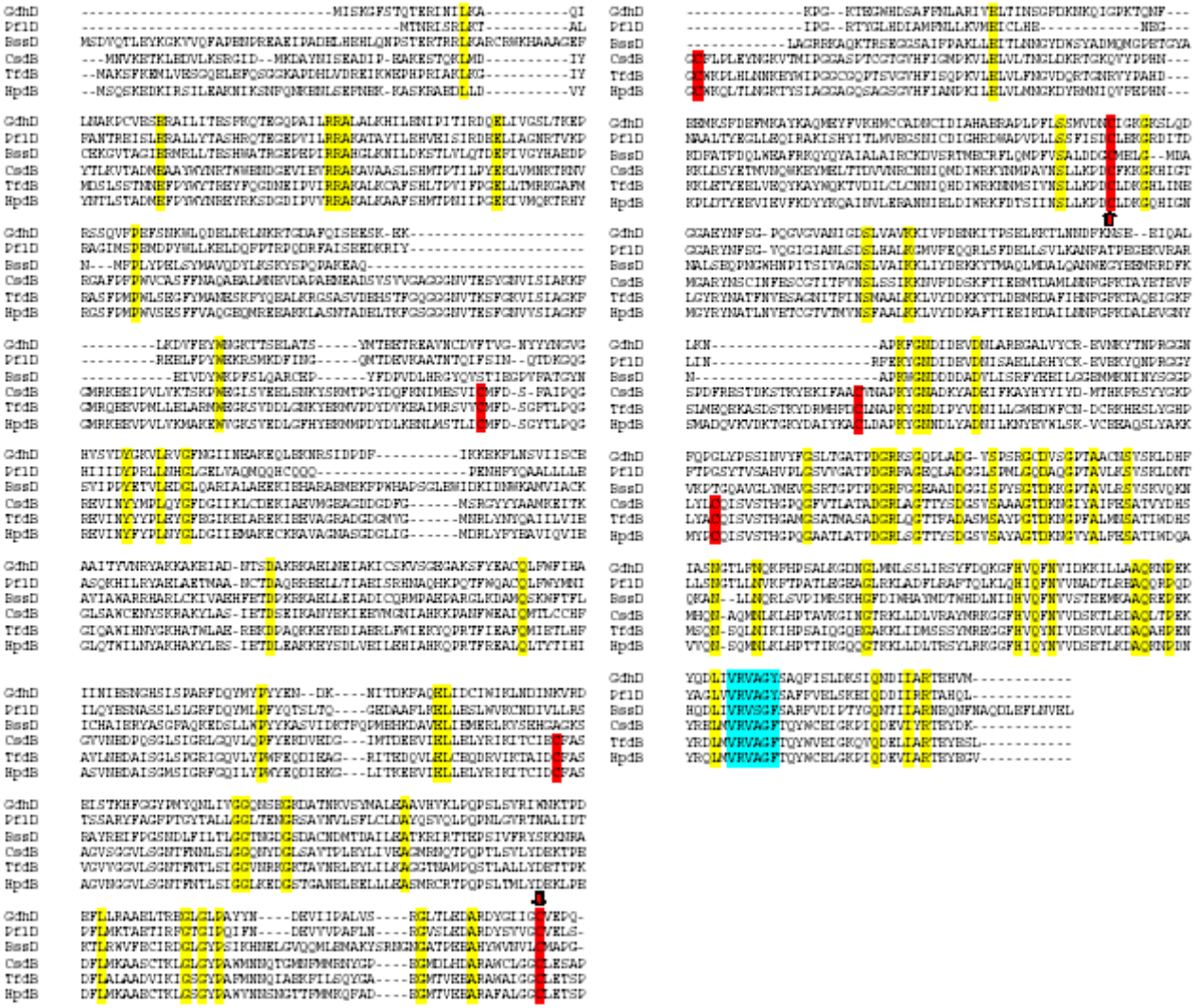


**Figure 3.11 Amino acid sequence alignment of CsdB from *C. scatologenes* with TfdB from *T. forsythensis*, HpdB from *C. difficile*, PflD from *E.coli*, BssD from *T. aromatica* and GhdD from *C.butyricum*.** The compleletely conserved amino acid seqnences are yellow boxed. The glycyl radical motif is shown in light blue and the 7 conserved cysteines of CsdB, TfdB and HpdB are red. The cysteines conserved in all glycyl radical enzymes are marked with arrows.

It has been previously suggested that the small subunit distinguishes Hpd from all other GREs. The finding of this polypeptide in both Csd and Tfd further supported this suggestion. The sequence identity between the small subunits was significantly lower (25 %) and the amino acid composition of these polypeptides differed strongly as indicated by the very different isoelectric points (table 3.1). Comparing these proteins, four cysteinyl residues were strictly

conserved. These residues were embedded in three individual motifs of reasonable conserved amino acids, which were separated by short variable amino acid sequences (figure 3.12).

```
CsdC    ---MRHYDCKNYINLDCEKGLCALTKGMVPIDGEGREASPNFKPAEKCGNCKNF-CNPDK
HpdC    --MRKHSDCMNFCAVDATKGICRLSKQMINLDD---AACPEIKVMPKCKNCKNF-VEAND
TfdC    MEERIYKNSLNYIPIDVAKGIDRRTGKRVNADD----VDPNYERMPKCMHCVNFTLNQEK

CsdC    YGLGTCTGLEKENWAYATCGASACPSYKAE--
HpdC    EGIGKCVGLEKEDWVYSTLNAITCEGHVFNE-
TfdC    IGLGICR-MGKEFIAYPDMAAVTCTGYKERGA
```

**Figure 3.12 Sequence alignment of the small subunits.** Strictly conserved residues are shown in red, identical residues between two individual subunits are boxed light blue.

The three decarboxylase AEs shared an overall sequence identity of about 40 % (figure 3.13). Interestingly, these enzymes contained an about 61 amino acids long insert between the SAM radical motif and the SGG stretch, which are common for all SAM radical enzymes [55]. Within these inserts, 8 cysteinyl residues are strictly conserved among all decarboxylase AEs. Noteworthy, similar inserts are also found in all GRE-AEs with the remarkable exception of Pfl-AE and Nrd-AE [33].

```
CsdA          --MKEK-----GLIFDIQSFSVHDGPGCRTSGPGCRTSVFFIGCPLQCKWCANPESWTKK
HpdA          --MSSQ-KQLEGMIFDVQSFSVHDGP-------GCRTTVFLNGCPLSCKWCANPESWTVR
TfdA          MLMENKEKEVKGLIFSIQSYSVHDGP-------GTRTTVFLNGCPLMCKWCCNPEGLFRY
BssD          --MKIP------LITEIQRFSLQDGP-------GIRTTIFLKGCPLRCPWCHNPETQDAR
PflA          --MSVIG-----RIHSFESCGTVDGP-------GIRFITFFQGCLMRCLYCHNRDTWDTH

CsdA          KHIMVAENVCKWKNGCRSCINACSHDSIKFSE----DGKLKISWDTCEKCETFDCVNMCP
HpdA          PHMMFSELSCQYENGCTVCHGKCKNGALSFNL----DNKPVIDWNICKDCESFECVNSCY
TfdA          PVMLHSHAKCK---KCGACIQVCPYHAISIGE----DHEPIFDRSICDTCTTIECVEACL
BssD          QEFYFYPDRCVG---CGRCVAVCPAETSRLVRNSDGRTIVQIDRTNCQRCMR--CVAACL
PflA          G-------------------------------------------------------

CsdA          NNALKQCVKEYTVDELMTILKRDFNNWG-SDGGVTFTGGDPLMHHEFLVEVLKKCYDSQI
HpdA          YNAFKLCAKPYTVDELVQVIKRDSNNWR-SNGGVTFSGGEPLLQHEFLHEVLLKCHEVNV
TfdA          HEGNSISGKYYTIDELIHRLDRDRPYWG-DHGGVTFSGGEPLLQKDFILPMLMACKEHYM
BssD          TEARAIVGQHMSVDEILREALSDSAFYRNSGGGVTISGGDPLYFPDFTRQLASELHARGV
PflA          -------GKEVTVEDLMKEVVTYRHFMNASGGGVTASGGEAILQAEFVRDWFRACKKEGI
```

**Figure 3.13 Alignment of the N-terminal amino acid sequences of glycyl radical enzyme activating enzymes (GRE-AE).** The N-terminal half of the amino acid sequences of HpdA, CsdA and TfdA are compared with the sequences of PflA from *E. coli* and BssD from *Thauera aromatica*. The SAM-binding motif is boxed. The completely conserved amino acid sequences are shadowed in yellow. The cysteins of the SAM-cluster are red lettered. The 8 conserved cysteinyl residues within an approximately 60 amino acids long insert as compared

to PflA are shadowed in red. Note that only about 1/3 of the amino acid sequences of the individual AEs is shown.

**Cloning and expression of the *csd* and *tfd* genes**

Expression clones for the *hpdBC* genes and for the *hpdA* gene in pASK-IBA7 vectors have been previously made and were kindly provided by Dr. Paula I. Andrei for this work.
Similar expression plasmids were generated for the recombinant production of CsdA, TfdA, CsdBC and CsdA. Therefore, specific primers were deduced based on the genomic DNA sequences of the *csd* and *tfd* genes. The sense primers for the PCR amplification of the individual genes were deduced including the start ATG of the genes and carried a 5'-mutagenic extension, which provided a *SacII*-cleavage site for the N-terminal fusion with the Streptag II amino acid sequence provided by pASK-IBA7 vector. In order to allow a directed cloning of the genes, the corresponding antisense primers contained a mutagenic 3'-extension generating a *BamHI*-site downstream of the stop codons.
Based on the experience with the Hpd system, the AE genes were cloned separately from the decarboxylase genes, which were co-amplified and introduced together in the same vector. The genes were amplified with proofreading DNA polymerases in order to reduce the number of PCR-induced errors. The PCR products thus obtained were cut with *SacII*/*BamHI*, purified and subsequently ligated into pASK-IBA7 cut with the same enzymes. The target plasmids were amplified in either DH5α or Bl21 (DE3) cells, purified and double stranded sequenced. In order to avoid PCR errors, three individual clones derived from three individual PCRs were sequenced. However, the *tfd*-genes were intermediately cloned into pBluescript, sequenced and subsequently subcloned via *SacII*/*BamHI* into pASK-IBA7.
For the *tfd* genes additional expression clones in pPR-IBA2 were generated. While the pASK-IBA7 vector is designed for the expression of genes under control of the Tet-respressor and induced with anhydrotetracycline, the pPR-IBA2 expression is under control of the T7 promotor and induced by IPTG. Subcloning of the genes in the pASK-IBA7 vectors using the *NheI* and *BamHI* sites of both vectors generated these clones.
In order to elucidate the possible function of the small subunits, expression plasmids were generated in pASK-IBA7, which allowed the co-expression of the individual glycyl radical subunits with all three small subunits individually. The *tfd* genes contained a natural *ClaI* cleavage site between the stop codon of *tfdB* and the ribosomal binding site (RBS) of *tfdC*. Since neither the *csdBC* nor the *hpdBC* genes contained natural *ClaI* site, such site was introduced in these systems by mutagenic PCR. This was achieved in a two-step procedures:

initially, PCR-products of glycyl radical subunit genes were generated, which carried a tandem of a *ClaI* and a *BamHI* sites downstream of the stop codon of *hpdB* or *csdB*, respectively. Note that these clones also allowed the expression of N-terminally *Strep*-tagged HpdB and CsdB alone. Subsequently, a *ClaI* site was generated upstream of the RBS of either *hpdC* or *csdC* with mutagenic primers. The *BamHI* site downstream of the *hpdC* or *csdC* stop condons was retained in this PCRs and the cloning of these fragments into the expression plasmids of the cognate glycyl radical subunit genes with the previously introduced *ClaI* sites regenerated the wildtype combinations of genes (designated pASK-IBA7-*hpdB(ClaI)C* and pASK-IBA7-*csdB(ClaI)C*, respectively), which were subsequently sequenced. Finally, the exchange of *ClaI*/*BamHI*-fragments between the parental wildtype plasmids generated all nine possible combinations of genes in expression plasmids. The resulting plasmids were used to various cell lines for expression studies. Initial experiments revealed, which of the generated expression strains yielded production of soluble proteins (table 3.2).

**Table 3.2 Conditions for expression of individual genes in *E. coli***

| Proteins | Strains* | | | T(°C) | | | Inducer | | | | | | | Atmosphere | | Results |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BL | T | R | 37 | 28 | RT | AHT(μg/L) | | IPTG(mM) | | | | | -O2 | +O2 | |
| | | | | | | | 50 | 100 | 0.2 | 0.4 | 0.6 | 0.8 | 1.0 | | | |
| pASK-IBA7-TfdBC | √ | | | √ | √ | | | √ | | | | | | | √ | soluble |
| | | √ | | √ | √ | | | √ | | | | | | | √ | soluble |
| | | | √ | √ | √ | | | √ | | | | | | | √ | soluble |
| pASK-IBA7-TfdA | √ | | | | | √ | √ | | | | | | | √ | √ | Partially soluble |
| | | √ | | | | √ | √ | | | | | | | √ | √ | Partially soluble |
| | | | √ | | | √ | √ | | | | | | | √ | √ | Partially soluble |
| pPR-IBA2-TfdBC | √ | | | | √ | | | | √ | √ | √ | √ | √ | | | soluble |
| | | √ | | | √ | | | | | √ | | | | | | soluble |
| | | | √ | | √ | | | | | √ | | | | | | soluble |
| pPR-IBA2-TfdA | √ | | | | | √ | | | √ | √ | √ | | | √ | √ | Partially soluble |
| | | √ | | | | √ | | | √ | | | | | √ | √ | Partially soluble |
| | | | √ | | | √ | | | √ | | | | | √ | √ | Partially soluble |

| Proteins | Strain | T (°C) | $C_{AHT}$ (µg/L) | Atmosphere | Ethanol Stress | Induction time (h) | result |
|---|---|---|---|---|---|---|---|
| CsdA | R* | 22-25 | 50 | + $O_2$ | 2% | 3 | soluble |
| CsdBC | R* | 28-30 | 100 | + $O_2$ | 2% | 3 | soluble |
| HpdB-CsdC | R* | 28-30 | 100 | + $O_2$ | 2% | 3 | soluble |
| HpdB-TfdC | R* | 28-30 | 100 | + $O_2$ | 2% | 3 | soluble |
| TfdB-CsdC | R* | 28-30 | 100 | + $O_2$ | 2% | 3 | insoluble |
| TfdB-HpdC | R* | 28-30 | 100 | + $O_2$ | 2% | 3 | soluble |
| CsdB-TfdC | R* | 28-30 | 100 | + $O_2$ | 2% | 3 | insoluble |
| CsdB-HpdC | R* | 28-30 | 100 | + $O_2$ | 2% | 3 | insoluble |

*: the strains used were BL21™ (DE3) Codon Plus RIL (BL), Rossetta™ (DE3) pLysS ® (R) and Tuner™ (DE3) pLysS (T).

The investigations addressing the solubility of recombinant HpdA and HpdBC in *E. coli* hosts have shown that a carefully controlled growth of the cells was required in order to obtain preparations for high yield purification of recombinant proteins. The crucial parameters identified were the particular host strain, the growth temperature and the maintainance of exponential growth throughout the production phase. Unless otherwise stated, the recombinant proteins were produced in Rosetta-pLysS cells, which were cultivated in aerated media. The AEs were produced from cells growing at 25-28 °C, while the decarboxylases were produced at 28-30 °C in LBG medium supplemented with 50 µg/mL carbenicillin and 34 µg/mL chloramphenicol. Colonies from freshly transformed cells were used for the inocculation and kept growing exponentially until an $OD_{578nm}$ of about 0.6 in 1.5 L of medium was reached. At this point, 2 % ethanol was added (over-expression of chaperons that often improve the yields of recombinant soluble proteins). After additional 30 min of growth, the appropriate inducer was added together with Fe(III)-citrate and L-cysteine (yielding final concentrations of 50 µM and 5 mM, respectively). The cells were harvested 3 to 4 h post-induction, washed with TBS and stored frozen at – 80 °C.

**Protein purification**

The heterologue expression of genes from pASK-IBA7-derived vectors let to the production of recombinant proteins, which were N-terminally fused to a *Strep*-tag for affinity purification

on *Strep*-Tactin columns. The recombinant AEs were purified at pH 8.0 and the decarboxylases at pH 7.5 in 100 mM Tris-HCl-buffers supplemented with 150 mM sodium chloride and 5 mM each ammonium sulfate, dithiothreitol and magnesium chloride under strictly anoxic conditions provided by an anoxic glove box.

The frozen cells were resuspended in the appropriate buffer supplemented with 5 mM ATP, 5 mM L-cysteine and 200 µM ammonium-Fe(II)-sulfate (Mohr's salt). After addition of DNase/RNase (10 µg/mL each), the cells were oppened by sonication. The crude extracts were supplemented with 40 µg avidin per gram of wet packed *E. coli* cells and subcequently centrifuged at 100,000x g for 1 h at 18 °C. Noteworthy, in particular the activator preparation showed an intensive "browning" during this step, indicating that an iron-sulfur-cluster reconstitution took place.

The clear extract was loaded on 5 mL *Strep*-Tactin macroprep columns (IBA, Göttingen) equilibrated with the appropriate purification buffer. Since the binding affinity of the recombinant AEs was substancially lower than for the decarboxylases, the washing with purification buffer was restricted to 3 column volumes for purifications of AEs, while 5 column volumes were used for the decarboxylases. The bound proteins were then eluted with purification buffer supplemented with 2 to 2.5 mM desthiobiotin and subsequently concentrated to yield final concentrations > 5 mg/mL with Vivaspin® centrifuge membrane concentrators of the appropiated cut-off. The final preparation was was either directly used in subsequent experiments or split in suitable aliquots and frozen at – 80 °C. Individual aliquots were thawn only once.

As shown in figure 3.14, 3.15 and 3.16, the resulting proteins were pure according to SDS-PAGE and Coomassie staining after affinity purification. As expexted, all decarboxylases that have been purified were composed of both, the putative glycyl radical and the small subunits, while the activator consisted of only one polypeptide.

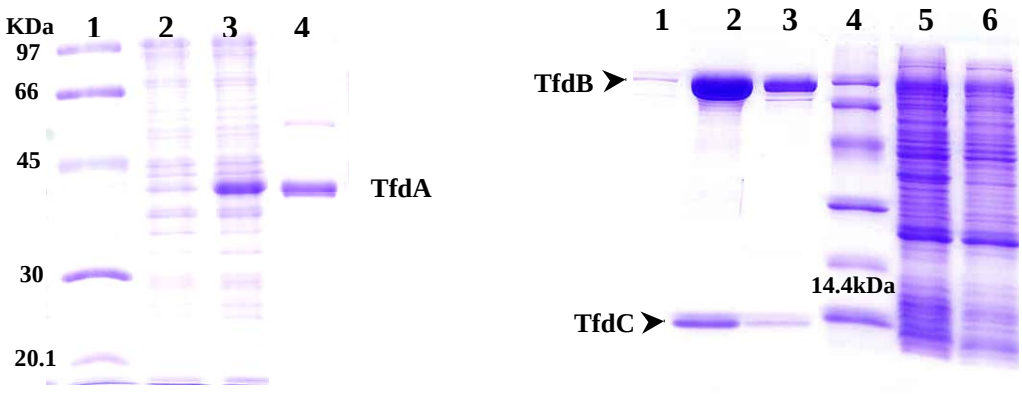



**Figure 3.14 SDS-PAGE of purified recombinant Tfd proteins.** The gel was stained with Coomassie Brilliant Blue. Left (12% gel) 1: molecular weight marker (Amersham). 2: uninduced; 3: induced; 4: purified TfdA. Right (15% gel), 1-3: successive fractions of purified TfdBC. 4: molecular weight marker. 5: induced; 6: uninduced TfdBC.

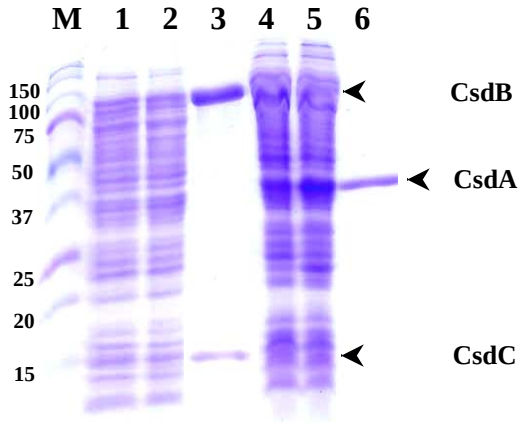



**Figure 3.15 SDS-PAGE of purified recombinant Csd proteins.** M: molecular weight marker (Precision Plus Protein ™ Unstained Standards, Bio-Rad). 1: uninduced; 2: induced; 3: purified CsdBC. 4: uninduced; 5: induced; 6: purified CsdA.

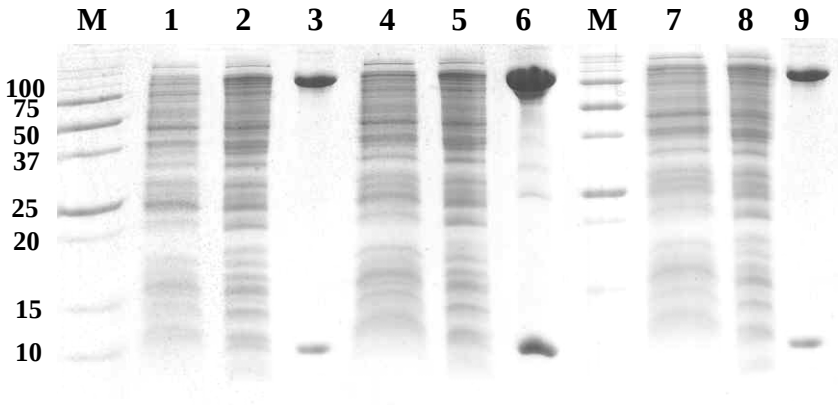



**Figure 3.16 SDS-PAGE of purified hybrid decarboxylases.** The 15% gel was stained with Coomassie Brilliant Blue. M: molecular weight marker (Precision Plus Protein™ unstained standards, Bio-Rad). 1: uninduced; 2: induced; 3: purified HpdB-TfdC. 4: uninduced; 5: induced; 6: purified TfdB-HpdC. 7: uninduced; 8: induced; 9: purified HpdB-CsdC.

**Physical and chemical characterisation of the recombinant proteins**

Interestingly, all the recombinant proteins exhibited a clearly visible brownish colour, indicating the presence of iron-sulfur centres in both, decarboxylases and AE. This notion was further supported by UV/Vis spectroscopies (figure 3.17) and by chemical determinations of non-heme iron and acid-labile sulfur.

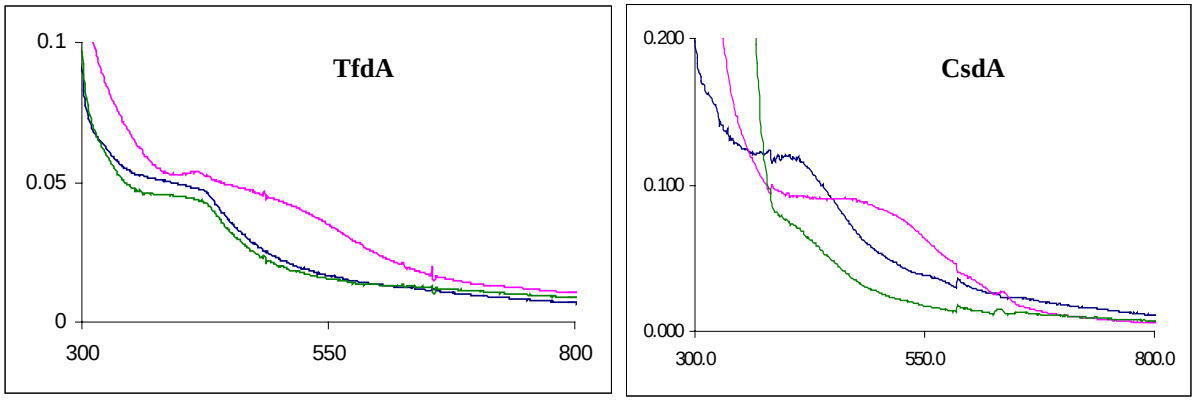



**Figure 3.17 The UV-vis absorption spectra of purified CsdA and TfdA.** Green: reduced; Pink: oxidised; Blue: as purified. The purified protein samples (5 µM) were in 100 mM Tris-HCl buffer, pH 7.5 with 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $(NH_4)_2SO_4$ and 5 mM DTT. The reduction was done with 10-fold molar excess of sodium dithionite in the glove box for 1h. Afterwards, the sample was oxidised under air for 1h with occasunally shaking.

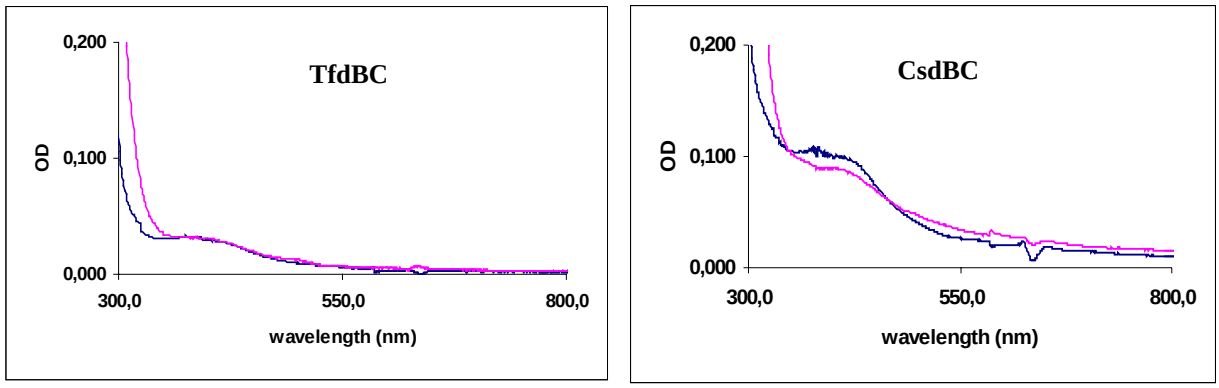



**Figure 3.18 The UV-vis absorption spectra of as-purified (blue line) and dithionite reduced CsdBC/ TfdBC (pink line).** The purified protein samples (3 µM for TfdBC; 5 µM for CsdBC) were in 100 mM Tris-HCl buffer, pH 7.5 supplemented with 150 mM NaCl, 5 mM $MgCl_2$, 5 mM $(NH_4)_2SO_4$ and 5 mM DTT. After recording the spectra, 10x molar excess dithionite was added and the reduction was performed anaerobically for 1 h.

The UV/Vis spectroscopies clearly showed broad charge-transfer bands indicative for iron-sulfur centres. Redox-properties and the lack of visible fine structure in the spectra suggested that these centres are most likely [4Fe4S]-centers. The observed $\varepsilon_{400nm}$ of above 30 $mM^{-1}cm^{-1}$ for the AEs and of about 14 $mM^{-1}cm^{-1}$ for individual decarboxylases indicated the presence of two metal centres in AE monomers and of one metal centre in the decarboxylase hetero-dimers. The stoichiometry of 8 irons and 7-8 acid-labile sulfurs in the AEs, and of >4 irons and >4 sulfurs in the decarboxylases, which was determined chemically, further supported this view.

The subunit compositions of the individual decarboxylases were determined by RP-HPLC and size exclusion chromatographies. Therefore, the glycyl radical (B) and the small (C) subunits were separated on AquaporeRP300_Butyl columns (Baker, 2.1 x 100 mm, 7µm). The individual subunits were baseline-separated and the intergrated signal intensities at 280 nm were corrected for the very different $\varepsilon_{280nm}$ in order to establish the molar ratio of the individual polypeptides in the preparations (data not shown). Irrespectively of the particular decarboxylases, the molar ratios of both subunits were always very close to 1.0 (range: 0.98 to 1.04), indicating an equimolar stoichiometry of both subunits in all decarboxylase polypeptides, which yielded soluble protein, including the hybrid decarboxylases.

The native molecular masses of the individual decarboxylases and of the AEs were determined by size exclusion chromatographies in purification buffer on Superdex200 (HR 1.0 / 10) columns (Pharmacia, Freiburg), which was calibrated with appropriate molecular mass standard proteins. The elution volumes indicated that all AEs were monomeric enzymes with native molecular masses of 35-40 kDa. In contrast, the molecular masses of the three wild-type decarboxylases and of the hybride decarboxylases varied significantly. While both HpdBC and CsdBC exhibited molecular masses of about 440-450 kDa, TfdBC was significantly smaller (220 kDa). Interestingly, the HpdB-derived hybrids with CsdC and TfdC exhibited also lower native molecular masses of 220 kDa, while the TfdB-HpdC hybride was even smaller (110 kDa). Taken together, these data suggested that CsdBC and HpdBC are hetero-octameric proteins ($\beta_4\gamma_4$), while TfdBC, HpdB-CsdC and HpdB-TfdC are hetero-tetramers ($\beta_2\gamma_2$) and TfdB-HpdC is a hetero-dimer ($\beta\gamma$).

Unpublished work of Dr. Paula I. Andrei suggested that the glycyl radical subunit of recombinant HpdBC was serine-phosphorylated. The recombinant proteins obtained throughout this work were, therefore, also analysed by Western blots for the occurence of this modification by Dr. T. Selmer. These studies revealed that the glycyl radical subunits

are serine phosphorylated in all the recombinant enzymes. The results of the physico-chemical characterisation of the recombinant proteins are briefly summarized in table 3.3.

**Table 3.3 The results of the physico-chemical characterisation of the recombinant proteins**

| Enzyme | Mw (kDa) | Oligomeric state | Fe and Scontent | phosphorylation | solubility |
|---|---|---|---|---|---|
| HpdA | 35-40 | $\alpha$ | 7-8Fe/monomer<br>6-7S/monomer | no | soluble |
| HpdBC | 440-450 | $\beta_4\gamma_4$ | 5-8Fe/heterodimer<br>5-8S/heterodimer | yes | soluble |
| CsdA | 35-40 | $\alpha$ | 7-8Fe/monomer<br>6-7S/monomer | no | soluble |
| CsdBC | 440-450 | $\beta_4\gamma_4$ | 5-8Fe/heterodimer<br>5-8S/heterodimer | yes | soluble |
| TfdBC | 220 | $\beta_2\gamma_2$ | 4Fe/heterodimer | yes | soluble |
| TfdA | 35-40 | $\alpha$ | 7-8Fe/monomer<br>6-7S/monomer | no | Partial soluble |
| HpdB-CsdC | 220 | $\beta_2\gamma_2$ | 4Fe/heterodimer<br>4S/heterodimer | yes | soluble |
| HpdB-TfdC | 220 | $\beta_2\gamma_2$ | 4Fe/heterodimer<br>4S/heterodimer | yes | soluble |
| TfdB-CsdC | - | - | - | - | insoluble |
| TfdB-HpdC | 110 | $\beta\gamma$ | 4Fe/heterodimer<br>4S/heterodimer | yes | soluble |
| CsdB-TfdC | - | - | - | - | insoluble |
| CsdB-HpdC | - | - | - | - | insoluble |

## Functional characterisation of the individual systems *in vitro*

A proper understanding of the experiments, which were carried out with the recombinant enzymes relies on the knowledge of recent observations in the Hpd system from *C. difficile*. This system has been previously studied by Dr. Paula I. Andrei and is currently investigated by Martin Blaser. With kind permission of these colleagues, the most important but yet unpublished findings are briefly summarized here.

The characterisation of the Hpd systems was initially hampered by the observation that the post-translational activation of the decarboxylase precursor by its recombinant activator proceeded smoothly in cell-free extract of recombinant *E. coli* cells, but was complicated with the purified enzymes. In particular, the maximum specific activities observed with pure enzymes were only about 10-20 % of the estimated activity in cell-free extracts [12]. However, it has been only recently found that these difficulties most likely arose from the unexpected fact that the AE does not only catalyse the SAM-dependent activation of the decarboxylase precursor, but also causes a subsequently time-dependent quenching of the glycyl radical. Therefore, the experiments performed with the recombinant enzymes essentially attempted an understanding of the processes involved in this property, which is apparently found only in glycyl radical decarboxylases, but not in the better studied Pfl and Nrd systems.

## Functionality and substrate specificity of the recombinant decarboxylases

The characterisation of the Csd and Tfd systems was initially complicated by the unknown substrate specificity of both systems. While no biochemical or physiological information concerning the organism's ability to decarboxylate arylacetates is available for *T. forsythensis*, *C. scatologenes* has been described to produce *p*-cresol from tyrosine and scatole from tryptophane. Therefore, the Csd system could either encode a 4-hydroxyphenylacetate or an indol-3-acetate decarboxylase, or alternatively provide a bi-functional enzyme catalysing both reactions. In order to circumvent the problems, which have been encountered with the *C. difficile* Hpd, initial testing for functionality and substrate specificity was carried out in cell-free extracts and later confirmed with the purified enzymes.

In order to test for the ability of the recombinant decarboxylases to decarboxylate various arylacetates, the individual compounds were disolved at 25 mM final concentration in solution A (100 mM Tris-HCl, pH 7.5, 5 mM each DTT, $(NH_4)_2SO_4$, $MgCl_2$ and cysteine)

supplemented with 1 mM sodium dithionite. About 100 µg total protein from freshly prepared cell-free extracts from *E. coli* cells producing either CsdBC, CsdA, TfdBC or TfdA and positive controls from the Hpd system were added to the test mixtures and pre-reduced for 30 min at 30 °C. Then, SAM (250 µM) was added and the anticipated product formation was monitored time-dependently by RP-HPLC. Individual tests ommitting the designated AEs or SAM, and combinations of the individual decarboxylases with the AEs of the other systems were additionally performed.

The individual decarboxylases were tested with 4-hydroxyphenylacetate, indole-3-acetate, 3,4-dihydroxyphenylacetate and 4-aminophenylactetate as substrates and assayed by RP-HPLC for the corresponding products *p*-cresol, skatole, 4-methylcatechol and *p*-toluidine, respectively. Irrespectively of the substrates used, no activity was observed for the Tfd system or for any of the hybrid decarboxylases. However, both the Hpd control and the Csd system exhibited very similar substrate specificity. Both recombinant enzymes showed a time-dependent decarboxylation of 4-hydroxyphenylacetate and 3,4-dihydroxyphenylacetate to yield *p*-cresol and 4-methylcatechol, respectively. These data suggested that the *csd* locus of *C. scatologenes* encodes a 4-hydroxyphenylacetate decarboxylase.

It turned out that the individual decarboxylases and their designated AEs were freely exchangable between Hpd and Csd, while no cross reactivity was observed with the Tfd system. No activity was found with the decarboxylase containing cell-free extracts alone, while the addition of SAM was not essentially required in cell-free extracts. From these data it might be concluded that the decarboxylase activation was strictly dependent on the dedicated AE, while the cell-free extracts provided sufficient SAM to allow certain activation in the absence of additionally added SAM.

These data were perfectly reproducible with purified recombinant enzymes. However, when the pure decarboxylases (2 µg/mL) were activated with recombinant activator (2.7 µg/mL), time- and strictly SAM-dependence was observed. As shown in figure 3.19, the time-course of such activation experiments showed clearly a rapid activation of CsdBC/HpdBC by CsdA/HpdA. Once the full activation was achieved, the cresol concentration in the test increases linear in time, sugessting specific activities of more than 20 U/mg for the decarboxylation of HPA by CsdBC. Similar specific activities were also observed for the activated HpdBC.

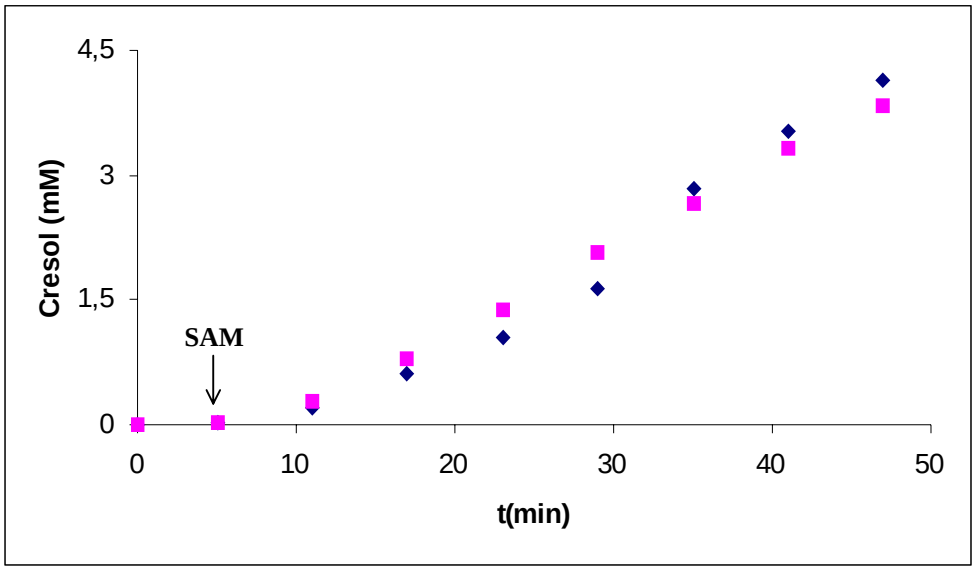



**Figure 3.19 The time-course of cresol formation by CsdBC (blue)/HpdBC (pink).** Activation of the decarboxylase precursors by either CsdA or HpdA was started by addition of SAM.

## Kinetic studies

### Transient activation

It has been recently shown for the Hpd system, that the activation of HpdBC by HpdA also takes place in the absence of substrate [12]. However, in such pre-activation experiments the observation was recently made, that the maximum specific activity of the decarboxylase was rapidly reached, but that the activity declined thoughout extended periods of incubation (M. Blaser, unpublished). As shown in figure 3.20, the activity loss was apparently accompanied by a loss of the glycyl radical in HpdBC. Immediately after the addition of SAM, the glycyl radical was rapidly formed as indicated by the oxygen-induced cleavage of HpdB, which was monitored by SDS-PAGE (figure 3.20 B). However, it was rather obvious that this band diminished in the course of the experiment, suggesting that the radical transiently formed is subsequently quenched. The reactions involved in this process are currently in the focus of our research.

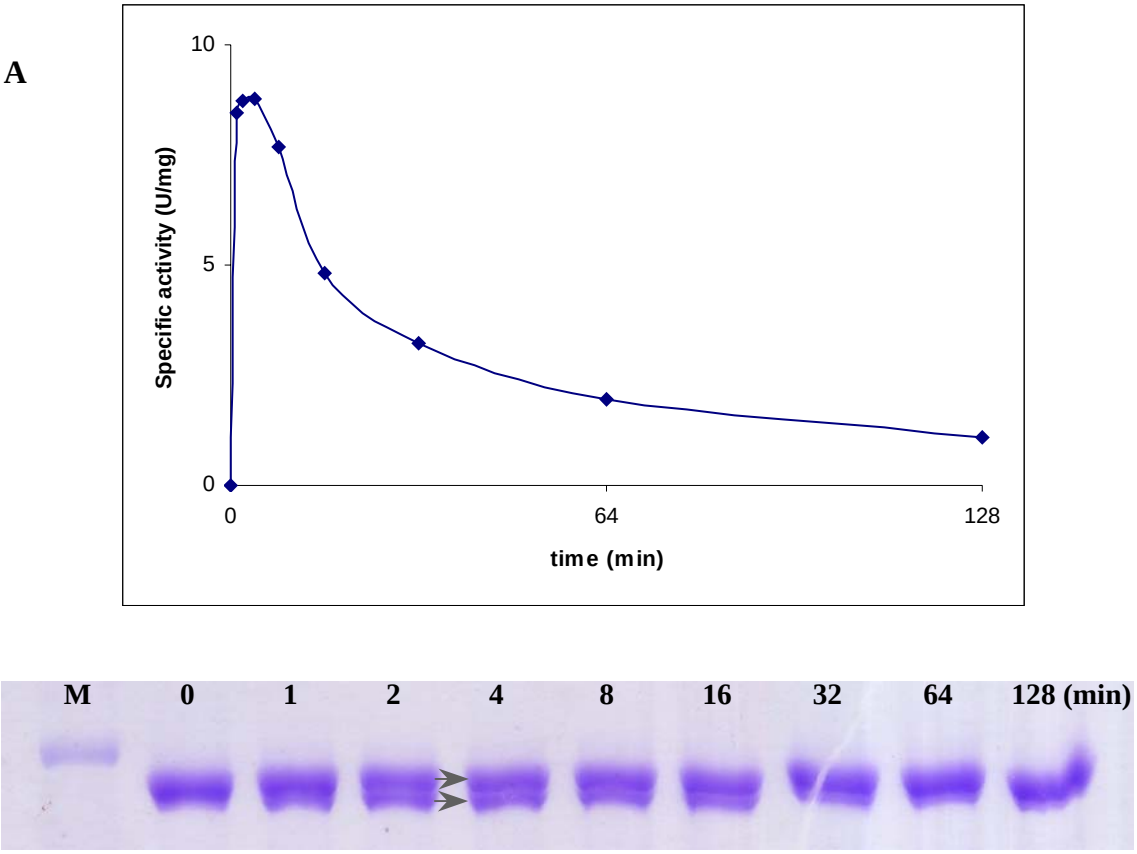



**Figure3.20 The transient activation of HpdBC.** A: Specific activity versus time. B: the $O_2$-induced cleavage monitored by SDS-PAGE. Note that the slightly smaller protein band showing up and dimishing is due to the C-terminal truncation of the polypeptide by

oxygen-induced cleavage at the glycyl radical site. M: Precision Plus Prestained ™ Marker (Bio-Rad). 0 was the negative control, which was the activation mixture without SAM, 1 to 128 is the time intervals (minutes) post-addition. Note that the first time point for activity measurements was taken immediately after addition of SAM.

For the Csd system, the transient activation was also observed, but due to the very poor resolution of cleaved and non-cleaved CsdB, EPR spectroscopy was used instead to follow the radical quenching process, which will be shown later.

**Different A/BC ratio activation**

In order to study the activation process of the Csd system, different amounts of CsdA were used to activate recombinant decarboxylase precursor. Therefore, the concentration of CsdBC was kept constant (1.7 μM) and the A/BC ratio was varied from 0.125 to 8. After 30 min of pre-reduction at 30 °C and samples were withdrawn for activity tests and analysis of oxygen-induced cleavage by SDS-PAGE. Then, SAM (250 μM) was added and additional sample pairs were taken at various time points to monitor activity and glycyl radical formation.

As shown in figure 3.21, the activation of CsdBC by CsdA was rapidly achieved with a molar excess ≥ 1 of CsdA vs. CsdBC. Within 5 to 10 min, maximum specific activities of about 20 U/mg CsdBC were reached. However, this rapid activation was followed by a much slower decline in the specific activity during extended incubation period. Although the transient formation of the glycyl radical was visible in Coomassie-stained SDS-PAGE, the poor resolution of the precursor and the radical-derived, oxygen-cleaved degradation product did not reproduce sufficiently and were therefore not shown. However, as will see below, the radical quenching was readily detectable by EPR.

Interestingly, a molar excess of approximately two CsdA per glycyl radical subunit appeared to be necessary to accomplish maximum activation of CsdB. Moreover, from these low-resolution experiments there was no indication for a significant acceleration of the activation process at higher excess. This saturation behavior was seen even nicer during the re-inactivation at later time points, which apparently occured with the same rate for CsdA: CsdBC ratios at and above 2.

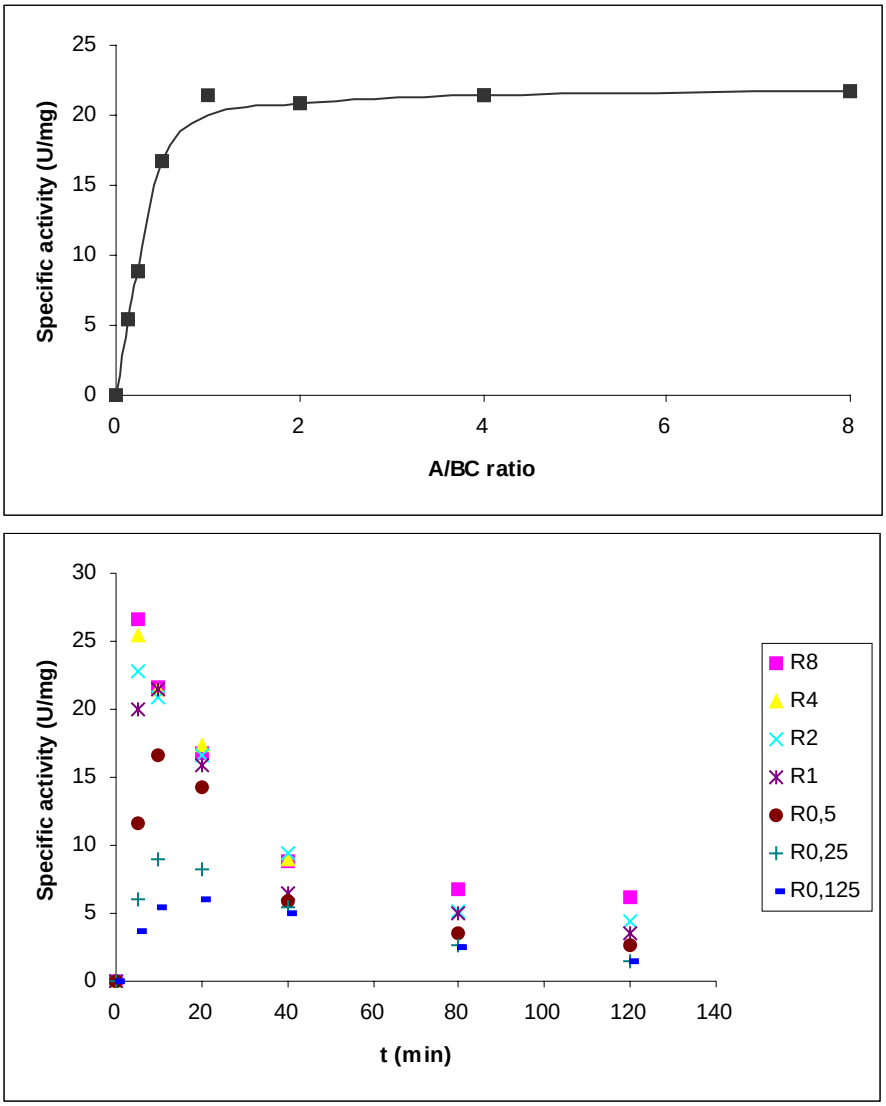



**Figure3.21 Different CsdA/CsdBC ratio activation.** The ratios of CsdA/CsdBC in individual enxperiments are given in the legend of the lower panel. The concentration of CsdBC was kept constant (~1.7µM) and the CsdA/CsdBC ratio was varied from 0.125 to 8. After 30min of pre-reduction at 30°C in solution A (100 mM Tris/HCl, pH 7.5, 25 mM NaCl, 5 mM DTT, 5 mM cysteine, 5mM $MgCl_2$, 5mM $(NH_4)_2SO_4$) supplimented with 2.5mM Ti-citrate and 2.5mM $Na_2SO_3$, samples were withdrawn and mixed with assay buffer for activity tests. Then, SAM was added to the pre-activation (0.25mM) and additional samples were taken at indicated time points and measured.

## Alternative chemical reductants for the AE

Most previously described experiments were carried out with commercially available preparations of sodium dithionite. However, it is well established in the literature that even the purest commercially available sodium dithionite preparatations contain about 15 % [w/w] sodium sulfite [88]. Hence it was desired to replace sodium dithionite by an alternative electron source in the *in vitro* activation of CsdBC. A possible substitute for dithionite was provided by the low potential electron source titanium(III)-citrate [89].

$Ti^{3+}$ could not replace sodium dithionite in the *in vitro* activation assays and only negligible activation of CsdBC (less than 2 % of the dithionite control) was obtained using this compound as electron source. However, $Ti^{3+}$-citrate was not inhibitory since an equimolar mixture of sodium dithionite with the compound yielded equal or even slightly higher specific activities.

Interestingly, a pre-incubation of $Ti^{3+}$-citrate with sodium dithionite was accompanied by a clearly visible bleaching of the purple Ti-complex, indicating its oxidation to $Ti^{4+}$ in the presence of commercially available dithionite preparations. These observations suggested that sodium hydrogensulfite (or more likely trace sulfurdioxide ($SO_2$) at pH 7.5) might act as an electron acceptor. This reaction would directly yield the sulfuranyl radical anion, which has been shown to provide the reducing species in dithionite-dependent reactions [88]. Indeed, it turned out that a mixture of sodium hydrogensulfite and $Ti^{3+}$ provides a very effective system for the reduction of the iron-sulfur clusters in CsdA and also in HpdA (M. Blaser, unpublished results), and was therefore used in many of the subsequently described experiments.

## Michaelis-Menten kinetics for CsdBC.

The previously described characterization of the pre-activation of CsdBC and HpdBC provided the opportunity to study the enzyme kinetics of these enzymes under defined starting conditions. In order to obtain comparable results, standard conditions for the activation of the decarboxylases by the cognate AEs were established, which allowed highly reproducible enzymatic testing.

Based on the previously described features of the activation process, the decarboxylases (0.2 mg/mL) were pre-reduced with a 4-fold excess of the cognate AEs at pH 7.5 by sodium hydrogensulfite/titanium-(III)-citrate for 30 min at 30 °C. Then, SAM was added to yield final concentrations of 250 µM and the activation was allowed to proceed for 5

min. The pre-activated decarboxylase was subsequently assayed at various substrate concentrations. Therefore, the enzyme was diluted 20-fold in the assay buffer. The time-dependent product formation was monitored in 2-minute intervals for 10 min by RP-HPLC of perchloric acid inactivated samples withdrawn from the assay mixture. The substrate concentrations in individual assays were varied in serial dilutions from 64 mM down to 125 µM and the initial velocities were plotted against the substrate concentrations (figure 3.22). The $K_m$- and $V_{max}$-values were obtained from non-linear fits of the data to the Michaelis-Menten-equation using the enzyme kinetics package provided by IgorPro.

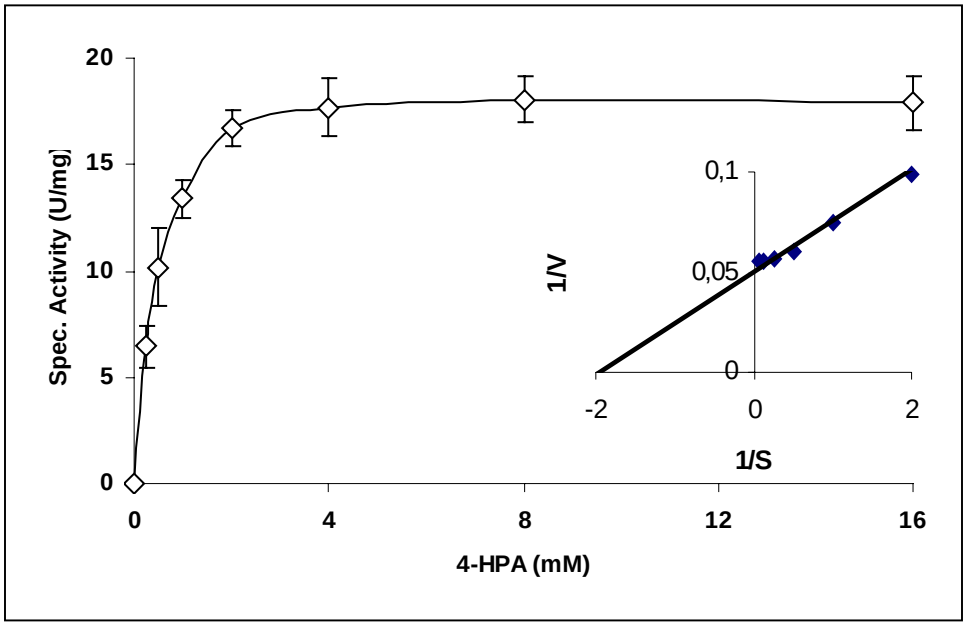



**Figure 3.22 Michaelis-Menten kinetics for CsdBC with HPA as substrates.** The details are described in text. The inset shows the corresponding Lineweaver-Burk plot of the data.

The individual results obtained for CsdBC and HpdBC with 4-hydroxyphenylacetate and 3,4-dihydroxyphenylacetate were summarized in table 3.4. These data indicated that the enzymes are very similar with regard to kinetic parameters and catalysed the decarboxylations of both substrates with very similar efficiencies.

Table 3.4 Kinetic properties of CsdBC and HpdBC

| Enzyme | Substrate | $K_m$ (mM) | $V_{max}$ (U/mg) | $k_{cat}$ ($s^{-1}$) | Specificity constant ($k_{cat}/ K_m$) |
|---|---|---|---|---|---|
| HpdBC | 4-HPA | 0.6 | 15 | 95 | $9.5x10^4$ |
| | 3,4-DHPA | 0.4 | 9 | 17 | $4.2x10^4$ |
| CsdBC | 4-HPA | 0.4 | 19 | 85 | $1.8x10^5$ |
| | 3,4-DHPA | 0.4 | 9 | 17 | $4.2x10^4$ |

The products of the decarboxylation reactions, *p*-cresol and 4-methylcatechol, were subsequently tested for product inhibition. Therefore, the effect of increasing *p*-cresol concentrations on the 3,4-dihydroxyphenylacetate decarboxylase activities was studied, while 4-methylcatechol was used in the 4-hydroxyphenylacetate decarboxylation assays. However, neither of these compounds showed any significant inhibition of the reaction in concentrations up to 2 mM. It is worth noting, that the *p*-cresol concentrations in *C. difficile* cells grown in 4-hydroxyphenylacetate-containing media was always much lower (< 100 μM), suggesting that product inhibition does not play any role in the regulation of HpdBC or CsdBC.

## EPR results

### Detection of the glycyl radical in HpdBC, CsdBC and TfdBC

EPR samples for the detection of the glycyl radical in activated recombinant decarboxylase were prepared from freshly prepared enzymes and activated under standard conditions (e.g. Tris-HCl buffer, pH 7.5, 5 mM each DTT, ammonium sulfate and magnesium chloride and 1 mM sodium dithionite). As shown in figure 3.23, partially purified endogenous HpdBC from *C. difficile* (A) and *in vitro* activated recombinant HpdBC (B), TfdBC or CsdBC (C) yielded virtually identical EPR signals, which showed the characteristic features of glycyl radicals. The G-values of about 2.0023 and the unique lineshapes due to the spin coupling of the radical with the "residual" hydrogen atom at the α-carbon of the glycyl residue are clearly visible.

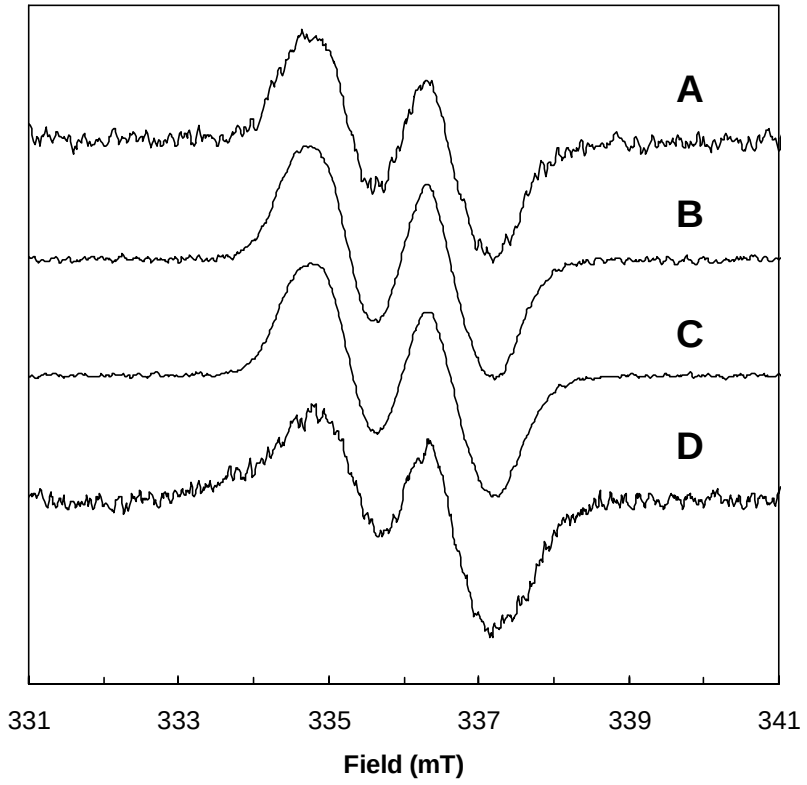



**Figure 3.23 EPR spectra of the glycyl radical in *p*-hydroxyphenylacetate decarboxylases.** Trace A, partially purified endogenous *p*-hydroxyphenylacetate decarboxylase from *Clostridium difficile* (14 U/ml; spin concentration ~1.25 μM) from *Clostridium difficile*. Trace B, *p*-hydroxyphenylacetate decarboxylase (4 mg/ml HpdBC; spin concentration ~9.5 μM, equal to 1Gly$^{\bullet}$ /hetero-octamer) (Hpd). Trace C, Recombinant *p*-hydroxyphenylacetate decarboxylase from *Clostridium scatologenes* (Csd) (4 mg/ml

CsdBC; spin concentration ~8 µM, equal to ~1 $Gly^{\bullet}$ /hetero-octamer) obtained after heterologous expression in *E.coli*. Trace D, recombinant Tfd (4 mg/ml TfdBC; spin concentration ~1 µM) from *Tannerella forsythensis*. The samples were prepared in buffer A containing 4 fold molar excess of the activating enzyme and reduced with 1 mM dithionite for 30 min at 30 °C. Then SAM was added to a final concentration of 0.25 mM and incubated for another 5 min prior to freezing in liquid nitrogen. The activity was measured for all steps separately. Spectra were recorded at 60 K (trace A and C), 77 K (trace D) or 100 K (trace B); microwave power, 13 µW; microwave frequency, 9.437 GHz (trace A), 9.458 GHz (trace B and C) or 9.433 GHz (trace D); modulation frequency, 100 kHz; modulation amplitude, 0.45 mT (0.9 mT for trace D). Averages of 8 scans were used (12 scans for trace D).

The spin concentrations of the glycyl radical in the EPR samples (8 µM $Gly^{\bullet}$ for CsdBC, 9.5 µM $Gly^{\bullet}$ for HpdBC, 1 µM $Gly^{\bullet}$ for TfdBC) were calculated and compared to the known protein concentration (4 mg/ml, equal to 9 µM for the hetero-octameric complexes HpdBC and CsdBC or to 18 µM for the hetero-tetrameric complex of TfdBC). These data suggest that there was only one of the 4 potential glycyl radical sites in the hetero-octameric complexes was activated during the activation process. Concerning the previously reported activation of half of the sites in other GREs, the less glycyl radical formation maybe due to the incomplete activation of glycyl radical decarboxylases. For the TfdBC, the glycyl radical content was much lower and only about 10 % of the radical content of the other decarboxylases, although protein concentration was twice as high due to the lower native complex mass. However, the radical concentration of the TfdBC sample was almost equal to the spin concentration of a previously prepared sample of partially purified endogenous Hpd from *C. difficile*. The latter sample exhibited volume activities of 14 U/mL. Hence the inability to detect any cresol formation by TfdBC in standard assays reflects the fact that the substrate of this enzyme is most likely not HPA.

.

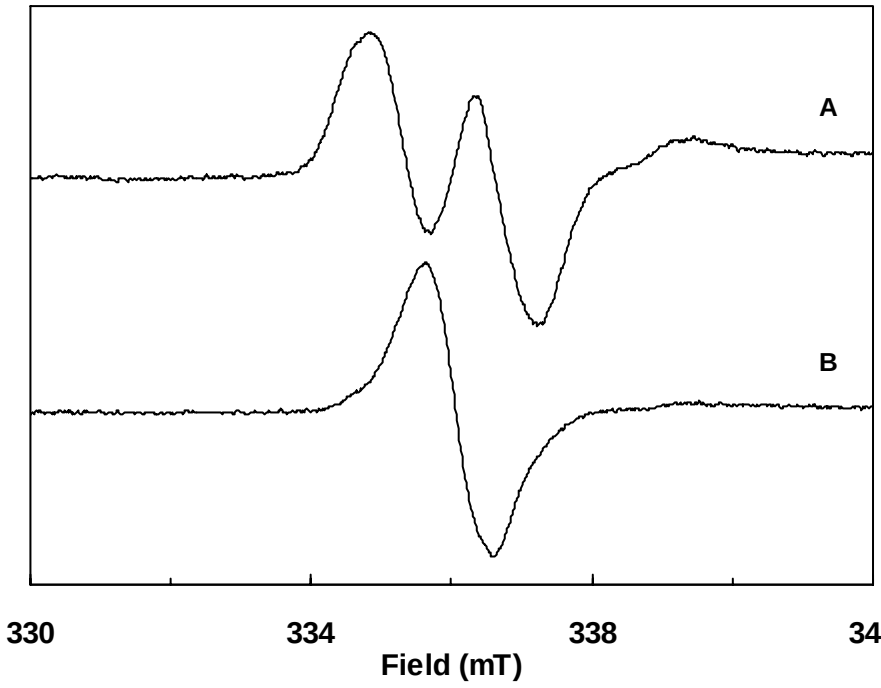



**Figure 3.24 EPR spectra of the glycyl radical in *Clostridium scatologenes p*-hydroxyphenylacetate decarboxylase.** The samples contained 2 mg/ml CsdBC, 2.7 mg/ml CsdA, 0.5 mM $Na_2SO_3$, 0.5 mM titanium (III) citrate in buffer A. After the addition of SAM and incubated for another 5min, the sample was frosen in liquid nitrogen for EPR measurement. Trace A, in $H_2O$; trace B in 90 % v/v $D_2O$. Spectra were recorded at 77 K; microwave power, 50 μW; microwave frequency, 9.434 GHz; modulation frequency, 100 kHz; modulation amplitude, 0.4 mT. Averages of 12 scans were used. Weak broad signals between 339 and 341 mT are from residual titanium (III) citrate.

As shown in figure 3.24, the "residual" hydrogen in CsdBC readily exchanges with solvent deuterium in the absence of substrate when the activation was carried out in $D_2O$.
It has been earlier stated that the activation of CsdBC by CsdA in the absence of substrate caused an activation, which was monitored by activity measurements and oxygen-induced cleavage of the polypeptide. However, EPR spectroscopy allowed a direct monitoring of the radical quenching throughout such experiments. Therefore, concentrated samples containing CsdBC or HpdBC, CsdA or CsdA, reductant and SAM were prepared. At various time-points samples were withdrawn and analysed in enzymatic assay and by EPR spectroscopies. As shown in figure 3.25 and 3.26, increase and decline of specific activities correlated very well with the appearance and disappearence of the glycyl radical signal.

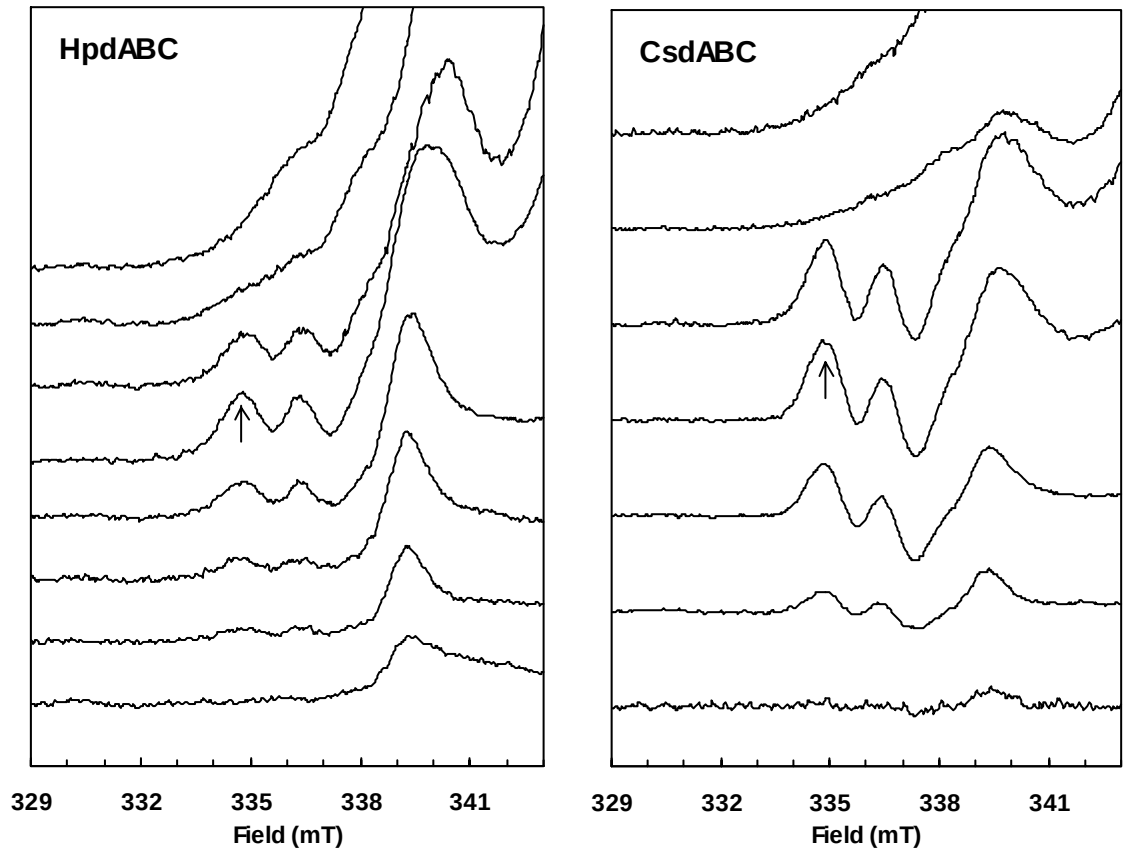



**Figure 3.25 Formation and stability of the glycyl radical in *p*-hydroxyphenylacetate decarboxylases.**

Left panel, from top to bottom: fresh titanium (III) citrate (2.5 mM), *C. difficile* *p*-hydroxyphenylacetate decarboxylase (HpdABC) in absence of SAM, idem in the presence of 0.25 mM SAM after 5, 10, 30, 60, 120 min incubation at 30 °C, respectively. Samples contained 4 mg/ml HpdBC, 5 mg/ml HpdA, 2.5 mM $Na_2SO_3$, 2.5 mM Ti(III)-citrate in buffer A and were frozen at the indicated times after addition of SAM. The lowest trace was from 2.5 mM Ti(III) citrate after 90 min.

Right panel, from top to bottom: fresh Ti(III) citrate (0.5 mM), *C. scatologenes* *p*-hydroxyphenylacetate decarboxylase (CsdABC) in absence of SAM, idem in the presence of 0.25 mM SAM after 5, 10, 30, 60, 120 min incubation at 30 °C, respectively. Samples contained 4 mg/ml CsdBC, 5 mg/ml CsdA, 0.5 mM $Na_2SO_3$, 0.5 mM Ti(III) citrate in buffer A and were frozen at the indicated times after addition of SAM.

Pronounced signals, which dominate the right part of the spectrum, are from Ti(III)-citrate. Spectra were recorded at 77 K; microwave power, 50 μW; microwave frequency, 9.433 ±0.002 GHz; modulation frequency, 100 kHz; modulation amplitude, 1.2 mT. Averages of 2 scans were used.

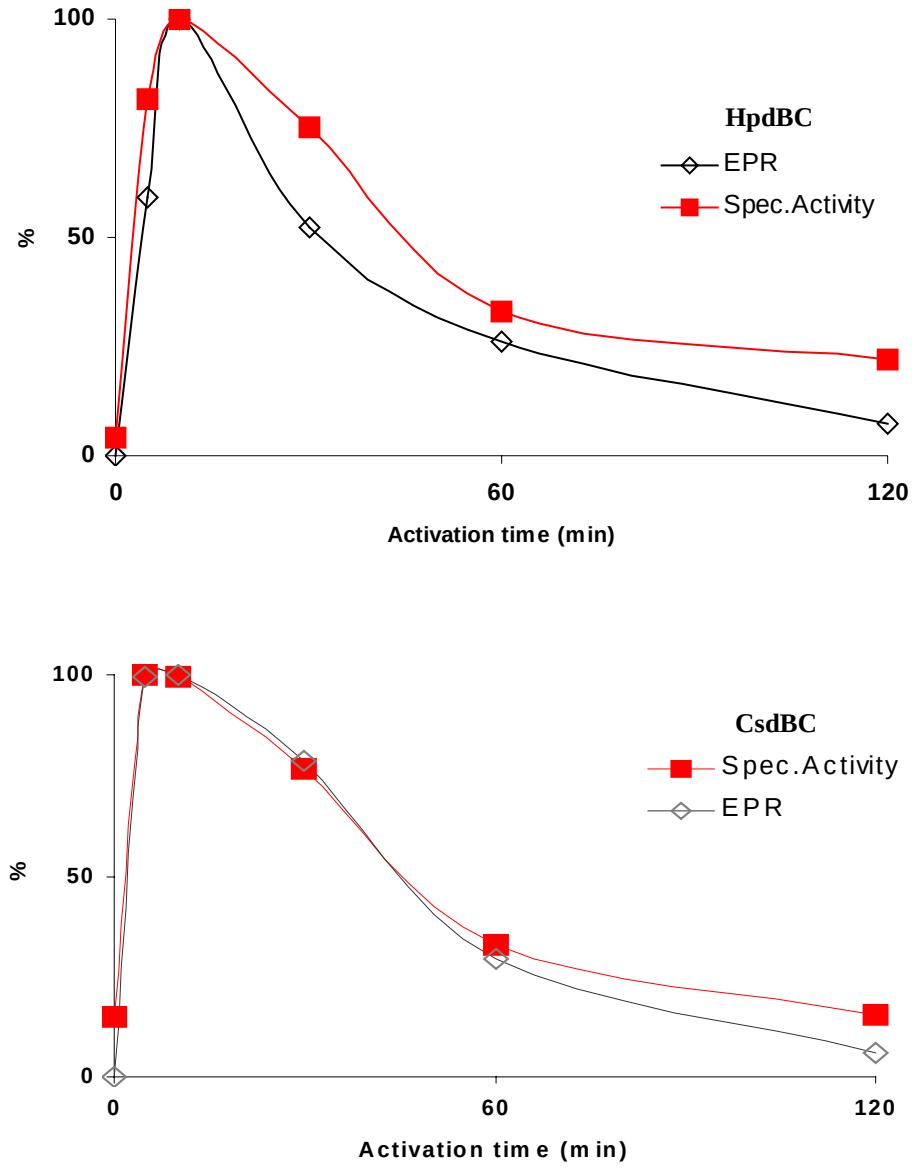



**Figure 3.26 Time dependence of the intensity of the glycyl radical EPR signal (in percent) in HpdBC and CsdBC** The intensities of the absorption-shaped feature of the glycyl radical at 334.81 mT ± 0.06 mT (see arrows in Fig. 3.25) were corrected for a background from the left shoulder of the Ti(III)-citrate EPR signal. This background was estimated for each sample from the average EPR intensity at 334.23 ± 0.06 mT and at 335.40 ± 0.06 mT. The error caused by this subtraction was equivalent to maximally 2 % glycyl radical based on application of this method for fresh and aged Ti(III)-citrate. 100 % equaled to ~9.4 µM radical or ~8.4 U/mg specific activity.

**Initial characterisation of the metal centers in Hpd and Csd**

The metal centres in the AEs and in the decarboxylase of the Hpd and Csd systems were initially characterised by EPR spectroscopies at low temperature. For these measurement series individual samples were prepared and subsequently analysed by X-band EPR in order to establish properties of the individual constituents of the systems and also to obtain initial indications for interactions between the metal centres throughout the activation process.

It should be stated, that the purified AEs were entirely silent in EPR (data not shown). Since [3Fe-4S] in the oxidised stated would yield EPR signals, there is no evidence from this measurements for cluster decay. However, as shown in figure 3.27, very clear signals of reduced cubanes were obtained from samples, which have been reduced with a 10-fold molar excess of sodium dithionite. These signals are characteristic for spin 1/2 systems, while no additional signals indicating high-spin clusters (spin 3/2) were detectable. The temperature-dependence of the signal clearly showed that these signals arose predominantly from [4Fe-4S] clusters. The detection of very poor signals remaining at higher temperatures suggested that only negligible [2Fe-2S] clusters were present in the samples. Although the overall appearance of the EPR signals and the g-values were similar for HpdA and CsdA, there were also clear differences. The HpdA EPR was significantly broader and less intense than the signal observed for CsdA.

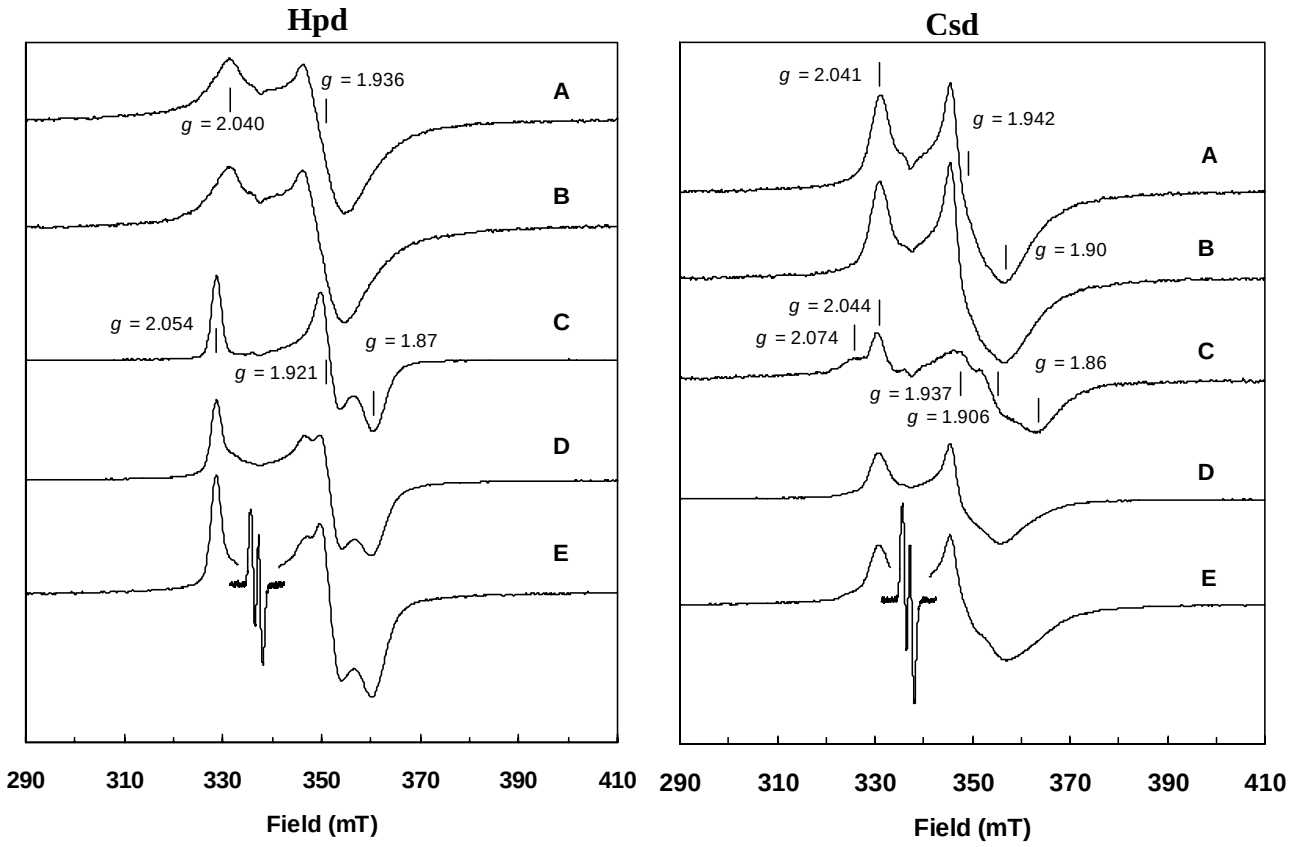



**Figure 3.27 EPR spectra of the *p*-hydroxyphenylacetate decarboxylase systems from *C. difficile* (left) and *C. scatologenes* (right).** Trace A, HpdA/CsdA (5 mg/ml) in buffer A plus 1 mM dithionite, reduced at 30 °C for 30 min; Trace B, as the samples in traceA plus 0.25 mM SAM for 5 min; Trace C, 4 mg/ml HpdBC/CsdBC in buffer A plus 1 mM dithionite (trace C in Csd was enlarged 2-fold for comparison purpose), reduced at 30 °C for 30 min; Trace D, 4 mg/ml HpdBC/CsdBC and 5 mg/ml HpdA/CsdA in buffer A plus 1 mM dithionite, reduced at 30 °C for 30 min; Trace E, the sample was treated as in trace D, then SAM was added to 0.25 mM. After 5 min incubation, sample was frozen for EPR measurement (the central part with the very intense glycyl radical signal was saturated and therefore the inset has been recorded with different settings). Spectra were recorded at 10 K (60 K for the inset of trace E); microwave power, 0.8 mW (0.13 mW for the inset of trace E); microwave frequency, 9.459-9.460 GHz; modulation frequency, 100 kHz; modulation amplitude, 1.25 mT (0.45 mT for the inset of trace E).

Interestingly, there are virtually no changes in the appearance of the EPR signal of HpdA upon addition of SAM. The CsdA spectrum also shows only minor changes, predominantly

a small increase in signal intensity rather than the expected changes, which have been reported for other GRE-AEs to occur when SAM binds to the SAM-cubane.

Very clear rhombic signals indicating [4Fe-4S] centres were also observed in reduced preparations of both decarboxylases. However, although these signals are very intense in reduced samples, about 30 % (HpdBC) or 50 % (CsdBC) of the cubanes were reduced also in the purified enzymes without sodium dithionite (data not shown). These findings suggested that either the redox potential is quite high and a reduction by DTT might cause partial reduction, or that these clusters are very stable and, therefore, were purified from *E. coli* in a partially reduced state. However, while the EPR of reduced HpdBC suggested that the metal centres in this enzyme are homogenous, the signal obtained for CsdBC is more complex and may indicate the presence of subspecies with different local environments.
When the decarboxylases were reduced in presence of their cognate AEs, the resulting EPR spectra showed basicly the features expected from a summation of those spectra obtained for the individual enzymes. However, the individual protein concentrations in the particular samples are the same and the spectra have been scaled. Therefore, the signal intensities are directly comparable. Considering the signal intensities, the presence of the decarboxylases apparently slightly reduced the reduction of the AEs in the mixed samples. This view was further supported by the observation that the addition of SAM did not only cause the glycyl radical signal to show up, but also affected the relative signal intensities of the cubane signals from the individual constituents. However, a quantitative analysis of these changes was not possible at present and must await the characterisation of the cubanes in the individual proteins by Mössbauer-spectroscopy. However, since there are only minor changes in the overall appearance of the iron-sulfure centres in the shown spectra, there is no clear indication for magnetical interaction between the individual clusters or any of the clusters with the glycyl radical.

## Disscussion

This work aimed to increase the number of known bacterial arylacetate decarboxylases systems from anaerobes. The recombinant production of the individual proteins encoded by these systems in *E. coli* and *in vitro* biochemical studies were used to define unique properties of these novel glycyl radical enzymes in comparison to the well known pyruvate formate-lyases (Pfl) and class III ribonucleotide reductases (Nrd).

### Genetic arrangements of the decarboxylase loci

While the genomic DNA sequences encoding the individual components of the 4-hydroxyphenylacetate (HPA) decarboxylase system from *C. difficile* (Hpd) and a very similar system of unknown function from *T. forsythensis* (Tfd) were readily accessible in genome and gene databases [11, 12], the locus encoding a hitherto unknown HPA decarboxylase from *C. scatologenes* (Csd) was *de novo* sequenced by a combination of PCR-based techniques.

Glycyl radical enzymes (GREs) are versatile biocatalysts, which enable strict or facultative anaerobes to perform chemically difficult reactions under anoxic conditions [90-92]. The analysis of recently available whole genome data from various microorganisms allowed the identification of at least eight distinct classes within the large group of GREs [93, 94]. However, at present only five of these groups are functionally characterized, while the other three systems (‘Pfl2’, ‘Pfl3’ and ‘SHORT’) are wide spread in anaerobes, but their physiological function is still unknown [93, 94]. The GREs of known function catalyze chemical reactions as different as the reversible cleavage of pyruvate to yield acetyl-CoA and formate (pyruvate formate-lyase, Pfl [94]), the reduction of ribonucleoside-triphosphates to the corresponding deoxyribonucleotides (anaerobic ribonucleotide reductase, Nrd [95]) and the elimination of water from glycerol (glycerol dehydratase, Gdh [96]) as well as the addition of the methylgroup of toluene to the double bond of fumarate yielding benzylsuccinate (benzylsuccinate synthase, Bss [97]) or the decarboxylation of arylacetates to the corresponding arylmethanes (Hpd [11] and Csd [this work]).

All these GRE systems are composed of the catalytically competent GRE itself and of an activating enzyme (AE), which is essentially required for the posttranslational modification of the inactive GRE precursor protein to the radical-bearing, catalytically active GRE. For the particular GRE systems, these individual components exhibit a characteristic genetic

arrangement (figure 3.10). GRE decarboxylase (GRE-DC, [12] and this work) and Bss systems [97] vary from this general genetic arrangement by the presence of additional small subunit genes in their genetic context as well as by the presence of auxiliary genes, which probably provide substrate transport or regulatory systems (GRE-DCs and Bss) and specific maturation factors for the GRE (Bss).

The insertion of a small subunit between the glycyl radical subunit genes and the AE genes in GRE-DCs allows a rather selective identification of these glycyl radical systems in the genomes of microorganisms. All GREs contain the highly conserved amino acid consensus sequence RVxG[FWY]$x_{6-8}$[FL]$x_4$Q$x_2$[IV]$x_2$R close to the C-terminus of the GRE [92], which carries the glycyl radical in the activated state. Moreover, all AEs contain a highly conserved GCPLxCKWCxN-motif close to their N-terminus, which provides the cysteinyl ligands for the functionally essential SAM-cluster. Using degenerated DNA-primers deduced from these sequences (the sense primer covering the glycyl radical motif and the anti-sense primer the SAM radical motif, respectively), the expected PCR products amplified from GRE-DCs are significantly larger (500-1500 bp) than those obtained for any other GRE system (no products for Bss and Pfl3, or 200-300 bp products for Pfl, Nrd, Gdh, Pfl2 or SHORT). This approach indeed yielded the initial DNA sequence information for the *de novo* sequencing of the *csd*-locus from *C. scatologenes*.

**Recombinant production of the GRE decarboxylases**

The recombinant production of 4-hydroxyphenylacetate decarboxylase from *C. difficile* (Hpd) was initially hampered by the fact that the small subunit was lost during purification of endogenous enzyme and its gene escaped detection in the first analysis of the *hpd* locus [11]. Hence, all attempts to express the *hpdB* gene alone in *E. coli* were not successful yielding exclusively insoluble protein. However, the co-expression of *hpdB* with the small subunit gene (*hpdC*) yielded soluble and, more importantly, functional enzyme [12]. Consequently, the two genes encoding the GRE decarboxylases from *C. scatologenes* and *T. forsythensis* were co-expressed and soluble proteins were obtained. Interestingly, the genetic combination of glycyl radical subunit genes with the small subunit genes of the other two systems yielded hybrid decarboxylases, which were soluble for some of the combinations (*hpdB-tfdC*, *hpdB-csdC* and *tfdB-hpdC*), while others caused the formation of insoluble inclusion bodies in *E. coli* (*csdB-hpdC*, *csdB-tfdC* and *tfdB-csdC*). Since individually produced glycyl radical subunits were always found in inclusion bodies ([12]

and this work), the small subunits are probably essential for the proper folding of the large subunit and the cross-system compatibility of the subunits accounts for the rescue of the nascent polypeptide chain from denaturation in *status nascendi*. However, none of the hybrid decarboxylases was active in enzyme tests and analysis of the native molecular masses by gel filtration indicated that only the natural subunit of the individual systems was able to promote the functionally important oligomeric state of the decarboxylases (see below).

Remarkably, in spite of the extreme sensitivity of the purified AEs and the activated decarboxylases towards molecular oxygen, the decarboxylases as well as the activating enzymes were produced in aerobically grown *E. coli*. Though most of the proteins thus obtained were functional, the Tfd-AE showed a strongly reduced iron content and was therefore purified from anaerobically grown cells.

A carefully controlled growth of the recombinant *E. coli* strains was essential for the production of high quality cells for enzyme preparations. In particular all three AEs strongly benefit from a pre-induction ethanol stress, up-regulating endogenous chaperons and assisting folding of the recombinant proteins [98]. Moreover, these proteins benefit strongly from an *in situ* reconstitution of the metal-centres in cell-free extracts, since a clearly visible browning of the extracts upon centrifugation at elevated temperatures in the presence of $Fe^{2+}$, cysteine and ATP was observed.

The poor intracellular solubility of the AEs was also reflected by the post-purification solubility of the enzymes. Initially, it was almost impossible to concentrate AE preparations to yield concentrations above 5 mg/mL. However, the addition of sodium glutamate and arginine hydrochloride strongly improved the solubility [99].

**The biochemical characteristics of GRE decarboxylases**

Glycyl radical enzymes (GREs) are generally synthesized as catalytically inactive precursor proteins that require a post-translational activation by a specific iron-sulfur protein (activating enzyme, AE) [74, 100-103,54]. The AEs are members of the SAM radical enzyme superfamily [55] and catalyze the reductive cleavage of S-adenosylmethionine (SAM) to yield methionine and a 5'-desoxyadenosyl radical, which is subsequently used in the stereospecific abstraction of the C-2 *pro-S* hydrogen from a strictly conserved glycyl residue [54].

While these two essential components are common to all glycyl radical systems, benzylsuccinate synthase [97, 104] and the GRE decarboxylases contain small subunits. When the glycyl radical subunits of the individual decarboxylases were genetically combined with the small subunits of the other GRE decarboxylases, soluble recombinant proteins were formed and purified for some of these combinations. Interestingly, in all combinations yielding soluble enzymes, the molar ratio of the B and C subunits was 1:1, while the oligomeric state differed from the wildtype complexes, exhibiting a lower order of complexity. These data suggest that the small subunit is directly involved in the regulation of the oligomeric state of GRE decarboxylases. Moreover, all three wild-type decarboxylases and all hybrid decarboxylases contained 4 irons and 4 sulfurs per glycyl radical subunit.

The oligomeric state of the GRE decarboxylases provides a second unique feature of these enzymes as compared to other GREs. While HpdBC ($\beta_4\gamma_4$) and CsdBC ($\beta_4\gamma_4$) were hetero-octamers, which were smoothly activated *in vitro* to specific activities of about 20 U/mg, TfdBC (($\beta_2\gamma_2$) was a hetero-tetramer lacking activity. The available data strongly suggest that the Tfd system differs from the former systems by its substrate-specificity and it seems reasonable to assume that a hetero-octameric state in this system might be formed preferentially upon binding of substrate. Accepting this assumption, all GRE decarboxylases are hetero-tetrameric with respect to the numbers of glycyl radical subunits within the native complexes. This is in sharp contrast to all other known GREs, which contain only 2 glycyl radical subunits. However, recently the purification of GRE from *Archaeoglobus fulgidus* (Pfl2) has been reported, which probably forms a homo-trimer [93].

It is a rather important finding, that the hetero-octameric complexes of the recombinant decarboxylases remained stable after *in vitro* activation, while inevitable decay of this complex was observed for the endogenous Hpd from *C. difficile* [11, 12]. Though not yet fully investigated, preliminary findings strongly suggest that the oligomeric state and thereby the specific activity of GRE decarboxylases is controlled by a reversible serine phosphorylation of the glycyl radical subunits (P. Andrei & T. Selmer, unpublished).

The GRE activating enzymes contain a labile [4Fe-4S] $^{1+/2+}$ cubane, which upon reduction generates the adenosyl radical from SAM. Initial analysis was hampered by reductant- and SAM-dependent cluster conversion. Upon binding of SAM, the EPR spectrum of the reduced cubane cluster changes from rhombic to axial [44, 74]. Cleavage of SAM [42] and oxidation of the cubane [40] occurs stoichiometrically with formation of the glycyl radical.

Although no X-ray crystallographic structural information is available yet for GRE-AEs, the direct coordination of SAM to the cubane has been elucidated by Mössbauer [43] and ENDOR spectroscopy [44, 45]. The mode of binding is in agreement with X-ray crystallographic data on SAM-binding in coproporphyrinogen III decarboxylase [46] and biotin synthase [47].

The SAM cluster anchors the methionine of SAM, whereas the adenosyl group is bound by a glycine-rich SGG-sequence motif [55]. This motif is directly adjacent to the cysteines coordinating the SAM cluster in Pfl-AE and Nrd-AE, whereas other GRE-AEs contain a 31 to 64 amino acid insertion between these functional motifs. Within this insert, 4 to 8 cysteinyl residues are located, which form up to two ferredoxin-like $Cx_{2-4}Cx_2Cx_{7-34}C$ motives. It has been previously suggested, solely based upon bioinformatic information, that this 'complex' group of AEs might differ in its metal content from the well-studied Pfl- and Nrd-AEs, but GRE decarboxylases AEs are the first enzymes for which this assumption has been supported by biochemical data. All three decarboxylase AEs contain 8 irons and 8 sulfurs per molecule, and the EPR data strongly suggest the presence of two redox active cubanes in these proteins, the common SAM-cluster and a second one, which has been tentatively named the insert or I-cluster. Remarkably, the previously reported, SAM-dependent changes in the appearance of the EPR spectra of Pfl- and Nrd-AE [44, 74] were not seen for the decarboxylase AE, suggesting that the signals arising from the I-cluster might cover these spectral changes.

Upon activation of GREs the stable protein-bound glycyl radical is generated by its cognate AE. The seminal work of Knappe on Pfl established that the $sp^2$ radical was centred at the α-carbon atom of a specific glycine residue [23, 34]. It turned out that the characteristic glycyl radical EPR signal has a remarkably constant appearance (a doublet with a hyperfine coupling of 1.4-1.5 mT centred around $g_{iso} = 2.0035$ at X-band) in Pfl [23, 34], Nrd [35, 36], GRE decarboxylases (this work), Bss [37] [38] and methylpentylsuccinate synthases [39]. Combined with the sensitivity of detection (down to concentrations of 0.2 μM) the fingerprint EPR signal allows convenient studies on the kinetics and extent of activation [40] as well as identification of GREs in whole cells and cell-free extracts [38, 39]. The glycyl radical is stabilised by delocalisation of the free electron over the adjacent peptide bonds in the protein backbone and is regarded as a "storage form" of the radical. In the presence of oxygen, the glycyl radical becomes highly unstable, and the enzyme is irreversibly inactivated by cleavage of the polypeptide chain at the site of the radical. Spectroscopic characterization of sulfinyl and peroxyl radical

breakdown products of dioxygen inactivated Pfl [105] allows assessment of inadvertent exposure to oxygen. Temperature and microwave power dependence of the EPR signal pointed out that glycyl radicals occur in a magnetically isolated form. In addition, the limited *g*-anisotropy in high frequency EPR spectroscopy [41] provided evidence for a similar local environment. A rhombic *g*-tensor was ressolved for Pfl, Nrd and Bss, which was in agreement with EPR studies on irradiated *N*-acetylglycine crystals. However, certain differences have been observed. Whereas the 'residual' hydrogen atom at the α-carbon atom of the glycyl radical can be exchanged in $D_2O$ in Pfl [34], Bss [37, 38] and GRE decarboxylases (this work), exchange in Nrd was not detected [35].

The reaction cycle of all known glycyl radical enzymes is thought to be initiated by transfer of the radical to the thiol group of a conserved cysteine located in the middle of the large subunits amino acid sequence, which is close to the glycyl radical site in the tertiary structure [28, 106, 107]. The resulting highly reactive thiyl radical then initiates a radical reaction path with the substrates [93].

**Both, HpdBC and CsdBC, were smoothly activated by the cognate AEs to yield specific activities of about 20 U/mg with 4-hydroxyphenylacetate as substrate. Both enzymes also catalysed the decarboxylation of 3,4-dihydroxyphenylactetate, but were entirely inactive towards 4-aminophenylactetate or indole-3-acetate. Interestingly, the quantification of the glycyl radical by EPR suggested that only one glycyl radical is formed per hetero-octameric complex. However, it must be stated that the radical content of fully active decarboxylases could be significantly higher and the observed radical content may reflect the equilibrium content, which is cause by two individual reversible processes (see below). In contrast, TfdBC was completely inactive with the substrates tested. Though to a much lower extent (5%), glycyl radical formation mediated by TfdA was indeed detected. Since the radical concentration in these samples was sufficient to promote measurable turnover of substrates, we conclude that none of the tested compounds provided the natural substrate of TfdBC. The very low efficiency of radical formation in TfdBC is probably due to the inability of the enzyme to form significant amounts of hetero-octameric complex in absence of the natural substrate. This view is supported by the observation that no radical formation was observed for any of the hybrid decarboxylases either, which differed also in their oligomeric state from wild-type HpdBC.**

The glycyl radical is stable in anoxic environments for several hours to days in Pfl and Nrd. However, the catalytically active radical form of Pfl and the radical-free precursor are

freely interchangeable [108], allowing facultative anaerobes to rescue Pfl when changing from anoxic to oxic environments. It is another unique property of GRE decarboxylases that the radical formation occurs transiently in these systems. Though rapidly formed, the radical is clearly quenched during extended incubation of the enzyme in absence of substrates. However, when the substrate is present, the radical is apparently stabilized and there is no indication for activity losses caused by radical quenching. Only recently, the formation of 5'-deoxyadenosine was followed during the time course of the reaction, suggesting an almost complete turnover (Blaser, unpublished). Since SAM was initially present in a 10-fold molar excess and hydrolysis of the compound does not yield 5'-deoxyadenosine, these data suggest that 1) each individual decarboxylase molecule undergoes several cycles of activation and radical quenching, and 2) that a limitation of SAM causes the observed decline in specific activity and radical content towards the end of these experiments.

It is an interesting question, how this radical quenching is achieved. Indeed, the loss of radical could be an experimental artefact, which is caused by a chemical reaction of the active site thiyl radical with the thiols (DTT and cysteine) in the test mixtures. Since an exchange of the 'residual' hydrogen at the glycyl radical site in $D_2O$ has been observed, it is rather clear that a transfer of the radical to the thiol takes place in the absence of substrate. However, when these small molecules were removed from the mixture by gel filtration, the radical quenching was still seen (Blaser, unpublished). These data suggest, that the radical quenching is an intrinsic property of the system. Indeed, it is not difficult to envisage that the transfer of a second electron to either the glycyl radical (forming a carbanion) or – more likely – to the thiyl radical (yielding a thiolate anion) followed by subsequent protonation could readily mediate radical quenching. While the site of radical quenching via the first pathway suggest the participation of the AE in the reaction, the second pathway would more likely present an intrinsic feature of the GRE decarboxylase itself. The required electron is most likely donated by additional iron sulfur centers, which are found the AEs (the I-cluster) but also in the decarboxylases.

## Conclusions

The GRE decarboxylases from *C. difficile*, *C. scatologenes* and *T. forsythensis* form a distinct group of GREs, which is well defined and differs from all other members of the glycyl radical enzyme superfamily. These novel enzymes contain a small subunit in

addition to the glycyl radical subunit, form large, hetero-octameric complexes and contain redox active iron-sulfur centres as prosthetic groups in addition to the glycyl radical.

The decarboxylase AEs are distinguished from other AEs by the metal content. While the former enzymes contain two cubanes (the SAM-cluster and the novel I-cluster), only the SAM cluster is found in other AEs.

The structural differences are accompanied by clearly defined biochemical features, which are not known for other GREs. In particular the only transiently occurring activation of GRE decarboxylases by the cognate AEs *in vitro* suggests that the regulation of the radical content provides the possibility to control cresol formation *in vivo*. Moreover, there is increasing evidence that a reversible serine phosphorylation of the glycyl radical subunit influences the enzymes oligomeric state and thereby its specific activity. These regulatory features are novel properties for GREs and allow the formation of toxic compounds in response to the metabolic state of the producers (figure 4.1).

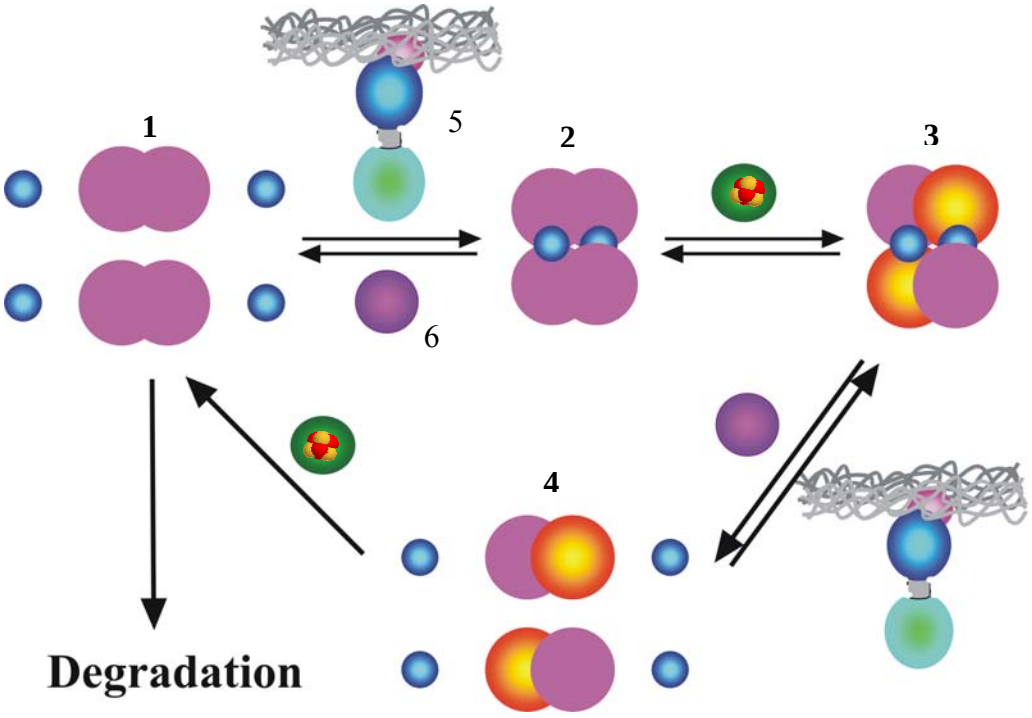



Figure 4.1: A proposed scheme for the regulation of cresol formation *in vivo*. The individual subunits of the decarboxylases (1) might occur individually. The serine-specific phosphorylation of the glycyl radical subunits causes the formation of the activation-competent, hetero-octameric complex (2), which is activated by its cognate AE to the active decarboxylase (3). The dephosphorylation of the enzyme causes complex decay, yielding catalytically inactive, radical-bearing homo-dimers of the glycyl radical subunits

(4), which are predicted to be the natural substrate for radical quenching, preceding protein degradation *in vivo*. Aputative protein kinase (5) and a corresponding phosphorylase (6) might be involed. Note that the complexes 2, 3 and 4 have been experimentally observed. The reversible conversion of complexes 2 and 3 *in vitro* is most likely artificially caused by the use of low potential chemical electron donors in the test system. The red-colored part in complex 3 and 4 represents the activated glycyl radical subunit. It is speculated that the radical formation here is the same as that of Pfl and Nrd, and that maximum two glycyl radicals per octamer can be formed.

**Outlook**

**The data presented here show clearly, that GRE decarboxylases show a number of novel properties, which are not found in other systems and the future work will focus on the elucidation of these features. The proper understanding of the individual processes involved in the SAM-dependent radical formation and in the SAM-independent quenching of the radical will hopefully contribute to a more detailed understanding of the molecular mechanisms involved in the enzymatic handling of these highly reactive intermediates. The putative ligants of metal centers of I-cluster in the AE will be identified by alanine or serine scanning mutagenesis experiments applied on the cysteinyl residues, mutational analysis of the GRE consensus motif will probably yield information about the interaction between the GRE and its cognate AE.**

**The analysis of the structural and functional properties of the decarboxylases will be facilitated in the future by the crystal structure of 4-hydroxyphenylacetate decarboxylase from *C. scatologenes*, which has been solved only very recently by Prof. Holger Dobbek and Dr. Berta Martins at Bayreuth University. Although these data became available so recently that a detailed analysis of the structure cannot be included here, the available data confirmed the proposed overall structure as well as the presence of two metal centres in these enzymes, which are unequivocally bound by the small subunits (figure 4.2).**

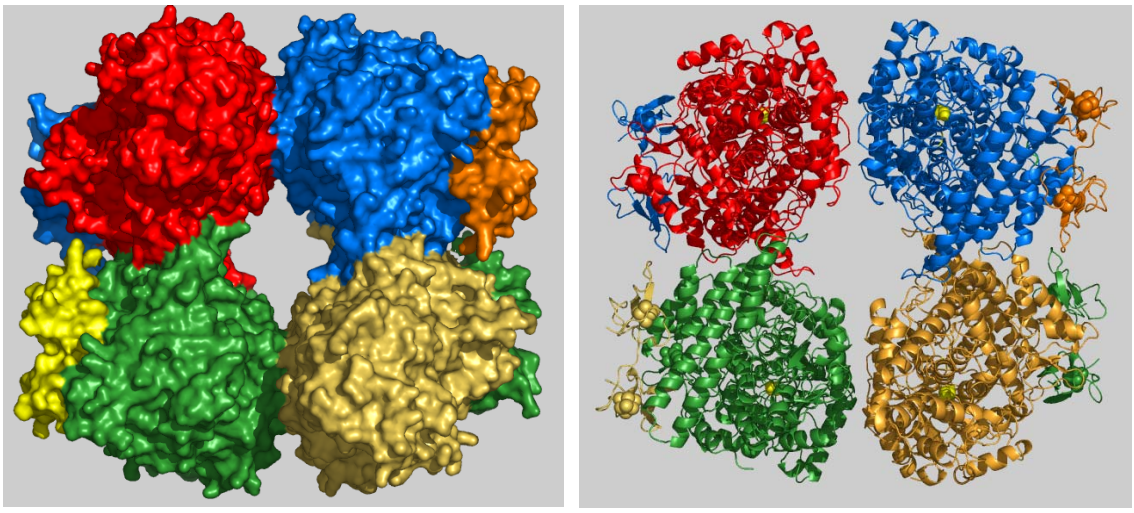


Figure 4.2 Hetero-octameric packing of Csd subunit crystal structure. **The left picture shows a surface presentation, and the right is the Ribbon diagram, which shows the**

**binding of the small subunit and the Fe-S clusters in these unique constituents of GRE decarboxylases.**

The repeated, reversible formation of the radical and its subsequent quenching generates a futile cycle, which is energetically costly and should, therefore, not occur *in vivo*. *In vitro* studies, however, were always conducted with low potential electron donors, which have been proven to reduce all three metal centers at the same time. *In vivo*, alternative electron donors for individual steps may decouple both processes and regulate decarboxylase activities. It should be noted in this respect that a selective reduction of either the AEs alone or of both enzymes by photoreduction could provide experimental access to this question. It will be necessary to establish the participation of individual metal centers in the particular steps. However, the properties of the individual clusters in the AE are difficult to study and recombinant proteins with genetically removed or disrupted cluster motifs might be useful to study both clusters individually.

Other experiments must address the serine phosphorylation of the glycyl radical subunits. First the phosphorylation site needs to be identified, in order to allow the generation of phosphomimetic (S->E) or dephosphomimetic (S->A) mutants, which could be useful in studies of the phosphorylation effects. Moreover, the protein kinase(s) and phosphatase(s) responsible for the *in vivo* regulation of GRE decarboxylases must be identified.

Moreover, the biochemical studies with the pure enzymes must be extended to physiological studies with *C. difficile*. While the effect of varying growth conditions on cresol formation can be directly followed in growth experiments, the induction and the processing of Hpd can be monitored by immunological methods employing the recently generated polyclonal antiserum against HpdBC and HpdA (P. Andrei & T. Selmer. unpublished), or commercially available antibodies against phosphoserine. The recombinant, N-terminally affinity-tagged decarboxylase might be also useful for *in situ* studies with cell-free extracts of the organism.

And finally, the effect of cresol on competitive microbes in the gut and also on the host needs to be studied. Since we predict that cresol formation allows the suppression of other organisms and could therefore provide an important virulence factor for the emerging human pathogen *Clostridium difficile*, the importance of such studies is rather obvious. Moreover, cresol is a known cytotoxin and studies addressing its impact on cultivated human cells are planned. In particular, synergistic effects of cresol and the two exotoxins

of the organism have not yet been studied and might provide novel insights in *C. difficile*-associated diseases.

## References

1. Elsden, S. R., Hilton, M. G. and Waller, J.M. (1976) The end products of the methabolism of aromatic amino acids by *Clostridia*. *Arch. Microbiol. 107*, 283-288.

2. Probst, E. L., Flickinger, E. A., Bauer, L. L., Merchen, N. R. & Fahey, G. C. (2003) A dose-response experiment evaluating the effects of oligofructose and inulin on nutrient digestibility, stool quality, and facal protein catabolites in healthy adult dogs. *J. Anim. Sci. 81*, 3057-3066.

3. Smith, E. A. & Macfarlane, G. T. (1997) Formation of Phenolic and Indolic Compounds by Anaerobic Bacteria in the Human Large Intestine. *Microb. Ecol. 33*, 180-188.

4. Smith, E. A. & MacFarlane, G. T. (1996) Enumeration of human colonic bacteria producing phenolic and indolic compounds: effects of pH, carbohydrate availability and retention time on dissimilatory aromatic amino acid metabolism. *J. Appl. Bacteriol. 81*, 288-302.

5. Birkett, A. M., Jones, G. P., de Silva, A. M., Young, G. P. & Muir, J. G. (1997) Dietary intake and faecal excretion of carbohydrate by Australians: importance of achieving stool weights greater than 150 g to improve faecal markers relevant to colon cancer risk. *Eur. J. Clin. Nutr. 51*, 625-632.

6. Birkett, A. M., Jones, G. P. & Muir, J. G. (1995) Simple high-performance liquid chromatographic analysis of phenol and p-cresol in urine and feces. *J. Chromatogr. B Biomed. Appl. 674*, 187-191.

7. Birkett, A., Muir, J., Phillips, J., Jones, G. & O'Dea, K. (1996) Resistant starch lowers fecal concentrations of ammonia and phenols in humans. *Am. J. Clin. Nutr. 63*, 766-772.

8. Brooks, J. B., Nunez-Montiel, O. L., Wycoff, B. J. & Moss, C. W. (1984) Frequency-pulsed electron capture gas-liquid chromatographic analysis of metabolites produced by *Clostridium difficile* in broth enriched with amino acids. *J. Clin. Microbiol. 20*, 539-548.

9. Muramatsuo, T. (1990) Gut microflora and tissue protein turnover in vivo in animals. *Int. J. Biochem. 22*, 793-800.

10. D'Ari, L. & Barker, H. A. (1985) p-Cresol formation by cell-free extracts of *Clostridium difficile*. *Arch. Microbiol. 143*, 311-312.

11. Selmer, T. & Andrei, P. I. (2001) p-Hydroxyphenylacetate decarboxylase from *Clostridium difficile*. A novel glycyl radical enzyme catalysing the formation of p-cresol. *Eur. J. Biochem. 268*, 1363-1372.

12. Andrei, P. I., Pierik, A. J., Zauner, S., Andrei-Selmer, L. C. & Selmer, T. (2004) Subunit composition of the glycyl radical enzyme p-hydroxyphenylacetate decarboxylase. A small subunit, HpdC, is essential for catalytic activity. *Eur. J. Biochem. 271*, 2225-2230.

13. Spencer, R. C. (1998) Clinical impact and associated costs of *Clostridium difficile*-associated disease. *J. Antimicrob. Chemother. 41 Suppl. C*, 5-12.

14. Williams, C. N. (1999) *Clostridium difficile* diarrhea. *Can. J. Gastroenterol. 13*, 293-294.

15. Borriello, S. P. (1998) Pathogenesis of *Clostridium difficile* infection. *J. Antimicrob. Chemother. 41 Suppl. C*, 13-9.

16. Borriello, S. P., Davies, H. A., Kamiya, S., Reed, P. J. & Seddon, S. (1990) Virulence factors of *Clostridium difficile*. *Rev. Infect. Dis. 12 Suppl. 2*, S185-191.

17. Borriello, S. P. & Wilcox, M. H. (1998) *Clostridium difficile* infections of the gut: the unanswered questions. *J. Antimicrob. Chemother. 41 Suppl. C*, 67-69.

18. Mead, G. C. (1971) The amino acid-fermenting *Clostridia*. *J. Gen. Microbiol. 67*, 47-56.

19. Weinberg, M. G., B. (1927) Donnees recentes sur les microbes anaerobies et leur role en pathologie. *Paris: Masson et Cie*, 1-291.

20. Smith, L. D. S. (1975) Common mesophilic anaerobes, including *Clostridium botulimun* and *Clostridium tetani*, in 21 soil specimens. *Appl. Environ. Microbiol. 29*, 590-594.

21. Wyss, C. (1989) Dependence of peroliferation of *Bacteriodes forsythus* on exgenous N-Acetylmuramic acid. *Infect. Immun. 57*, 1757-1759.

22. Lehtio, L. & Goldman, A. (2004) The pyruvate formate lyase family: sequences, structures and activation. *Protein. Eng. Des. Sel. 17*, 545-552.

23. Wagner, A. F., Frey, M., Neugebauer, F. A., Schafer, W. & Knappe, J. (1992) The free radical in pyruvate formate-lyase is located on glycine-734. *Proc. Natl. Acad. Sci. U S A. 89*, 996-1000.

24. Fontecave, M., Mulliez, E. & Ollagnier-de-Choudens, S. (2001) Adenosylmethionine as a source of 5'-deoxyadenosyl radicals. *Curr. Opin. Chem. Biol. 5*, 506-511.

25. Leuthner, B., Leutwein, C., Schulz, H., Horth, P., Haehnel, W., Schiltz, E., Schagger, H. & Heider, J. (1998) Biochemical and genetic characterization of benzylsuccinate synthase from *Thauera aromatica*: a new glycyl radical enzyme catalysing the first step in anaerobic toluene metabolism. *Molecular Microbiology. 28*, 615-628.

26. Raynaud, C., Sarcabal, P., Meynial-Salles, I., Croux, C. & Soucaille, P. (2003) Molecular characterization of the 1,3-propanediol (1,3-PD) operon of *Clostridium butyricum*. *Proc. Natl. Acad. Sci. U S A. 100*, 5010-5015.

27. Becker, A., Fritz-Wolf, K., Kabsch, W., Knappe, J., Schultz, S. & Wagner, A. F. V. (1999) Structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase. *Nature Structural Biology. 6*, 969-975.

28. O'Brien, J. R., Raynaud, C., Croux, C., Girbal, L., Soucaille, P. & Lanzilotta, W. N. (2004) Insight into the mechanism of the $B_{12}$-independent glycerol dehydratase from *Clostridium butyricum*: preliminary biochemical and structural characterization. *Biochemistry. 43*, 4635-4645.

29. Knappe, J. & Wagner, A. F. (1995) Glycyl free radical in pyruvate formate-lyase: synthesis, structure characteristics, and involvement in catalysis. *Methods Enzymol. 258*, 343-362.

30. Parast, C. V., Wong, K. K., Kozarich, J. W., Peisach, J. & Magliozzo, R. S. (1995) Electron paramagnetic resonance evidence for a cysteine-based radical in pyruvate formate-lyase inactivated with mercaptopyruvate. *Biochemistry. 34*, 5712-5717.

31. Eklund, H. & Fontecave, M. (1999) Glycyl radical enzymes: a conservative structural basis for radicals. *Structure Fold Des. 7*, R257-R262.

32. Logan, D. T., Andersson, J., Sjoberg, B. M. & Nordlund, P. (1999) A glycyl radical site in the crystal structure of a class III ribonucleotide reductase. *Science. 283*, 1499-1504.

33. Selmer, T., Pierik, A. J. & Heider, J. (2005) New glycyl radical enzymes catalysing key metabolic steps in anaerobic bacteria. *Biol. Chem. 386*, 981-988.

34. Unkrig, V., Neugebauer, F. A. & Knappe, J. (1989) The free radical of pyruvate formate-lyase. Characterization by EPR spectroscopy and involvement in catalysis as studied with the substrate-analogue hypophosphite. *Eur. J. Biochem. 184*, 723-728.

35. Mulliez, E., Fontecave, M., Gaillard, J. & Reichard, P. (1993) An iron-sulfur center and a free radical in the active anaerobic ribonucleotide reductase of *Escherichia coli*. *J. Biol. Chem. 268*, 2296-2299.

36. Young, P., Andersson, J., Sahlin, M. & Sjoberg, B. M. (1996) Bacteriophage T4 anaerobic ribonucleotide reductase contains a stable glycyl radical at position 580. *J. Biol. Chem. 271*, 20770-20775.

37. Krieger, C. J., Roseboom, W., Albracht, S. P. & Spormann, A. M. (2001) A stable organic free radical in anaerobic benzylsuccinate synthase of *Azoarcus sp*. strain T. *J. Biol. Chem. 276*, 12924-12927.

38. Verfurth, K., Pierik, A. J., Leutwein, C., Zorn, S. & Heider, J. (2004) Substrate specificities and electron paramagnetic resonance properties of benzylsuccinate synthases in anaerobic toluene and m-xylene metabolism. *Arch. Microbiol. 181*, 155-162.

39. Rabus, R., Wilkes, H., Behrends, A., Armstroff, A., Fischer, T., Pierik, A. J. & Widdel, F. (2001) Anaerobic initial reaction of n-alkanes in a denitrifying bacterium: evidence for (1-methylpentyl) succinate as initial product and for involvement of an organic radical in n-hexane metabolism. *J. Bacteriol. 183*, 1707-1715.

40. Henshaw, T. F., Cheek, J. & Broderick, J. B. (2000) The $[4Fe\text{-}4S]^{1+}$ cluster of pyruvate formate-lyase activating enzyme generates the glycyl radical on pyruvate formate-lyase: EPR-detected single turnover. *J. Am. Chem. Soc. 122*, 8331-8332.

41. Duboc-Toia, C., Hassan, A. K., Mulliez, E., Ollagnier-de Choudens, S., Fontecave, M., Leutwein, C. & Heider, J. (2003) Very high-field EPR study of glycyl radical enzymes. *J. Am. Chem. Soc. 125*, 38-39.

42. Knappe, J., Neugebauer, F. A., Blaschkowski, H. P. & Ganzler, M. (1984) Post-translational activation introduces a free radical into pyruvate formate-lyase. *PNAS. 81*, 1332-1335.

43. Krebs, C., Broderick, W. E., Henshaw, T. F., Broderick, J. B. & Huynh, B. H. (2002) Coordination of adenosylmethionine to a unique iron site of the [4Fe-4S] of pyruvate formate-lyase activating enzyme: a Mössbauer spectroscopic study. *J. Am. Chem. Soc. 124*, 912-913.

44. Walsby, C. J., Hong, W., Broderick, W. E., Cheek, J., Ortillo, D., Broderick, J. B. & Hoffman, B. M. (2002) Electron-nuclear double resonance spectroscopic evidence that S-adenosylmethionine binds in contact with the catalytically active [4Fe-4S](+) cluster of pyruvate formate-lyase activating enzyme. *J. Am. Chem. Soc. 124*, 3143-3151.

45. Walsby, C. J., Ortillo, D., Broderick, W. E., Broderick, J. B. & Hoffman, B. M. (2002) An anchoring role for FeS clusters: chelation of the amino acid moiety of S-adenosylmethionine to the unique iron site of the [4Fe-4S] cluster of pyruvate formate-lyase activating enzyme. *J. Am. Chem. Soc. 124*, 11270-11271.

46. Layer, G., Moser, J., Heinz, D. W., Jahn, D. & Schubert, W. D. (2003) Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J. 22*, 6214-6224.

47. Berkovitch, F., Nicolet, Y., Wan, J. T., Jarrett, J. T. & Drennan, C. L. (2004) Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. *Science. 303*, 76-79.

48. Sun, X., Harder,J., Krook,M., Jornvall,H., Sjoberg,B. & Reichard, P. (1993) A Possible Glycine Radical in Anaerobic Ribonucleotide Reductase from *Escherichia coli:* Nucleotide Sequence of the Cloned nrdD Gene. *Proc. Natl. Acad. Sci. U S A. 90*, 577-581.

49. King, D. S. R., P. (1995) Mass spectroscopic determination of the radical scission site in the anaerobic ribonucleotide reductase of *Escherichia coli*. *Biochem. Biophys. Res. Commun. 206*, 731-735.

50. Mulliez, E., Fontecave, M., Gaillard, J. & Reichard, P. (1993) An iron-sulfur center and a free radical in the active anaerobic ribonucleotide reductase of *Escherichia coli*. *J. Biol. Chem. 268*, 2296-2299.

51. Sun, X., Eliasson, R., Pontis, E., Andersson, J., Buist, G., Sjoberg, B. M. & Reichard, P. (1995) Generation of the glycyl radical of the anaerobic *Escherichia coli* ribonucleotide reductase requires a specific activating enzyme. *J. Biol. Chem. 270*, 2443-2446.

52. Ollagnier, S., Mulliez, E., Gaillard, J., Eliasson, R., Fontecave, M. & Reichard, P. (1996) The anaerobic *Escherichia coli* ribonucleotide reductase. Subunit structure and iron sulfur center. *J. Biol. Chem. 271*, 9410-9416.

53. Broderick, J. B., Duderstadt, R. E., Fernandez, D. C., Wojtuszewski, K., Henshaw, T. F. and Johnson, M. K. (1997) Pyruvate formate-lyase activating enzyme is an iron-sulfur protein. *J. Am. Chem. Soc. 119*, 7396-7397.

54. Kulzer, R., Pils, T., Kappl, R., Huttermann, J. & Knappe, J. (1998) Reconstitution and characterization of the polynuclear iron-sulfur cluster in pyruvate formate-lyase-activating enzyme. Molecular properties of the holoenzyme form. *J. Biol. Chem. 273*, 4897-4903.

55. Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res. 29*, 1097-1106.

56. Krebs, C., Broderick, W. E., Henshaw, T. F., Broderick, J. B., and Huynh, B. H. (2002) Coordination of adenosylmethionine to a unique iron site of the [4Fe-4S] of pyruvate formate-lyase activating enzyme: A Mössbauer spectroscopic study. *J. Am. Chem. Soc. 124*, 912-913.

57. Cosper, N. J., Booker, S. J., Ruzicka, F., Frey, P. A., and Scott, R. A. (2000) Direct FeS cluster involvement in generation of a radical in lysine 2,3-aminomutase. *Biochemistry. 39*,15668-15673.

58. Cosper, M. M., Jameson, G. N. L., Davydov, R., Eidsness, M. K., Hoffman, B. M., Huynh, B. H., and Johnson, M. K. (2002) The [4Fe-4S] $^{2+}$ cluster in reconstituted biotin synthase binds S-adenosyl-L-methionine. *J. Am. Chem. Soc. 124,* 14006-14007.

59. Frey, P. A. (2001) Radical mechanisms of enzymatic catalysis. *Annu. Rev. Biochem. 70,* 121-148.

60. Walsby, C. J., Ortillo, D., Yang, J., Nnyepi, M. R., Broderick, W. E., Hoffman, B. M. & Broderick, J. B. (2005) Spectroscopic approaches to elucidating novel iron-sulfur chemistry in the "radical-SAM" protein superfamily. *Inorg. Chem. 44,* 727-741.

61. Broderick, J. B., Henshaw, T.F., Cheek, J., Wojtuszewski, K., Smith, S.R., Trojan, M.R., McGhan, R. M., Kopf, A., Kibbey, M. and Broderick, W. E. (2000) Pyruvate formate-lyase-activating enzyme: strictly anaerobic isolation yields active enzyme containing a [3Fe-4S](+) cluster. *Biochem. Biophys. Res. Commun. 269,* 451-456.

62. Krebs, C., Henshaw, T. F., Cheek, J., Huynh, B. H. and Broderick, J. B. (2000) Conversion of 3Fe-4S to 4Fe-4S Clusters in Native Pyruvate Formate-Lyase Activating Enzyme: Mössbauer Characterization and Implications for Mechanism. *J. Am. Chem. Soc. 122,* 12497-12506.

63. Petrovich, R. M., Ruzicka, F. J., Reed, G. H. & Frey, P. A. (1992) Characterization of iron-sulfur clusters in lysine 2,3-aminomutase by electron paramagnetic resonance spectroscopy. *Biochemistry. 31,* 10774-10781.

64. Duin, E. C., Lafferty, M. E., Crouse, B. R., Allen, R. M., Sanyal, I., Flint, D. H. & Johnson, M. K. (1997) [2Fe-2S] to [4Fe-4S] cluster conversion in *Escherichia coli* biotin synthase. *Biochemistry. 36,* 11811-11820.

65. Tse Sum Bui, B. F., D., Marquet, A., Benda, R. & Trautwein, A. X. (1999) Mössbauer studies of *Escherichia coli* biotin synthase: evidence for reversible interconversion between [2Fe-2S](2+) and [4Fe-4S](2+) clusters. *FEBS Lett. 459,* 411-414.

66. Ugulava, N. B., Gibney, B. R. & Jarrett, J. T. (2001) Biotin synthase contains two distinct iron-sulfur cluster binding sites: chemical and spectroelectrochemical analysis of iron-sulfur cluster interconversions. *Biochemistry. 40,* 8343-8351.

67. Ugulava, N. B., Sacanell, C. J. & Jarrett, J. T. (2001) Spectroscopic changes during a single turnover of biotin synthase: destruction of a [2Fe-2S] cluster accompanies sulfur insertion. *Biochemistry. 40,* 8352-8358.

68. Ugulava, N. B. S., K. K. & Jarrett, J. T. (2002) Evidence from Mössbauer spectroscopy for distinct [2Fe-2S](2+) and [4Fe-4S](2+) cluster binding sites in biotin synthase from *Escherichia coli*. *J. Am. Chem. Soc. 124,* 9050-9051.

69. Tse Sum Bui, B., Benda, R., Schünemann, V., Florentin, D., Trautwein, A. X. and Marquet, A. (2003) Fate of the (2Fe-2S)(2+) cluster of *Escherichia coli* biotin synthase during reaction: a Mössbauer characterization. *Biochemistry. 42, 8*791-8798.

70. Busby, W. B., Schelvis, J. P. M., Yu, D. S., Babcock, G. T., and Marletta, M. A. (1999) Lipoic acid biosynthesis: LipA is an iron-sulfur protein. *J. Am. Chem. Soc. 121, 4*706-4707.

71. Ollagnier-de Choudens, S. & Fontecave, M. (1999) The lipoate synthase from *Escherichia coli* is an iron-sulfur protein. *FEBS Lett. 453, 2*5-28.

72. Cicchillo, R. M., Lee, K. H., Baleanu-Gogonea, C., Nesbitt, N. M., Krebs, C. & Booker, S. J. (2004) *Escherichia coli* lipoyl synthase binds two distinct [4Fe-4S] clusters per polypeptide. *Biochemistry. 43,* 11770-11781.

73. Layer, G., Moser, J., Heinz, D. W., Jahn, D. and Schubert, W. D. (2003) Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J. 22, 6*214-6224.

74. Ollagnier, S., Mulliez, E., Schmidt, P. P., Eliasson, R., Gaillard, J., Deronzier, C., Bergman, T., Graslund, A., Reichard, P. & Fontecave, M. (1997) Activation of the anaerobic ribonucleotide reductase from *Escherichia coli*. The essential role of the iron-sulfur center for S-adenosylmethionine reduction. *J. Biol. Chem. 272, 2*4216-24223.

75. Mulliez, E., Padovani, D., Atta, M., Alcouffe, C. & Fontecave, M. (2001) Activation of class III ribonucleotide reductase by flavodoxin: a protein radical-driven electron transfer to the iron-sulfur center. *Biochemistry. 40, 3*730-3736.

76. Blaschkowski, H. P., Neuer, G., Ludwig-Festl, M. & Knappe, J. (1982) Routes of flavodoxin and ferredoxin reduction in Escherichia coli. CoA-acylating pyruvate: flavodoxin and NADPH: flavodoxin oxidoreductases participating in the activation of pyruvate formate-lyase. *Eur. J. Biochem. 123, 5*63-569.

77. Padovani, D., Thomas, F., Trautwein, A. X., Mulliez, E. & Fontecave, M. (2001) Activation of class III ribonucleotide reductase from *E. coli*. The electron transfer from the iron-sulfur center to S-adenosylmethionine. *Biochemistry. 40, 6*713-6719.

78. Henshaw, T. F., Cheek, J. and Broderick, J.B. (2000) The $[4Fe\text{-}4S]^{1+}$ Cluster of Pyruvate Formate-Lyase Activating Enzyme Generates the Glycyl Radical on Pyruvate Formate-Lyase: EPR-Detected Single Turnover. *J. Am. Chem. Soc. 122, 8*331-8332.

79. Boquet, P., Munro, P., Fiorentini, C. and Just, I. (1998) Toxins from anaerobic bacteria: specificity and molecular mechenisms of action. *Curr. Opin. Microbiol. 1, 6*6-74.

80. Skerra, A. (1994) Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene. 151,* 131-135.

81. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72,* 248-254.

82. Pierik, A. J., Wolbert, R. B. G., Mutsaers, P. H., Hagen, W. R. & Veeger, C. (1992) Purification and biochemical characterisation of a putative [6Fe-6S] prismane-cluster-containing protein from *Desulfovibrio vulgaris* (Hildenborough). *Eur. J. Biochemistry. 206,* 679-704.

83. Paulsen, I. T., Brown, M. H. & Skurray, R. A. (1996) Proton-dependent multidrug efflux systems. *Microbiol Rev. 60,* 575-608.

84. Pao, S. S., Paulsen, I. T. & Saier, M. H. Jr. (1998) Major facilitator superfamily. *Microbiol. Mol. Biol. Rev. 62,* 1-34.

85. O'Connor, L., Coffey, A., Daly, C. & Fitzgerald, G. F. (1996) AbiG, a genotypically novel abortive infection mechanism encoded by plasmid pCI750 of *Lactococcus lactis sub sp. cremoris* UC653. *Appl. Environ. Microbiol. 62,* 3075-3082.

86. Pei, J. G., N. V. (2001) Type II CAAX prenyl endopeptidases belong to a novel superfamily of putative membrane-bound metalloproteases. *Trends Biochem. Sci. 26,* 275-277.

87. Diep, D. B., Havarstein, L. S. & Nes, I. F. (1996) Characterization of the locus responsible for the bacteriocin production in *Lactobacillus plantarum* C11. *J. Bacteriol. 178,* 4472-4483.

88. Mayhew, S. G. (1978) The Redox Potential of Dithionite and $SO_2$ from Equilibrium Reactions with Flavodoxins, Methyl Viologen and Hydrogen plus Hydrogenase. *Eur. J. Biochem. 85,* 535-547.

89. Zehnder, A. J. B. W., K. (1976) Titanium(III) Citrate as a Nontoxic Oxidation-Reduction Bufferring System for the Culture of Obligate Anaerobes. *Science. 194,* 1165-1166.

90. Buckel, W. & Golding, B. T. (1999) Radical species in the catalytic pathways of enzymes from anaerobes. *FEMS Microbiol. Rev. 22,* 523-541.

91. Frey, P. A. (2001) Radical Mechanisms of Enzymatic Catalysis. *Annu. Rev. Biochem. 70*, 121-148.

92. Selmer, T., Pierik, A. J. & Heider, J. (2005) New glycyl radical enzymes catalysing key metabolic steps in anaerobic bacteria. *Biol. Chem. 386,* 981-988.

93. Lehtiö, L. & Goldman, A. (2004) The pyruvate formate lyase family: sequences, structures and activation. *Protein Eng. Des. Sel. 17,* 545-552.

94. Sawers, G. & Watson, G. (1998) A glycyl radical solution: oxygen-dependent interconversion of pyruvate formate-lyase. *Mol. Microbiol. 29, 945*-954.

95. Fontecave, M. (1998) Ribonucleotide reductases and radical reactions. *Cell Mol. Life Sci. 54, 68*4-695.

96. Raynaud, C., Sarcabal, P., Meynial-Salles, I., Croux, C. & Soucaille, P. (2003) Molecular characterization of the 1,3-propanediol (1,3-PD) operon of *Clostridium butyricum. Proc. Natl. Acad. Sci. U S A. 100,* 5010-5015.

97. Leuthner, B., Leutwein, C., Schulz, H., Horth, P., Haehnel, W., Schiltz, E., Schagger, H. & Heider, J. (1998) Biochemical and genetic characterization of benzylsuccinate synthase from *Thauera aromatica*: a new glycyl radical enzyme catalysing the first step in anaerobic toluene metabolism. *Mol. Microbiol. 28, 615*-628.

98. Mogk, A., Volker, A., Engelmann, S., Hecker, M., Schumann, W. & Volker, U. (1998) Nonnative proteins induce expression of the *Bacillus subtilis* CIRCE regulon. *J. Bacteriol. 180, 28*95-2900.

99. Golovanov, A. P., Hautbergue, G. M., Wilson, S. A. & Lian, L. Y. (2004) A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc. 126, 89*33-8939.

100. Ollagnier, S., Mulliez, E., Gaillard, J., Eliasson, R., Fontecave, M. & Reichard, P. (1996) The anaerobic *Escherichia coli* ribonucleotide reductase. Subunit structure and iron sulfur center. *J. Biol. Chem. 271, 94*10-9416.

101. Sun, X., Eliasson, R., Pontis, E., Andersson, J., Buist, G., Sjoberg, B. M. & Reichard, P. (1995) Generation of the glycyl radical of the anaerobic *Escherichia coli* ribonucleotide reductase requires a specific activating enzyme. *J. Biol. Chem. 270, 24*43-2446.

102. Tamarit, J., Mulliez, E., Meier, C., Trautwein, A. & Fontecave, M. (1999) The anaerobic ribonucleotide reductase from *Escherichia coli*. The small protein is an activating enzyme containing a [4fe-4s](2+) center. *J. Biol. Chem. 274,* 31291-31296.

103. Tamarit, J., Gerez, C., Meier, C., Mulliez, E., Trautwein, A. & Fontecave, M. (2000) The activating component of the anaerobic ribonucleotide reductase from *Escherichia coli*. An iron-sulfur center with only three cysteines. *J. Biol. Chem. 275, 15*669-15675.

104. Coschigano, P. W., Wehrman, T. S. & Young, L. Y. (1998) Identification and analysis of genes involved in anaerobic toluene metabolism by strain T1: putative role of a glycine free radical. *Appl. Environ. Microbiol. 64, 16*50-1656.

105. Reddy, S. G., Wong, K. K., Parast, C. V., Peisach, J., Magliozzo, R. S. & Kozarich, J. W. (1998) Dioxygen inactivation of pyruvate formate-lyase: EPR evidence for the formation of protein-based sulfinyl and peroxyl radicals. *Biochemistry. 37, 5*58-563.

106. Becker, A., Fritz-Wolf, K., Kabsch, W., Knappe, J., Schultz, S. & Volker Wagner, A. F. (1999) Structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase. *Nat. Struct. Biol. 6, 96*9-975.

107. Logan, D. T., Andersson, J., Sjoberg, B. M. & Nordlund, P. (1999) A glycyl radical site in the crystal structure of a class III ribonucleotide reductase, *Science. 283, 14*99-1504.

108. Kessler, D., Herth, W. & Knappe, J. (1992) Ultrastructure and pyruvate formate-lyase radical quenching property of the multienzymic AdhE protein of *Escherichia coli*. *J. Biol. Chem. 267,* 18073-18079.

# Acknowledgements

I would like to give my gratefully acknowledge to all the people who gave me supports and helps to my Ph.D program and my life in Germany. I also acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG).

I give my deeply thanks to my supervisor Dr. Thorsten Selmer who led me to the research field of the biology and biochemistry. With his constant guidance and encouragement, and also his patients, I was able to get through my work and finish this dissertation. He gave me all the opportunities to practice and enabled me to gain more experiences in my research field. I have learned quite a lot from his extensive knowledge, especially in molecular biology as well as biochemistry and many brilliant and creative ideas. His kind and all-aspect help has made the past years an ever-good memory in my life.

I am especially grateful to Prof. Dr. Wolfgang Buckel. He lets me to get the chance to study in this friendly and knowledgeable interdisciplinary team. And I am also greatly thankful to all the members of his team for their friendship and help. And great thanks to Prof. Dr. R. K. Thauer for allowing us to take the EPR measurements in MPI.

I am greatly thankful to Dr. Antonio Pierik for his powerful help about EPR analysis and for his kind help and good ideas in all my Ph.D works.

I would like to give my thanks to Dr. Paula Andrei for her former works in this project and Martin Blaser for our really nice cooperations.

I would also say thanks to all my Chinese friends in Marburg for their help in the past years.

At last, I want to give my great thanks to my husband, Dr. Zhihong Jia, for giving me the most patients, loving supports and encouragements to my work and my life; my parents, for their ever-loving supports and understanding during the years of my studies and work.

# Publication

**Lihua Yu**, Martin Blaser, Paula I. Darley, Antonio J. Pierik and Thorsten Selmer. Glycyl radical decarboxylases: defining the properties of a novel subclass of glycyl radical enzyme systems. In preparation.

# Curriculum Vitae

| | |
|---|---|
| 08. 02. 1975 | Born in Guanghua Village, Chifeng City, Innermongolia province, P. R. China. |
| 09. 1981 – 07. 1986 | Elementary school, Gunaghua, Innermongolia, P.R. China. |
| 09. 1986 – 07. 1991 | Middle school, Xinkaidi, Innermongolia, P.R. China. |
| 09. 1991 – 07. 1994 | High school, the First High School of Keshiketeng County, Innermongolia, P.R. China. |
| 09. 1994 – 07. 1998 | Study in the Department of Chemistry, Innermongolia Normal University, Hohhot, P. R. China.<br>BACHELOR's degree of Chemistry Education. |
| 09. 1998 -09. 1999 | Postgraduate study in the Department of Chemistry, Chemical Engineering Institute of Inner Mongolia Technical University, HohHot, Inner Mongolia, P. R. China. |
| 09. 1999-09. 2001 | Changchun Institute of Applied Chemistry of Chinese Academy of Sciences. Changchun, Jilin P. R. China<br>MASTER`s degree of Science. |
| from 15.10.2002 | Ph.D student in the research group of Dr. Thorsten Selmer, the Department of Biology, Philipps-University of Marburg, 35032 Marburg, Germany. |

# Erklärung

ich versichere, daß ich meine Dissertation

Two novel glycyl radical decarboxylase systems from *Clostridium scatologenes* and *Tannerella forsythensis*

selbständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

_____ _____

(Ort/Datum) (Unterschrift mit Vor- und Zuname)