Construction of the Genomic Library of *Proteus vulgaris* and its Screening for the Gene Encoding (2R)-Hydroxycarboxylate-Viologen-Oxidoreductase

Zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaftenan der
Fakultät für Bio- und Geowissenschaften der
Universität Karlsruhe
genehmigte

Dissertation

Von

Tayseer M. A. Abdel-Rassol

aus

Ismailia, Ägypten

2003

Tag der mündlichen Prüfung: 07.05.2003
Referent: Prof. Dr. J. Rétey
Korreferent: Prof. Dr. W. Zumft

Acknowledgments

This work was carried out during the period from October 1999 to February 2003 at the Institute of Organic Chemistry, Chair for Biochemistry, Karlsruhe University, Germany.

I would like to express my sincere gratitude to my supervisor, Prof. Dr. J. Retéy, whose expertise, understanding, and patience, added considerably to my graduate experience. I appreciate his vast knowledge and skill in many areas and his assistance during the course of this work.

My sincere gratitude is also due to *Prof. Dr. W. Zumft* for reading the manuscript and serving as an examiner.

I could not have maintained my studies as a graduate student in molecular biology without the support of my wife, Nessren. and my young children, Mohamed and Mazen. Through their innocent eyes I see the world full of joy and hope.

I have made many friends along the way. They have helped me, one way or another, in my struggle to complete my Ph.D. While not being able to name all, I would like to acknowledge a few in particular: C. Weber for his unfailing help with a smile; M. Gaby and B. Nadia for being always there and for giving hand whenever they were needed.

This is dedicated to the memory of my late father and to my mother, who instilled in their children the love of learning.

Table of contents

List of abbreviations:

Amp.	Ampicillin
APS	Ammonium persulfate
ATP	Adenosin-5'-triphosphate
bp	Basepairs
Bisacrylamide	*N, N'*-methylene bisacrylamide
BSA	Bovine Serum Albumin
BV	Benzoyl viologen
c	Circa
CIAP	Calf intestinal alkaline phosphatase
Cys	Cysteine
deion.	Deionized
dist.	Distilled
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTP	Desoxy-Nucleotide-5'-triphosphate
DSM	Deutsche Sammlung von Mikororganismen
DTE	1,4-Dithioerythritol
DTT	Dithiothreitol
Є	Molar extension coefficient
E	Extension
EC	Enzyme Commision
EDTA	Ethylenediamine-N,N,ŃŃ-tetraacetic acid
EPR	Electron paramagnetic resonance
Fig.	Figure
FPLC	Fast performance liquid chromatography
FS	AmpliTaq DNA Polymerase
h	Hour
IEF	Isoelectric Focusing
IPTG	Isopropyl-ß-thiogalactoside

Kan.	Kanamycine
Km	Michaelis-Menten constant
Min	Minute
Moco	Molypdopterin-cofactor
MOPS	3-[N/morpholino]Propanesulfonic acid
Mpa	Megapascal
MT	Metallothionein
nm	Nanometer
OD_{600}	Optical density (masured at 600 nm)
ORFs	Open reading frames
p.a.	Pro analysis
PAGE	Polyacrylamide gel-electrophoresis
Pfu	Plaque forming unit
pg	Pictogram
Phpy	Phenyl pyruvate
Polidocanol	Polyoxyethylene-9-laurylether
pp_i	Free pyrophosphate
PVDF	Polyvinylidene difluoride membranes
Py	Pyruvate
rpm	Rounds per second
SA	Specific activity
SDS	Sodium dodecyl sulfate or sodium lauryl sulfate
sec	Second
SEC	Size exclusion chromatography
st.	Sterile
TAE	Tris/Acetate/EDTA
TEMED	*N, N, N', N'*-tetramethylethylenediamine
T_m	Melting temperature
T_{opt}	Optimal hybridization temperature
Tris	Tris (hydroxymethyl) amino methane
TTM	Tetrathiomolybdate
VA	Volume activity

1. Introduction

Molybdenum is not one of the more abundant elements in the universe, nor in the earth as a whole, nor in the earth's crust. However, in the oceans, where life likely arose and still thrives, molybdenum is the most abundant of the redox-active transition metals. Therefore, it is not surprising that living systems from the simplest bacteria to higher plants and animals use molybdenum at the active centers of their redox-active enzymes.

Molybdenum is widely available to biological systems due to the solubility of its high-valent oxides in water and is found in two basic forms: as an integral component of the multinuclear M center of nitrogenases and as the mononuclear active sites of a much more diverse group of enzymes that in general function catalytically to transfer an oxygen atom either to or from a physiological acceptor/donor molecule (Hille, 1996).

Molybdenum enzymes are ubiquitous in the biosphere. They are present in anaerobic and aerobic organisms and play prominent roles in the metabolism of microorganisms, plants and animals. The mononuclear molybdenum enzymes constitute a fairly large class of enzymes that can be divided into two subcategories on the basis of the reaction catalyzed (Hille, 1996).

The first of these is a quite large family of enzymes whose members catalyze the oxidative hydroxylation of a diverse range of aldehydes and aromatic heterocycles in reactions that necessarily involve the cleavage of a C-H bond. These enzymes are properly considered hydroxylases. The second is a family made up of bacterial enzymes such as DMSO reductase and biotin-S-oxide reductase, as well as the bacterial dissimilatory (or respiratory) nitrate reductase; those periplasmic or membrane associated enzymes that function as terminal respiratory oxidases.

1.1 Molybdenum acquisition

Organisms have evolved the capacity to take up molybdenum via active transport systems, and at least in some cases, appear to have elaborated proteins that store molybdenum for use in biosynthesis. In most natural environments the form of molybdenum is the soluble anionic molybdate ion, MoO_4^{2-}.

1.2 Importance of molybdenum

This metal's importance has been discovered in recent years. Molybdenum is now considered one of our essential trace elements. It has been found to be essential in most mammals, as well as in all plants. We obtain it primarily from food, but since it is often scarce in the earth's crust and

therefore deficient in many soils, molybdenum deficiency can be a problem. In fact, it was recently discovered that molybdenum deficiency in the soil in an area of China was responsible for the highest known incidence of esophageal carcinoma over many generations. In nature, molybdenum is found as part of other metal complexes. In the soil, it serves as a catalyst to the nitrogen-fixing process; thus, decreased soil molybdenum can lead to deficient plant growth.

The body contains minute amounts, about 9 mg of molybdenum. It is found mainly in the liver, kidney, adrenal glands, bones and skin, but it is present in all tissues. It is important to several enzyme systems, most significantly that of xanthine oxidase, which supports many functions, including uric acid metabolism and mobilization of iron from the liver for body use. Molybdenum is fairly easily absorbed from the gastrointestinal tract, though it competes with copper at absorption sites. It is eliminated through the urine and the bile.

Depletions or deficiencies of molybdenum are common, and its availability in foods is decreased through soil depletion and food technology. This mineral has come to the nutritional forefront in the last decade with the recognition of its essential nature and the concern about deficiency.

1.2.1 Sources

The food levels of molybdenum depend largely on soil content. The amount in food may be increased a hundredfold with molybdenum-rich soil. In certain areas, hard water may contain some molybdenum. Soft water and refined foods contain hardly any.

Whole grains, particularly the germ, usually have substantial amounts; oat, buckwheat, and wheat germ are some examples of grains containing molybdenum. Many vegetables and legumes are also good sources; these include lima beans, green beans, lentils, potatoes, spinach and other dark leafy greens, cauliflower, peas, and soybeans. Brewer's yeast also has some, and liver and organ meats are often fairly rich in molybdenum.

1.2.2 Functions

Molybdenum is a vital part of three important enzyme systems—***xanthine oxidase, aldehyde oxidase, and sulfite oxidase***—and so has a vital role in uric acid formation and iron utilization, in carbohydrate metabolism, and sulfite detoxification as well.

In the soil and possibly in the body, as the enzyme nitrate reductase, molybdenum can reduce the production or counteract the actions of nitrosamines, known cancer-causing chemicals, especially in the colon. Found more in molybdenum-deficient soils, nitrosamines have been associated with high rates of esophageal cancer.

Xanthine oxidase helps in the production of uric acid, an end product of protein (purine) metabolism. Though an excess of uric acid is known to cause gout, recent studies show that, in

proper concentrations in the blood, it has antioxidant properties and helps protect the cells and tissues from irritation and damage caused by singlet oxygens and hydroxyl free radicals.

This protection may prevent tissue wear and aging, in addition to other free-radical diseases. Thus, uric acid has a new image as being an important part of balanced human function and not just a waste product. With its different effects, uric acid is somewhat like cholesterol in its biochemistry. As with cholesterol, it is both made in the body and obtained through the diet; some people are genetically inclined to elevated levels; and, whereas the right amount is essential to important functions, excesses can lead to problems.

Xanthine oxidase may also help in the mobilization of iron from liver reserves. Aldehyde oxidase helps in the oxidation of carbohydrates and other aldehydes, including acetaldehyde produced from ethyl alcohol. Sulfite oxidase helps to detoxify sulfurs in the body, particularly sulfites, which are used to preserve food. These potentially toxic and harmful substances can cause nausea or diarrhea and precipitate asthma attacks in sensitive individuals.

The "salad bar" syndrome is caused by sulfite sprays used on vegetables to keep them "fresh" longer. It is possible that adequate tissue levels of molybdenum keep the sulfite oxidase activity levels high enough to counteract this chemical and reduce potential symptoms. Molybdenum deficiency may be a factor in those people who are more sensitive to sulfites. Luckily, though, the use of sulfites in food preservation is being made illegal.

1.2.3 Uses

Molybdenum may help prevent anemia by helping mobilize iron, provided there are sufficient iron stores. The suggestions that it protects the teeth from dental caries and that it prevents sexual impotence are not yet supported by definitive research.

Molybdenum deficiency may reduce uric acid formation; this was not previously thought to be a problem, but it may be important to supplement molybdenum to maintain uric acid levels in midnormal range for the antioxidant function.

There are few research findings to suggest that molybdenum may play a role in preventing cancer and definitely none to support its use in cancer treatment. Adding molybdenum to the soil and diet has helped reduce the incidence of esophageal cancer in the Lin Xian area of China's Hunan Province, which had the highest incidence in the world of this deadly disease.

It is unlikely, however, that lack of molybdenum in the soil and, thus, in the diet was a direct cause of the cancer. It was probably due to the production of nitrosamines in the soil that could not be metabolized because of a deficiency in the plants' roots activity of the molybdenum enzyme, nitrate reductase. Nitrates and nitrites, such as those in hot dogs, lunch meats, and other cured

meats, also increase food levels of nitrates, which can lead to the formation of carcinogenic nitrosamines in the stomach.

Both vitamin C, which helps detoxify nitrosamine, and nitrate reductase, which needs molybdenum to function, can help reduce the levels of this carcinogenic chemical as it has done for the Chinese esophageal cancer rates secondary to low soil molybdenum. It is also possible that molybdenum can help protect the body from nitrosamine formation after consumption of foods high in nitrates or nitrites, such as lunch meats.

1.2.4 Deficiency and toxicity

Molybdenum, like most trace minerals, is required in a specific narrow range of daily intake; amounts much greater than this may be toxic. Animals given large amounts experience weight loss, slow growth, anemia, or diarrhea, though these effects may be more the result of low levels of copper, a metal with which molybdenum competes. In people who are sensitive to it, high doses of molybdenum may lead to high uric acid levels and gouty arthritis symptoms related to increased action of the enzyme xanthine oxidase.

Information about molybdenum deficiency is limited as well. Low soil levels of molybdenum lead to increased soil and plant levels of nitrates and nitrosamines, which increase risk of cancer, especially in the esophagus and stomach.

Increased sensitivity to sulfites used in foods may be related to low molybdenum and deficient sulfite oxidase enzymes. In animals, molybdenum-deficient diets seem to produce anorexia, weight loss, and decreased life span. In humans, deficiency may lead to visual problems, rapid heart rate and breathing, and depression of consciousness.

1.2.5 Requirements

As with other newly recognized trace elements, there is no specific RDA for molybdenum. The amount provided by the average diet ranges widely, from 50 - 500 µg a day. A safe and sensible amount of added molybdenum is from 150 - 500 µg for adults and 50 - 300 µg for children. Molybdenum-rich yeast may be available as an added nutrient, which usually contains a lot of other metals and B vitamins. Sodium molybdate, which recently has come on the market, can be taken by people who want more molybdenum, though intake should be limited to 500 µg daily. It is probably best to take molybdenum in a general multivitamin and to take 2 - 3 mg of copper daily as well, because of the potential copper loss with molybdenum supplementation.

Elimination of molybdenum occurs via the kidney and usually is complete within several weeks. Molybdenosis (teart) is a form of molybdenum toxicity that produces a disease in ruminants similar to copper-deficiency.

Little data are available on the human toxicity of molybdenum. A gout-like syndrome and pneumoconiosis have been associated with excessive concentrations of molybdenum, but the inadequate design of the studies prevents an adequate determination of the etiology of these effects.

In animals the terminal step in the pathway for degradation of sulfur-containing amino acids is the oxidation of sulfite to sulphate. This reaction is catalysed by the enzyme sulfite oxidase.

The enzyme contains molybdenum and a cytochrome b_5 type haem, is localized in the mitochondrial intermembrane space and transfers electrons from sulfite to cytochrome c on the inner membrane.

Three cases of genetic sulfite oxidase deficiency in humans have been reported. The three affected children displayed mental retardation, neurological abnormalities and dislocated ocular lenses.

The biochemical basis for lack of enzyme activity in each case has been studied. All three have been shown to lack the sulfite oxidase protein, but in one case this appears to be secondary to a defect in synthesis of the molybdenum cofactor. Sulfite oxidase deficiency has been produced in the rat by administration of high levels of tungsten. Sulfite oxidase-deficient animals are particularly susceptible to the toxic effects of sulfite and atmospheric sulfur dioxide.

1.3 Molybdenum in biological cycles

Molybdenum enzymes play crucial roles in biogeochemical cycles due to their ability to transform small inorganic and organic molecules and ions.

1.3.1 Sulfur metabolism

In higher organisms sulfite is formed in the detoxification pathway of cyanide. The enzyme rhodanese catalyses the conversion:

$$CN^- + S_2O_3^{2-} \rightarrow SCN^- + SO_3^{2-}$$

Thiocyanate is readily eliminated through the kidneys but sulfite is toxic and must be oxidized to sulfate by the Mo enzymes, sulfite oxidase. The conversion of sulfite, SO_3^{2-} to sulfate, SO_4^{2-}, involves a formal oxygen atom transfer as do many of the reactions of molybdenum enzymes.

The molybdenum enzymes that is coming under increasing study is dimethyl sulfoxide (DMSO) reductase, which catalyzes the deoxygenation of dimethyl sulfoxide to form dimethyl sulfide. The oxidation of the volatile dimethyl sulfide to methylsulfonic acid has been implicated in cloud formation, especially above oceans, which has been postulated as a key part of the bioregulation of temperature as earth in the Gaia hypothesis.

1.3.2 Carbon metabolism

In anaerobic methanogenic organisms the reductive fixation of CO_2 leads to a formylmethanofurane intermediate by the formylmethanofurane dehydrogenase molybdoenzyme according to the following reaction:

$$\text{methanofurane} + CO_2 + 2e^- + 2H^+ \rightarrow \text{formylmethanofuran} + H_2O$$

In other anaerobic or facultative organisms (such as *E.coli*) the enzyme formate dehydrogenase catalyzes the dehydrogenation of formate to CO_2. This reaction may serve the function of a CO_2 reductase in autotrophic organisms and thus, again a Mo enzyme stands at the entry point of a simple inorganic substrate into a major metabolic cycle.

1.4 Molybdenum copper antagonism

In the anaerobic rumen of sheep and cattle, sulfate is reduced to sulfide by sulfate-reducing bacteria. Sulfide converts the molybdate present into tetrathiomolybdate, MoS_4^{2-}, which complexes and precipitates the copper. When excess Mo is present, all of the copper (Cu) is precipitated and made unavailable to the organism. Similarly, when excess copper is present all of the molybdenum is precipitated and hence unavailable for uptake. Mechanisms for removal of copper from metallothionein by tetrathiomolybdate were examined in vivo and in vitro using the LEC rat, which accumulates Cu as metallothionein owing to the hereditary disorder of this strain.

1.5 Molybdenum cofactor enzymes

All of the molybdenum enzymes, with the exception of the nitrogenases, contain the molybdenum cofactor, which has a single molybdenum atom and the organic ligand called molybdopterin, which together constitute Moco. Except for DMSO reductase, all of the Moco enzymes also contain other prosthetic groups in addition to Moco. These include hemes (in, for example, nitrate reductase and sulfite oxidase), flavins (e.g. in xanthine oxidase, aldehyde oxidase, and nitrate reductase), and iron-sulfur centers (in, for example, xanthine oxidase, formate dehydrogenase, and carbon monoxide oxidoreductase).

Probably in all Mo enzymes there is no direct transfer of electrons between the substrate and its redox partner. In a sense the enzymes is engineered as an electrochemical cell in which one half reaction occurs at one prosthetic group (e.g., Moco, the ‘anode’ in this case) and the other half reaction occurs at another prosthetic group (e.g., flavin, the ‘cathode’). The enzyme provides the electrical connection between the electrodes. In all cases, whether the substrate reaction is an oxidation, as in xanthine oxidase, or a reduction, as in nitrate reductase, the reaction of the substrate occurs at the Moco site.

In most Moco enzymes the molybdenum shuttles between the limiting oxidation states IV and VI during turnover. The intermediate Mo (V) site is detectable by EPR (Electron paramagnetic resonance) spectroscopy in many cases. The combined EXAFS and EPR studies implicate a sulfur-containing oxo or oxo/sulfido coordination sphere for most Mo enzymes. The following Table shows some of the purified molybdoenzymes and their substrates.

Substrate	Enzyme	Literature
Aldehyde	Aldehyde oxidoreductases	[White & Simon 1992]
CO-oxide	CO-dehydrogenases	[Meyer & Rajagopalan 1984]
Nitrate	Nitrate reductases	[Adams & Mortenson 1985]
Sulfite	Sulfite oxidases	[Rajagopalan 1980]
Dinitrogen	Nitrogenases	[Shah & Brill 1977]
Arsenite	Arsenite oxidases	[Williams et al. 1986]
Chlorate	Chlorate reductases	[Oltmann et al. 1979]
Tetrathionate	Tetrathionate reductases	[Oltman et al. 1975]
Formate	Formate dehydrogenases	[Schauer & Ferry 1986]
Trimethylamine-N-Oxide	TMANO-reductases	[Yamamoto et al. 1986]
Dimethylsulfoxide	DMSO-reductases	[Bilous & Weiner 1985]
Xanthine	Xanthine oxidases	[Coughlan 1980]
Purin	Purin hydroxylases	[Mehra & Coughlan 1984]
Pyridoxal	Pyridoxal oxidases	[Courtright 1976]
Nicotin	Nicotin hydroxylases	[Dilworth 1983]
Biotinsulfoxide	Biotinsulfoxide reductases	[Campillo-Campbell & Campbell 1982]
Formylmethonofuran	FMF-dehydrogenases	[Börner et al. 1991]
Quinoline, Iso-quinolin derivate	Quinoline oxidoreductases	[Bauder, Tshisuaka & Lingens 1990]
2-oxocarboxylates	HVOR*	[Trautwein & Simon 1994]
Aldehydes	FOR**	[Adams et al. 1999]

* (2*R*)-hydroxycarboxylate-viologen-oxidoreductase

** Formaldehyde ferredoxin oxidoreductase

List of some Molybdenum-containing enzymes and their substrate representatives

1.6 Molybdenum versus tungsten

The prevalence of molybdenum in microorganisms has recently been highlighted by the finding that certain thermophilic and extremely thermophilic organisms (hyperthermophiles) use tungsten

apparently in place of molybdenum. Many of these organisms have been isolated from deep-sea or shallow hydrothermal vents.

The best characterized W-dependent enzyme is an aldehyde oxidoreductase which appears to play a role in aldehyde oxidation that is similar to that played by the Mo-containing aldehyde oxidases in mesophilic organisms. Significantly, the tungsten seems to be associated with a pterin that may be identical to the molybdopterin of molybdenum enzymes.

1.7 Why do the thermophilic organisms use tungsten rather than molybdenum?

A plausible explanation lies in the comparative electrochemistry of W and Mo compounds. Specifically, the redox potentials for W complexes are known to be more negative than those of the corresponding molybdenum complexes.

Such a low potential W site could more effectively reduce Fe-S sites within the protein and in the electron-accepting ferredoxin, which, in turn, delivers electrons to hydrogenase for the production of dihydrogen. Thus, the use of W may be preferred to allow the organism to more effectively deliver redox equivalents for dihydrogen evolution.

However, despite the attractiveness of the above hypothesis there is potentially a simpler explanation. Specifically, the deep-sea vents, for example, have been reported to have no detectable molybdenum and significant tungsten in their effluent waters. This finding is consistent with the greater affinity of Mo for sulfide in the vent effluent leading to the precipitation of MoS_3 and MoS_2 and consequent unavailability of molybdenum.

If this finding is confirmed for deep-sea vents and extended to shallow hydrothermal vents, then organisms would have no choice but to use W then Mo, if they are to occupy the ecological niche provided by the hydrothermal vents.

It has been suggested that deep-sea hydrothermal vents are prime candidate sites for the origins of life on early earth. If this is so then W enzymes may actually have preceded Mo enzymes in an evolutionary sense. Molybdenum enzymes may have been developed in response to the greater availability of Mo to mesophilic organisms in low sulfide environments.

1.8 Classification of molybdopterin containing proteins

Molybdopterin containing proteins can be classified according to their biological function, molybdopterin centre type, and number of prosthetic centers, and sequence similarity.

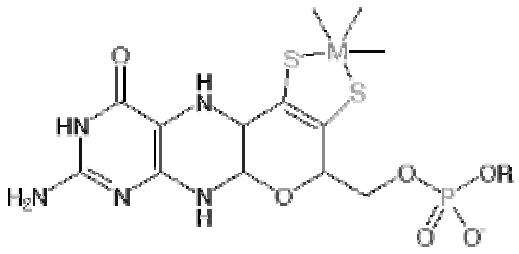
1.8.1 Molybdopterin containing proteins according to their biological function:

Oxidoreductases

- Aldehyde ferredoxin oxidoreductase
- Aldehyde oxidase
- Arsenite oxidase
- Carboxylic acid reductase
- CO dehydrogenase
- DMSO reductase
- Formaldehyde ferredoxin oxidoreductase
- Formate dehydrogenase
- Formylmethanofuran dehydrogenase
- Glyceraldehyde-3-phosphate ferredoxin oxidoreductase
- Nitrate reductases
- Polysulphide reductase
- Sulphite oxidase
- Trimethylamine oxide reductase
- Xanthine oxidase
- Xanthine dehydrogenase

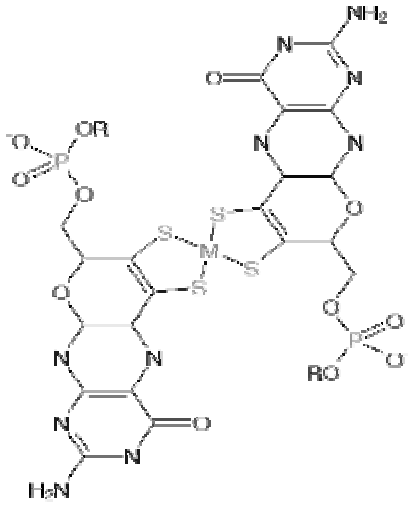
1.8.2 Molybdopterin containing proteins according to their molybdopterin centre type:

- Aldehyde ferredoxin oxidoreductase family and DMSO reductase family:



M·1molybdopterin cofactor (M = Mo or W; R = H or CMP)

- Sulphite oxidase family and xanthine oxidase family:



M·2 molybdopterin (M = Mo or W; R = H, AMP, GMP, HMP)

1.8.3 Molybdopterin containing proteins according to their type and number of prosthetic centres:

Simple	Complex
- DMSO reductase	- Moco-Fe-S-Proteins • Aldehyde ferredoxine oxidoreductase family • DMSO reductase family • Xantheine oxidase family - Moco-Fe-S-flavoproteins • Xantheine oxidase family - Moco-haem-(flavor)proteins • Sulfite oxidase family

1.8.4 Molybdopterin containing proteins in motif databases:

PRINTS ID	PRINTS AC	PROSITE/BLOCKS ID	PROSITE AC	BLOCKS AC
EUMOPTERIN	PR00407	MOLYBDOPTERIN_EUK	PS00559	BL00559
		MOLYBDOPTERIN_PROK_1 MOLYBDOPTERIN_PROK_2 MOLYBDOPTERIN_PROK_3	PS00551 PS00490 PS00932	BL00551

1.8.5 Molybdopterin containing proteins in *PROMISE*:

***PROMISE* ID**	**Description**
AOR	Aldehyde ferredoxin oxidoreductase family (including aldehyde ferredoxin oxidoreductase, formaldehyde ferredoxin oxidoreductase, glyceraldehyde-3phosphate ferredoxin oxidoreductase, carboxylic acid reductase and hydroxycarboxylate-viologen-oxidoreductase)
SULFOXIDASE	Sulphite oxidase family (including sulphite oxidase and plant and fungal assimilatory nitrate reductases)
DMSOR	DMSO reductase family (including DMSO reductase, dissimilatory nitrate reductases, formylmethanofuran dehydrogenase, trimethylamine N-oxide reductase, arsenite oxidase, formate dehydrogenase and polysulphide reductase)
XANTOXIDASE	Xanthine oxidase family (including xanthine oxidase, xanthine dehydrogenase, aldehyde oxidase, CO dehydrogenase)

1.8.6 Molybdopterin containing proteins in *Proteus* spp.:

An extensive work have been done for screening the molydenum and tungsten containing enzymes in different organisms that have been purified such as aldehyde ferredoxin oxidoreductase (AOR) from the strict anaerobic hyperthermophilic *Pyrococcus furiosus* (Mukund *et al*., 1991) and *Thermococcus* strain ES-1 respectively (Heider *et al*., 1995), Formaldehyde ferredoxin oxidoreductase (FOR) from *Pyrococcus furiosus* (Roy *et al*., 1999) *and* arsenite oxidase from *Alcaligenes faecalis* (Anderson *et al*., 1992).

Molybdenum and tungsten enzymes have received attention by Simon H. and his co-workers, who reported on the preparation of (*R*)-2-hydroxycarboxylic acids by reduction of the corresponding 2-oxo acids with resting cells of *Proteus vulgaris* and hydrogen gas or formate (Schummer *et al*., 1991). They also reported on the preparation of pyruvate from (*R*)-lactate using the resting cells of *Proteus* species (Schinschel *et al*., 1993) and the effect of medium composition on the formation of (*R*)-2-hydroxycarboxylate-viologen-oxidoreductase (HVOR) and dimethylsulphoxide reductase in *Proteus* spp. (Schinschel *et al*., 1993). Trautwein *et al*., 1994 have purified and characterized the only member of the aldehyde ferredoxin oxidoreductase (AOR) family that contains molybdenum instead of tungsten.

Compared to the voluminous literature on the purification, characterization and studying the mechanistic aspects of tungsten and molybdenum enzymes, genetic studies on this group of enzymes have received a relatively little attention.

From *P. vulgaris* genome, it is only the DNA of the Rts1 plasmid that has been isolated and characterized. It is a low-copy-number kanamycin resistence plasmid originally isolated from a clinical strain of *P. vulgaris* (Terawaki *et al*., 1967). Its molecular mass was originally estimated to be about 140 kDa (Ishihara, *et al*., 1978). It is the prototype for the T incompatibility group (Coetzee *et al*., 1972), and expresses pleiotropic thermosensitive phenotypes in autonomous replication (DiJoseph *et al*., 1974 and Terawaki *et al*., 1972), conjugative transfer (Terawaki *et al*.,

1967), host cell growth (DiJoseph *et al*., 1973 and Terawaki *et al*., 1968), and restriction of T-even phages (Ishaq *et al*., 1980, Janosi *et al*., 1994 and Yokota *et al*., 1969). The complete nucleotide sequence of Rts1 has been reported (Murata *et al*., 2002).

The genome is 217,182 bp in length and contains 300 potential open reading frames (ORFs). Among them, the products of 141 ORFs, including 9 previously identified genes, displayed significant sequence similarity to known proteins. The set of genes responsible for the conjugation function of it has been identified. Inspection of the overall genome organization revealed that the Rts1 genome is composed of four large modules, providing an example of modular evolution of plasmid genomes.

2. Aim of the work

The screening of the genome of the facultative anaerobic enterobacterium *Proteus vulgaris* for the presence of the gene encoding the (*R*)-2-hydroxycarboxylate-viologen-oxidoreductase (HVOR) protein would allow the isolation of the HVOR and the other related sequences.

The present study describes the isolation and initial characterization of several sequences related to the HVOR and the aldehyde ferredoxine oxidoreductases (AOR) from the *P. vulgaris* genome.

The presentation of the results is divided into two sections:

- ***The first*** covers the isolation and initial characterization of some of the AOR-related clones from the genomic library of the *P. vulgaris*.
- ***The second*** presents the evolutionary implications derived from the analysis of the sequences of the fished clones and their comparison with that of AORs from other organisms.

3. Materials

In the context of this work the chemicals (usually p.a. Quality), Instruments and consumed materials used were purchased, as far as different specifications did not refer, from the following companies:

3.1 Instruments

- Autoclave: - H+P	Labortechnik Varioklav
- Balance: - 1 digit PB3001	Mettler Toledo
- 3 digits	Sartorius analytic
- Centrifuges (floor): - J2-21	Beckman
- RC 5B Plus and Super T21	Sorvall
- Centrifuge (desk): - 202-MK	Sigma, München
- Biofugofresco	Heraeus Instruments
- Ultracentrifuge: - L8-60M	Beckman
- Clean Bench: - ET 130V/UV	Ehret, Emmendingen
- Evaporator: - Typ KL	Bühler Laborgeräte
- Fermentor: - model BIOFLO 3000, 10 L.	Edison, N.J., U.S.A.
- Fine Pipettes: 10 µl, 1.0 ml 40µl, 0.2 ml and 1 - 5 ml	Eppendorf Renner
- FPLC columns:	
- DEAE-S 2.3 cm X 20 cm	Pharmacia, Freiburg
- Hydroxyapatite 2.0 cm X 20 cm	Pharmacia, Freiburg
- Q-Sepharose 1.6 cm X 10 cm	Pharmacia, Freiburg
- Superdex-200 2.6 cm X 60 cm	Pharmacia, Freiburg
- FPLC: - LKB-P5000	Pharmacia, Freiburg
- Gel Elelectrophoresis- Agarose	Biometra Mini-Agarosegel-System
- Gel Electrophoresis-SDS, Mini-Protean II Cell	BioRad, München
- Gene Pulser	BioRad, München
- HPLC: L-6210 intelligent pump	Merck/Hitachi
- Incubator: G 25	New Brunswick Scientific Shaker
- Incubator (Shaking): KB 500	Heraeus
- Magnetic Shaker: MR 2000	Heidolph
- Oven: - Microwave	Privileg 2000
- Hybridization	
- Polaroid Camera	MP 4 Land Camera
- Power Supply Units: E532, E443, E722	Consort
- Pulse Controller	BioRad, München
- Refrigerator (-80 °C)	Kryotec, Hamburg
- Speed vacuum (Univapo 100 H)	Fröbel, Lindau
- Thermocycler - Express	HYBAID Limited, Ashford
- Techneq	Gesellschaft für Laborgeräte, Wertheim
- Transilluminator	Fröbel, Lindau
- Ultrasonic Transmitter	Bandelin Sonoplus
- UV/VIS Spectrophotometer	- Varian Cary 13E - Perkin Elmer Lambda 2

- Vortexer, Genie K-550-GE	Fröbel Laborgeräte
- Water bath	Julambo U3-7A
- Water bath (variable)	Gesellschaft für Labortechnik, Burgwedel

3.2 Consumed Materials

- Centricons	Millipore Ultrafree-5 und –10
- Cuvettes - for Electroporation	Peqlab & Promega
- Plastic	Hellma
- UV Quartz	Ratiolab
- Transformation	Peqlab, Erlangen
- Dialysis Membranes	Roth, Karlsruhe
- Eppendorf Microcentrifuge tubes 1.5 & 2	Sarstedt, USA
- Falcon Tubes 15 & 50 ml	Sarstedt, USA
- Glass beads, 0.25-0.5 mm	Serva, Heidelberg
- Microcentrifuge Tubes for PCR (0.5 ml)	Sarstedt, USA
- Millipore Filters 0.2 µm & 0.45 µm	Millipore Milex-GS
- Nitrocellulose Membrane Porablot 200 x 200	Macherey-Nagel
- Nylon membrane	Qiagen, Hilden
- Petri Diches 15 x 100	Sarstedt, USA
- Pipette Tips 10 µl - 1 ml	Sarstedt, USA
- Polaroid Films Type 667 (35 x 43)	Kodak BioMax MR

3.3 Chemicals

- Acrylamide & Bisacrylamide	BioRad, München
- Agar	Difco Laboratories, Michigan
- Agarose	GibcoBRL, Karlsruhe
- Ammonium acetate	Riedel-deHaen
- Ampicillin, sodium salt	Serva, Heidelberg
- Bactotrypton	Difco Laboratories, Michigan
- Bactoyeast	Difco Laboratories, Michigan
- Benzamidine	Merck, Darmstadt
- Boric acid	Roth, Karlsruhe
- Bromphenol Blue	Roth, Karlsruhe
- BSA	Sigma, München
- Canamycine (monosulfate)	ICN, Eschwege
- Casein hydrolysate	ICN, Eschwege
- Chloramphenicol	Serva, Heidelberg
- Coomassie Brillant Blue R-250	Fluka, Neu-Ulm
- DMSO	Riedel-deHaen
- Dithionite	Sigma, München
- DTT	Sigma, München
- DTE	Lancaster
- EDTA	Roth, Karlsruhe & Applichem
- Ethanol, absolute	Serva, Heidelberg
- Ethidium bromide	Serva, Heidelberg
- Gelatine	Serva, Heidelberg
- Glucose	Merck, Darmstadt
- Glycerol	Roth, Karlsruhe
- Glycin	ICN, Eschwege

- GSH-Agarose	Fluka, Neu-Ulm
- IPTG	Peqlab, Erlangen
- Kanamycine (monosulfate)	ICN, Eschwege
- Dl-lactic acid	Sigma, München
- Lysozyme	AppliChem
- Manganese sulfate	Merck, Darmstadt
- β-Mercaptoethanol	Roth, Karlsruhe
- Mineral oil	Fluka, Neu Ulm
- Polyoxyethylene-9-laurylether	Sigma, München
- K_2HPO_4/KH_2PO_4	Fluka, Neu-Ulm
- Sarcosine	Acros organics, Belgien
- Sodium chloride	Fluka, Neu-Ulm
- Sodium dithionite	Fluka, Neu-Ulm
- Sodium formate	Merck, Darmstadt
- Sodium selenite	ICN, Eschwege
- SDS	Serva, Heidelberg
- Sodium hydroxide	Fluka, Neu-Ulm
- Sodium pyruvate	Lancaster
- TEMED	BioRad, München
- Tetracycline	Serva, Heidelberg
- Tris	ICN, Eschwege
- Triton X-100	Sigma, München
- Tween 20	Roth, Karlsruhe
- Urea	ICN, Eschwege
- X-Gal	Peqlab, Erlangen

All the other not here mentioned chemicals were purchased in the chemical store of the Institute of organic chemistry, Karlsruhe University.

3.4 Markers

3.4.1 DNA markers:

1. λ-DNA-Sizer III (cut with *Eco* RI and *Hind* III)	Peqlab, Erlangen
2. DNA leiter plus	Peqlab, Erlangen
3. PeqGOLD DNA-Sizer X	Peqlab, Erlangen

3.4.2 Protein markers:

1. Premixed protein molecular weight marker (low range)	Boehringer, Mannheim
2. Precision Plus Protein Standards	BioRad, München
3. ß-lactoglobuline A	Sigma, München

3.5 Kits for molecular biology

λ-ZAP Express	Stratagene, Heidelberg
GigapackII XL Packaging Extracts	Stratagene, Heidelberg
E.Z.N.A. Plasmid Miniprep isolation kit I	Peqlab, Erlangen
Qiaquick Gel Extraction Kit	Qiagen, Hilden
Qiagen Genomic tip 100/G	Qiagen, Hilden
NucleoTrap® Nucleic Acid Purification Kit	Clontech
pecGold DNAPure^{TM}BA	Peqlab, Erlangen

3.6 Kit for DNA sequencing

ABI PRISM® BigDye™ Terminator Cycle Sequencing Kit	Perkin Elmer Applied Biosystems

3.7 Restriction enzymes

Enzyme	Recognized sequence	Manufacturer
Bam HI	5'-G↓GATCC-3'	Stratagene
Crf 142	5'-CCGC↓GG-3'	Hybaid,
Eco RI	5'-G↓AATTC- 3'	Hybaid
Eco RV	5'-GAT↓ATC- 3'	Hybaid
Hae III	5' –GG↓CC-3'	BioLabs
Hind III	5'-A↓AGCTT-3'	Hybaid
Nco I	5'-C↓CATGG- 3'	Hybaid
Nde I	5'-CA↓TATG-3'	Hybaid
Not I	5'-GC↓GGCCGC- 3'	BioLabs
Pst I	5'-CTGCA↓G-3'	Hybaid
Sac I	5'-GAGCT↓C- 3'	BioLabs
Sac II	5'-CCGC↓GG-3'	BioLabs
Sal II	5'-G↓TCGAC-3'	BioLabs
Sau 3AI	5'-↓GATC-3'	Stratagene

List of the used restriction enzymes

The used restriction enzymes as well as polymerases should be always stored at -20 °C.

3.8 Other enzymes and enzyme related substances

Alkaline phosphatase (CIAP)	Amersham, Braunschweig
Lysozyme	Sigma, München
Protease	Qiagen, Hilden
RNase	Peqlab, Erlangen
T4 DNA-Ligase	Hybaid
Taq-DNA Polymerase	Sigma, München
Pwo-Polymerase	Peqlab, Erlangen

3.9 Synthesized oligonucleotides

S. No.	Name	5'→ 3'	Length	Origin	GC%	Sequence 5' → 3'
1	Amp.1^{+*}	+	20	SK (vector)	50	GCT ATG ACC ATG ATT ACG CC
2	Amp.2^{-*}	-	20	SK (vector)	55	GGT CGA CGG TAT CGA TAA GC
3	Oligo.1	+	42	SK (vector)	45.2	TTC ACT GGC CGT CGT TTT ACA AAC GGT TGG ACA GGT AAC ATT
4	pBK1*	+	27	pBK (vector)	48.1	GGA AAC AGC TAT GAC CTT GAT TAC GCC
5	pBK2*	-	24	pBK (vector)	50	CGA CGG CCA GTG AAT TGT AAT ACG
6	T.1^{+}	+	21	HVOR N-ter.	42	AAC GGT TGG ACA GGT AAC ATT
7	N^{+}	+	20	ydhV gene	55	GGC TAA CGG TTG GAC AGG TA
8	N^{-}	-	21	ydhV gene	52.4	GTA GAA CAT CGT GAG CGA AGC
9	N.1^{+}	+	17	Clone 6	58.8	GCC TGG TGG CAG AAG TA
10	N.1^{-}	-	18	Clone 6	50	AAG CGG TGA TTA CCG TAG

11	N.2$^-$	-	15	Clone 6	60	GTT CCG GCA GTG CTT
12	N.3$^-$	-	16	Clone 6	50	CAG CGG CAT CAA CAA T
13	N.4$^-$	-	18	Clone 6	55.6	GAA CTG TCG AGC CAG TTG
14	N.5$^-$	-	23	Clone 6	56.5	GCT GCG TAC TCT CCA TCC GGA AT
15	F.1$^+$	+	23	ydhV gene	34.8	GGT TGG ACA GGT AAT ATA TTA AG
16	F.2$^+$	+	21	ydhV gene	42.9	GAA TTT GTT CGA TGA CTA CGG
17	F.3$^+$	+	20	ydhV gene	55	GTC TGC CAC GTT ACA AAG CG
18	P.340$^+$	+	19	ydhV gene	52.6	GAA GGG AAG GCG AAA TCA C
19	P.526$^-$	-	19	ydhV gene	52.6	CTT GTT CCT CTC TCT GGC T
20	P.3$^-$	-	20	ydhV gene	55	GCG AAA CAT TGC AGC GTC TG
21	P.4$^-$	-	19	ydhV gene	47.4	CA ATC TAC TGC CTG CGT AA
22	Test.1$^+$	+	20	Clone 2	50	GTT ATC TTT GCA TCC AGG CC
23	Test.1$^-$	-	20	Clone 2	50	CTG ATC CTG TTT CCT GTG TG
24	R.1$^-$	-	20	HVOR N-ter.	50	GCT CCT GTC GTC AGA TTC AT
25	R.2$^-$	-	18	HVOR N-ter.	50	GAT AGC TCC AGT CGT CAA
26	M.80.$^+$	+	20	HVOR N-ter.	40	AAY GGI TGG ACI GGI AAY AT
27	M.80.$^-$	-	20	HVOR N-ter.	40	ATR TTI CCI GTC CAI CCR TT
28	M.80.2$^+$	+	20	Cons. Seq.	46.7	ATH AAY YTN ACN ACN GGN CC
29	M.80.2$^-$	-	20	Cons. Seq.	46.7	GGN CCN GTN GTN AYR TTH AT
30	C.1$^+$	+	20	Homolog	55.3	CAA TCA TTG CAI CGG AAC AG
31	C.2$^+$	+	21	Homolog	54	TGG AAA AAG AAT CGA ATG TGG
32	C.1$^-$	-	21	Homolog	57.9	ATC ATA GAT CCA ATC GTC CCC
33	C.3$^+$	+	19	Homolog	57.9	GGC TCC GCA TCG TAC TAT G
34	C.3$^-$	-	19	Homolog	52.6	ATA GTA CGA TGC GGA GGC T
35	H.1$^+$	+	31	HVOR N-ter.	54.8	GAT CAA CGG CTG GAC AGG CAA CAT CTT GAG G
36	H.2$^+$	+	33	HVOR N-ter.	51.5	CTG CGG ATC AAT CTG ACG ACA GGA GCT ATA TCG
37	M.60.$^+$	+	18	Assoc. prot.	40.7	GCN GTN GGN GTN ATH GTN
38	M.60.$^-$	-	17	Assoc. prot.	43.1	HAC ATI ACI CCI ACI GC

***** Sequencing primers
Cons. seq.: Conserved sequences
HVOR N-ter.: N-terminus of the HVOR protein
Assoc. prot.: N-terminal of the 60 kDa associated protein

3.10 Bacterial strains

Proteus vulgaris (DSM 30115)	Gift from Prof. Dr. H. Simon
Escherichia coli K12-wild type	(DSM)
Escherichia coli JM109	Promega
Escherichia coli XLOLER	Stratagene
Escherichia coli XL10	Stratagene
Escherichia coli XL1-Blue	Stratagene
Escherichia coli XL1-Blue MRF'	Stratagene
Epicurian cloi XL1-Blue MRF' Kan Electroporation-competent cells	Stratagene

3.11 Media and buffers for molecular biology

- LB broth/L	10 g bactotryptone 10 g NaCl

	5.0 g bactoyeast add 1 l dist. H_2O, pH 7 - 7.5, autoclave
- LB agar/L	1 l LB broth 15.0 g agar pH 7 - 7.5, autoclave
- LB-Amp./L & LB-Kan./L	Short before dispensing the warm (48 °C) medium add 5 ml of 10-mg/ml-filter-sterilized ampicillin or 50 mg/l kanamycine
- NZY broth/L	5.0 g NaCl 2.0 g $MgSO_4.7H_2O$ 5.0 g yeast extract 10.0 g casein hydrolysate add 1l dist. H_2O pH 7.5, autoclave
- NZY agar/L	1 l NZY broth 15-20 g agar, pH 7.5, autoclave
- SOC-Medium/L	20.0 g trypton 5.0 g yeast extract 0.5 g NaCl Autoclave and then add: 10 ml of 1 M $MgCl_2$, 10 ml of 1 M $MgSO_4$ and 10 ml of a 2 M filter-sterilized glucose solution, Filter sterilize
- IPTG, 100mM	23.8 mg IPTG 1 ml sterile dist. H_2O, store at -20 °C
- X-Gal	25 mg X-Gal 1.25 ml DMF store at -20 °C in dark
- LB plates with Amp./IPTG/X-Gal	1l LB-amp. Separately add: - 0.5 mM of IPTG in LB pool - 80 µg/ml X-Gal in LB pool
- Ampicillin, 10^3x	100 mg amp./ml dist. H_2O, filter sterilize, store at-20 °C
- Kanamycine stock solution, 10^3x	50 mg kan./ml dist. H_2O, filter sterilize, store at -20 °C
- Tetracycline stock solution, 10^3x	25 mg tetra./ml dist. H_2O, filter sterilize, store at -20 °C

3.12 Buffers for preparing heat shock competent cells

- TFB1	100 mM RbCl 50 mM $MnCl_2$ 30 mM potassium acetate 10 mM $CaCl_2$ 15 % glycerol pH 5.8, filter sterilize
- TFB2	10 mM MOPS 10 mM RbCl 75 mM $CaCl_2$ 15 % glycerol pH 6.8 with KOH, filter sterilize

3.13 Buffers for agarose-gel electrophoresis

- TAE-buffer, 50x	242 g Tris-HCl 57.1 ml acetic acid 100 ml 0.5 M EDTA-solution complete to 1l with dist. H_2O, pH 8.0
- Sample buffer	25 mg bromphenol Blue 4 g saccharose add 10 ml dist.H_2O
- Ethidium bromide stock solution	10 mg ethidium bromide 10 ml dist. H_2O

3.14 Buffers for SDS-Polyacrylamide gel electrophoresis

- Buffer: A	36.6 g Tris-HCl 230 µl TEMED add 100 ml dist. H_2O, pH 8.8
- Buffer: B	19.2 ml of 1 M Tris-HCl, pH 6.8 1.6 ml of 10 % SDS solution 50 µl TEMED add 100 ml dist. H_2O, pH 8.8
- Solution: C	30 g acrylamide 2.4 g bisacrylamide add 100 ml dist. H_2O, pH 8.8
- Sample buffer, pH 8.0	60 mM Tris

	10 % glycine 2 % SDS 2.5 % β-mercaptoethanol 0.01 % bromophenol blue
- Electrophoresis buffer, 4x	12.0 g Tris 57.5 g lycin 4.0 g SDS add 1 l dist. H_2O
- staining solution	250 mg Coomassie G-250 50 ml acetic acid 450 ml ethanol add 1 l dist. H_2O
- Destaining solution	45 % methanol 10 % acetic acid 45 % dist. H_2O

3.15 Buffers for genome isolation

- Buffer B1	18.6 g EDTA-Na Salt 6.1 g Tris-HCl 50 ml of 10 % Tween-20 solution 50 ml Triton X-100 add dist. H_2O to 1l, pH 8.0
- Buffer B2	286.6 g guanidine HCl 200 ml of 100 % Tween-20 solution add dist. H_2O to 1l pH does not need to be adjusted
- Buffer QBT	43.8 g NaCl 10.5 g MOPS 150 ml absolute ehanol 15 ml Triton X-100 add dist. H_2O to 1l, pH 7.0
- Buffer QC	58.4 g NaCl 10.5 g MOPS 150 ml absolute ehanol add dist. H_2O to 1l, pH 7.0
- Buffer QF	73.1 g NaCl 6.1 g Tris-HCl 150 ml absolute ehanol add dist. H_2O to1l, pH 8.5
- Proteinase K stock solution	20 mg/ml in st. dist. H_2O Store at 2 – 8 °C or at -20 °C for prolonged storage

3.16 Buffers for hybridization

- Depurination solution	0.25 M HCl
- Denaturation solution	0.5 M NaOH 1.5 M NaCl
- Neutralization solution (for Southern transfer)	0.5 M Tris-HCl, pH 7.5 1.5 M NaCl
- Neutralization solution (for plaque hybridization)	1.0 M Tris-HCl, pH 7.4 1.5 M NaCl
- DIG-SSC, 20x	175.3 g NaCl 88.2 g sodium citrate add deion. H_2O to 1 l, pH 7.0
- DIG-SSC, 2x	100 ml DIG-SSC, 20x add 900 ml deion. H_2O
- DIG-SSC, 6x	300 ml DIG-SSC, 20x 700 ml deion. H_2O
- DIG-SSC, 0.2x	10 ml DIG-SSC, 20x 990 ml deion. H_2O
- Sarkosin, 10 %	1 g sarkosin 9 ml H_2O
- SDS solution, 10 %	10 g SDS add dist. H_2O to 100 ml, pH, 8.0
- DIG-Prehybridization buffer	250 ml SSC, 20x 100 ml of 10 % blocking solution 1 g sarkosin 2 ml of 10 % SDS solution add st. dist. H_2O to 1l, store at 4 °C
- DIG-Hybridization buffer	10 ml Prehybridization buffer labelled probe (variable concentrations), store at -20 °C
- DIG-buffer 1	17.6 g NaCl 23.2 g malic acid add 2 l st. dist. H_2O, pH 7.5 (with pellets of NaOH)
- DIG-buffer 2	900 ml DIG-buffer 1 100 ml DIG-blocking solution, 10 %

- DIG-buffer 3	12.1 g Tris-HCl 5.3 g NaCl 10.2 g $MgCl_2$ add 1 l st. dist. H_2O, pH 7.5
- DIG-buffer 4	1.21 g Tris-HCl 0.37 g Na_2EDTA add 1 l st. dist. H_2O, pH 8.0
- Tween 20, 10 %	4 ml Tween 20 36 ml dist. H_2O
- DIG-buffer 1/0.3 % Tween 20	970 ml DIG-buffer 1 30 ml Tween 20, 10 %
- 2x SSC/0.1 % SDS	990 ml 2x SSC 10 ml 0.1 % SDS
- 0.2x SSC/0.1 % SDS	990 ml 0.2x SSC 10 ml 0.1 % SDS
- DIG Blocking solution 10 %	25 g blocking reagent 250 ml DIG buffer 1, store at 4 °C
- DIG detection solution	20 ml DIG buffer 2 4µl Anti-DIG-AP conjugate For immediate use
- DIG color solution	10 ml DIG buffer 3 200 µl NBT/BCIP conc. Stock solution For immediate use

3.17 Buffers for common use

- 100 ml Triton X-100, 10 %	10 ml Triton X-100 90 ml dist. H_2O
- 100 ml Tween 20, 10 %	10 ml Tween 20 90 ml dist. H_2O
- Gelatine stock solution, 2%	1 g gelatine 50 ml dist. H_2O, autoclave
- SM buffer	1.2 g NaCl 0.4 g $MgSO_4.7H_2O$ 20 ml Tris-HCl, pH 8.0 1 ml of 2 % gelatine solution add dist. H_2O to 200 ml
- Tris-HCl, 1 M	121.1 g Tris-HCl 1 l dist. H_2O

	pH 8.0, autoclave
- Tris-HCl, 10 mM	400 µl of Tris-HCl, 1 M 39.6 ml dist.H_2O, pH 8.0, autoclave

3.18 Computer softwares

- HUSAR-Program package (Heidelberg Unix Sequence Analysis Resources) at the Deutschen Krebsforschungszentrums Heidelberg.
- Primer designer program for windows Version 3.0 from the Scientific & Educational Software 1994 -1996.

4. Methods & Principles

4.1 Protein Chemistry Techniques

4.1.1 Purification of HVOR

4.1.1.1 General considerations

The procedure of isolation and purification of HVOR was described by Neumann, (1985); Thanos *et al*. (1987); and Trautwein (1993) with some modifications. Due to the high degenerated nature of the N-terminus of the HVOR, it was necessary to purify it for further sequencing for obtaining more convenient amino peptide sequences from the C-terminus or from within the enzyme.

All the buffers and solutions were evacuated while boiling and cooling under a stream of N_2 or formier gas (5% H_2 / 95% N_2) to remove all O_2 that should be completely removed since the HVOR enzyme is very sensitive to even a trace amount of oxygen (Trautwein *et al*. 1994). FPLC columns were also washed up with evacuated buffers through the equilibration steps.

HVOR could only be chromatographed in the presence of the detergent (Polidocanol) and the dithionite in addition to the other components of buffers (Trautwein *et al*. 1994). Attempts to substitute the rather expensive non-ionic detergent Polidocanol by Triton X-100 were not successful.

The following chromatography columns were used starting with crude extract from 20 - 80 g wet packed cells:

DEAE-S	2.3 cm X 20 cm
Hydroxyapatite	2.0 cm X 20 cm
Q-Sepharose	1.6 cm X 10 cm
Superdex-200	2.6 cm X 60 cm

The purification has been conducted according to Trautwein (1993) with some deviations regarding the type and size of the used columns. Due to the higher sensitivity of the HVOR enzyme to oxygen (Trautwein *et al*. 1994), the selection of the fractions that have been collected for further investigations relied on the possibility of finding active fractions of considerable detectable activity.

If it were impossible to measure the activity of the selected fractions, it would be better to collect the peaks.

Polidocanol is a non-ionic detergent of great importance in the process of the purifying of the HVOR enzyme. This detergent has a considerable disturbing effect on the separation pattern of the bands (especially where IEF-gel electrophoresis is to be carried out). To reduce this disturbing effect, it was necessary to dialyze the fraction(s) of interest overnight in a suitable dialysis buffer followed by several washes with diluted acetic acid. All the separation steps were done at room temperature.

Buffer	10 mM	Dithionite (mM)	Polidocanol (%)	KCl (M)	DTE (mM)	pH
B1	Tris-HCl	-	-	-	1.0	7.0
B2	Tris-HCl	2.0	1.5	-	1.0	7.0
B3	Tris-HCl	2.0	0.8	-	1.0	7.0
B4	Tris-HCl	2.0	0.4	0.2	1.0	7.0
B5	Tris-HCl	2.0	-	0.2	1.0	7.0
B6	Tris-acetate	2.0	0.4	-	1.0	6.5

Table 4.1: Composition of the buffers used for purification of HVOR from *P.vulgaris*.

4.1.1.2 Cell material

Cells of *Proteus vulgaris* (DSM 30115) was a gift from Prof. Dr. H. Simon, Institut für Organische Chemie und Biochemie, Technische Universität München, Garching, Germany.

4.1.1.3 Growth of *P. vulgaris*

Growing and harvesting of the cells were carried out by determining pH, OD according to Kunst *et al*. 1984, using continuous culture technique in 10 L-Fermentor model: BIOFLO 3000 Batch/continuous Bioreactor, Edison, N.J., USA with a steady supplement of N_2. Pellets collected frequently in 10 h intervals by centrifugation for 20 min by 10000 rpm at 4 °C. The Pellets were then stored at -20 °C under nitrogen.

Different types of media according to Neumann, 1985, Thanos *et al*., 1987 & Trautwein, 1994 were used. Table 4.2, shows the medium composition for routine subculturing preparations, while table 4.3, indicates the components of the medium used for culturing the cells under strictly anaerobic conditions.

Substance	g x 1^{-1}
Yeast extract	**5.0**
Trypton	**5.0**
Glucose	**5.0**
K_2HPO_4	**5.0**

pH of 7.2 adjusted with NaOH

Table 4.2: Medium for routine culturing.

Substance	g x 1^{-1}
Tryptone	**5.0**
Yeast extract	**5.0**
H_2HPO_4	**5.1**
HCOONa	**1.0**
$MgSO_4.7H_2O$	**0.025**
NH_4Cl	**0.17**
$CaCl_2.2H_2O$	**0.04**
$Na_2MoO_4.2H_2O$	**0.014**
$MnSO_4.H_2O$	**0.0004**
$FeSO_4.7H_2O$	**0.0004**
***p*-aminobenzoic acid**	**0.0004**
$Na_2SeO_3.5H_2O$	**0.00026**
Biotin	**0.00002**
Sodium (*R*,*S*)-lactate	**7.2**

pH of 7.5 adjusted with NaOH

Table 4.3: Medium for induction of HVOR.

4.1.1.4 Cell lysis

20 – 80 g wet packed cells were suspended in double the amount of buffer B1 (see Table 4.1) and stirred for 15 min at room temperature. The lysis of cells was carried out with French Press Cell at a pressure of 130 Mpa. The cell wall fraction was removed by centrifugation at 10000 - 13000 rpm, 4 °C for 20 min. The sediment was then discarded and the supernatant was further used for solubilizing the enzyme from the membrane by ultracentrifugation for 3 h at 28000 - 42000 rpm, 4 °C. The precipitate containing the enzyme was warmed to room temperature.

The pellet was then resuspended in solubilization buffer B2 (see Table 4.1) and stirred at 4 °C, 500 rpm for 20 min. The enzyme was then separated from the membrane by centrifugation at 28000 – 42000 rpm, 4 °C for 90 min. The supernatant containing enzyme (membrane fraction) was then used for enzyme or stored at –20 °C under nitrogen.

4.1.1.5 Ion exchange chromatography on DEAE-S

The column was firstly equilibrated with buffer B3 untill it was made completely anerobic. The crude extract was then loaded and the column was washed with buffer B3 (see Table 4.1) untill a

constant absorption at 277 nm. The elution of the enzyme was carried out with a linear gradient of 0.1 – 0.7 M KCl. The peaks were then collected for further investigation. The column was regenerated by washing with 1 M KCl untill the run showed a stable baseline.

4.1.1.6 Separation on hydroxylapatite chromotography

The equilibration was carried out with buffer B6 (see Table 4.1). The pool of peaks from the first separation on DEAE-S was concentrated using Centricon system and the pH was readjusted again to 6.5. The probe was then loaded and the column washed with buffer B6 to read constant absorption at 277 nm. The fractions that showed higher specific activity were collected, pooled and concentrated with a Centricon system. The column was regenerated with the regeneration buffer (0.5 M Tris-HCl, pH 6.5).

4.1.1.7 Size exclusion chromatography on Q-Sepharose

The pool of peaks resulting from separation on hydroxylapatite column was loaded onto Q-Sepharose column that has been equilibrated with buffer B4 (see Table 4.1). The isoelectric elution of the enzyme was conducted with the equilibration buffer. The peaks were pooled and reconcentrated as described above.

4.1.1.8 SEC on Superdex-200

The final purification step was carried out on Superdex-200 column. The column was equilibrated with buffer B5 (see Table 4.1). The concentrated pool of the most active fractions or the peaks resulting from the separation on Q-Sepharose was loaded and the eluted fractions were collected and preserved under nitrogen at -20 °C

4.1.2 Enzyme assay

Reductase activity of the HVOR was determined at 37 °C under restricted anaerobic conditions (Thanos *et al.*, 1987). The 1 ml-glass cuvettes were tightly closed with round stoppers. A preliminary test would have been achieved for determining the appropriate concentrations of the enzyme, substrate and the artificial electron mediator since the natural one is not known yet (Simon and Günther, 1998). The enzyme in different forms was added. The start of the test was through the addition of the substrate using a Hamilton syringe. The conversion rate was estimated by measuring the extension increase or decrease of the oxidized or reduced electron mediators.

$$VA\ [\ U/ml\ ] = \frac{\Delta\ E\ [min^{-1}]\ x\ Vol.\ of\ cuvette}{2/n\ x\ \varepsilon\ [mM^{-1.\ cm-1}]\ x\ Vol.\ of\ enzyme\ [ml]\ x\ d\ [cm]}$$

The specific activity (SA) (U/mg) is the result of dividing the volume activity (VA) value on protein conc. (mg/ml). SA value indicates how many active molecules of enzymes are there.

HVOR reductase (Thanos *et al*., 1987)

The reaction mixture was composed of:

100 mM Tris-HCl pH 7.0
0.2-0.3 mM ≈ 150-250 µl reduced BV (E ≈ 2.0)
1-20 µl Enzyme
2.5 mM (Phpy) or (Py)

The Km value of phenyl pyruvate is 0.15, for pyruvate is 0.72 (Skopan, 1986) and for reduced benzyl viologen is 0.1 mM.

4.1.3 Dialysis of an enzyme preparation

For removing the detergent (polidocanol) from enzyme preparations as well as the removal of excess salts that might be found in enzyme solutions, it was necessary to carry out this step. The sample was poured into a special dialysis bag and closed tightly. The dialysis lasted overnight at 4 °C by letting the dialysis bag to swim on the surface of a dialysis buffer (5l) with gentle stirring. The buffer used was 10 mM K_2HPO_4, pH 7.2 or of 10 mM K_2HPO_4 / KH_2PO_4, pH 7.4.

4.1.4 Protein concentration

In order to concentrate large amounts of protein solutions, a stirring cell from Filtron was usually used. The cell has a nitrocellulose membrane that permits the flow of buffer and small proteins. Nitrogen gas (with a maximum pressure of 1.0 and up to 3.5 bar) was applied to enforce the buffer to penetrate the membrane. The protein solution to be concentrated was always placed on ice with constant stirring at 500 rpm until the desired volume was obtained.

For concentrating small amounts of protein solutions, the ultrafree filter units were used. The filter has a nitrocellulose membrane. It was filled in with the protein sloution and centrifuged at 4 °C, 5000 rpm for an adequate time to achieve the required concentration of the protein.

The buffer and small proteins passed through the membrane under the pressure resulted from the centrifugation. The remaining solution containing the protein of interest was then pipetted out with a fine pipet.

4.1.5 SDS-PAGE electrophoresis

4.1.5.1 General considerations

In SDS-polyacrylamide gel electrophoresis separations, migration is determined not by intrinsic electrical charge of polypeptides but rather by molecular weight. Sodium dodecylsulfate (SDS) is an anionic detergent that denatures proteins by wrapping around the polypeptide backbone. In so doing, SDS confers a net negative charge to the polypeptide in proportion to its length. When treated with SDS and a reducing agent, the polypeptides become rods of negative charges with equal "charge densities" or charge per unit length.

SDS-PAGE can resolve complex mixtures into hundreds of bands on a gel. The position of a protein along the lane gives a good approximation of its size, and, after staining, the band intensity is a rough indicator of the amount present in the sample.

There are two types of buffer systems in electrophoresis, continuous and discontinuous. A continuous system has only a single separating gel and uses the same buffer in the tanks and the gel. In a discontinuous system, a method first developed by Ornstein, 1964 and Davis, 1964, a non-restrictive large-pore gel called a stacking gel is layered on top of a separating (running) gel. Each gel layer is made with a different buffer, and the tank buffers are different from the gel buffers. Although a continuous system is slightly easier to set up and tends to have fewer sample precipitation and aggregation problems than a discontinuous system, much greater sample resolution can be obtained with a discontinuous system.

In a discontinuous system, a protein's mobility (a quantitative measure of the migration rate of a charged species in an electrical field) is intermediate between the mobility of the buffer ion in the stacking gel (leading ion) and the mobility of the buffer ion in the upper tank (trailing ion). When electrophoresis is started, the ions and proteins begin to migrate into the stacking gel. The proteins concentrate in a very thin zone, called the stack, between the leading ion and the trailing ion. The proteins continue to migrate in the stack until they reach the separating gel. In contrast, only minimal concentration effects are possible with continuous gels and proteins resolve into a zone nearly as broad as the height of the original samples in the sample wells, which results in bands that are poorly resolved.

The Laemmli system (Laemmli, 1970), a modification of those described by Ornstein, 1964 and Davis, 1964, is a discontinuous SDS system which is the most widely used electrophoretic system today. The resolution in a Laemmli-gel is excellent because the treated peptides are concentrated in a stacking gel before entering the separating gel.

The average pore size is determined by the acrylamide monomer concentration [% T for total monomer concentration (w/v)] and the concentration of cross linker (% C for percentage of the mass of monomer which is cross linker).

The most frequently used cross linker is bisacrylamide. Acrylamide monomer, when mixed with initiators (usually ammonium persulfate or riboflavin) and an accelerator (usually TEMED) will form a linear polymer, having the consistency of a viscous liquid.

The incorporation of a cross linker into the linear polymer joins the linear polymers together, side-to-side, to form a three-dimensional mesh. Average pore size is thus determined by the number of linear polymers per unit volume (a function of % T) and the frequency of interchain bridges, determined by the extent to which bisacrylamide is substituted for acrylamide (% C)(Gersten, 1996).

Coomassie Blue G-250 can be used as a colloidal dispersion. The particles of the colloid are too large to enter the interior of the gel. This results in a background gel which is perfectly clear. Silver stains have two major chemistries based on silver nitrate and silver diamine. The reaction mechanism(s) is/are still not clear.

Silver staining is usually used when the detection of nanogram levels of protein without radioactivity is to be achieved.

Procedure: (Minigel-Elektrophoresesystem from BioRad was used)

4.1.5.2 Gel casting and electrophoresis

The plates and spacers were assembled in the cassette as described by the manufacturer. The bottom of the spacers was to be accurately aligned with the bottom edge of the glass to prevent leaking. The cassette was then inserted into the stand and tightened using the off-center cams, avoiding exerting extra pressure which can damage the glass. The cassette was then leak tested with water.

Separating gel (10%) was prepared by setting up the following mixture:

625 µl	Buffer A
1,665 ml	Solution C
50.0 µl	10 % SDS
5.0 µl	TEMED
25.0 µl	10 % APS (radical starter of polymerization)
2.65 ml	st. dist. H_2O

The mixture was carefully mixed avoiding frothing and immediately poured between the plates. A lean space of c. 2 - 3 cm beneath the teeth of the well-forming comb was left for later pouring of the stacking gel. A layer of water was pipetted on top of the separation gel. The separation gel was left at room temperature to solidify; meanwhile the 5 % stacking gel was prepared as follows:
In 15 ml Falcon tube, the following components were pipetted:

0.5 ml	Buffer B
0.33 ml	Solution C
5.0 µl	TEMED
10.0 µl	10 % APS
1.2 ml	st. dist. H_2O

After the separating gel had been solidified, the water layer was carefully poured off and the stacking gel was instead poured into the space between the plates to the top. The comb was then inserted and the gel was allowed to polymerize. After the polymerization had been completed, the well-forming comb was carefully removed and the wells were washed with electrophoresis buffer. The gel cassette was then transferred from the casting stand to the upper reservoir and gently tightened. The reservoirs were then filled with the SDS-electrophoresis buffer.
The protein samples were then added to SDS sample buffer (1:1), mixed thoroughly and incubated for 2 - 5 min at 100 °C to ensure that the sample is completely complexed with SDS. The samples were then allowed to cool down to room temperature. 2 µl of 0.1 % (w/v) bromophenol blue per 25 µl was then added to each sample, mixed by pipetting. The samples were finally loaded into the wells and a current of 20 mA was applied. The separation was proceeded toward the positive electrode (top → bottom).

4.1.5.3 Staining the gel with Coomassie Blue G-250

After completed electrophoresis, the gel was removed from the cassette and fixed with an appropriate amount of the fixing solution at room temperature for 1 - 2 h with gentle shaking. The fixing solution was then decanted and replaced with filtered staining solution and left for at least 30 min at 55 °C or 0.5 - 1 min in the microwave. The staining solution was then replaced with destaining solution. The gel was incubated at room temperature on a shaker until the blue bands could be clearly seen. The destaining solution was to be changed when required. For subsequent transfer of the protein bands to nitrocellulose membrane, the gel should not be stained and fixed.

4.1.6 Gradient SDS-Polyacrylamide gel electrophoresis

During electrophoresis in gradient gels, proteins migrate until the decreasing pore size impedes further progress. Once the “pore limit” is reached, the protein banding pattern does not change

appreciably with time, although migration does not cease completely. There are three main advantages of gradient gels over linear gels:

1- The advancing edge of the migration protein zone is retarded more than the trailing edge, thus resulting in a sharpening of the protein bands.
2- The gradient in pore size increases the range of molecular weights that can be fractionated in a single gel run.
3- Proteins with close molecular weight values are more likely to separate in a gradient than a linear gel (Walker, 1984).

Procedure:

Minigel-Elektrophoresesystem from BioRad was used.

The same procedure as in the normal SDS-PAGE was also used with the exception that the separating gel was composed of several layers each having a definite concentration and prepared and poured individually. The gradient range can lie between 5 and 30. The most in use gradient lies between 5 - 20.

4.1.7 Determination of protein concentration

4.1.7.1 General considerations

The most accurate method of determining protein concentration is probably acid hydrolysis followed by amino acid analysis. Most other methods are sensitive to the amino acid composition of the protein and absolute concentrations cannot be obtained (Waterborg & Matthews, 1984).

In electrophoresis gels, there is an appreciable increase in the resolution of proteins and peptides as the amount of sample applied to the gel. The amount of sample loaded depends on the detection method. Wherever, the detection is to be done by staining, colorimetric methods are used. When detection by more sensitive methods is to be used and less protein is loaded into gel, fluorometric or amplified methods are available. In analytical experiments where the purpose is to resolve the components of a mixture, the strategy is to load as much protein as possible without allowing the bands or spots to merge. If the object is to assess the purity of a particular preparation, the strategy is to overload the gel in order to detect the presence of any potential contaminants (Gersten, 1996).

4.1.7.2 Estimation of protein concentration by Bradford technique

This method makes use of the fact that the dye, Coomassie Brilliant Blue G-250 (also called Serva Blue G-250 and Xylene Brilliant Cyanine G), undergoes a spectral shift in its absorption maximum from 465 to 595 nm when bound to the free amino groups of the protein.

Procedure:

The protein estimation was carried out according to the modified method of Read & Northcote (1981) as follows:

25 mg of Serva Blue G-250 was dissolved in 25 ml 85 % H_3PO_4 and 12 ml absolute ethanol. The mixture was stirred until the dye dissolved completely. Sterile distilled water was then added to the dye solution to a final volume of 250 ml. Dye solution was then filtered using 3 MM filter papers.

A stock solution of Bovine Serum Albumin (BSA) was prepared by dissolving 1 mg/ml of BSA in 0.1 % SDS solution. The solution was then divided into 1 ml aliquots in 1.5 ml microcentrifuge tubes and preserved at -20 °C until later use.

A standard curve was constructed by preparing 1 : 10 (0.1µg/1µl BSA) and 1 : 100 (1µg/µl BSA) dilutions of the BSA stock solution in sterile distilled water. Aliquots containing 1, 2, 4, 7, 10, 15 and 20 µg were completed with sterile distilled water to 100 µl final volume.

100 µl sterile distilled water served as a blank. To each microcentrifuge tube 900 µl Coomassie Brilliant Blue solution was added and the tubes mixed by vortexing. The tubes were then incubated at room temperature for 5 - 30 min. The optical density was read at 595 nm and the standard curve plotted (absorbance at 595 nm versus conc. of BSA dilutions).

4.1.7.3 Determination of protein concentration of an unknown sample

The protein sample was diluted (where it was necessary) with sterile distilled water to a final volume of 100 µl. 900 µl of the dye solution was then added and mixed gently. The absorbance was then measured at 595 nm. The reading was then plotted on the previously prepared standard curve and the protein concentration was calculated regarding the initial dilution factor.

4.1.7.4 Protein estimation according to Warburg & Christian, 1942.

For measuring the protein concentration with this method, a protein sample was diluted and the absorbance at 260 nm and at 280 nm was measured. As a blank, a buffer solution was used. The protein concentration was then calculated as follows:

Protein conc. (mg/ml) = $1.5 \times A_{280} - 0.75 \times A_{260}$

4.1.8 Isoelectric focussing

4.1.8.1 General considerations:

Isoelectric focussing takes place in a pH gradient and is limited to molecules which can be either positively or negatively charged (amphoteric molecules), like proteins, enzymes and peptides. Separation happens in a pH gradient which is formed by special amphoteric buffers (ampholytes)

having high buffer capacities at their pI (isoelectric point). The pH gradient is produced by an electric field. Before an electric field is applied, the gel has a uniform pH value and almost all the carrier ampholytes are charged. When an electric field is applied, the negatively charged ampholytes move towards the anode, the positively charged ones to the cathode and their velocity depend on the magnitude of their net charge. The carrier ampholytes align themselves in between the cathode and the anode according to their pI, and determine the pH of their environment. As the net charge of a protein depends in part on the environmental pH, it follows that, for any given protein, there is a pH, at the isoelectric point (pI), where the net charge of the protein is zero. That is to say, it will not migrate in an electric field. Isoelectric focussing, therefore, seeks to create a pH gradient across the gel, in which the proteins will cease migration at the position in the gel where the pH corresponds to their pI. A stable gradually increasing pH gradient depending on the initial mixture of ampholytes is formed.

Strips of filterpaper soaked in electrode solutions serve the purpose of stabilizing the gradient. An acid and a base are used as anolyte and catholyte respectively. When, for example, an acid carrier ampholyte reaches the anode, it aquires a positive charge from the medium and is attracted back towards the cathode.

Gels are usually made of polyacrylamide or agarose. A number of amphoteric buffer solutions and premade gels are avilable covering broad and narrow pH-ranges. High resolution is obtained when narrow pH-ranges are employed. In polyacrylamide gels pore size can be accurately controlled by the total acrylamide concentration and degree of crosslinking. When crosslinking is kept constant, and total concentration is increased, pore size will decrease (and diffusion will be reduced). Gel solution is made from appropriate amounts of acrylamide (~ 5 %) ampholyte (c. 2 %) double destilled H_2O and riboflavin 5´. Gels (approx. 250 x 120 x 1 mm) are mold between two glass-plates and polymerized overnight in UV-light (requires a pH in the solution of 5 - 6).

Samples are usually applied cathodically (approx 3 – 5 mm from electrode strip) on the gel, but in many cases **its necessary to find the optimum spot at which the protein penetrates the gel without any trouble, do not aggregate or are unstable**. Prefocussing of the gel allow the pH gradient partly to be established prior to application. In isozyme studies, the enzymes will (when applied at the cathode) migrate towards the anode until they reach the point where their net charge is zero. In case of diffusion to an adjacent pH-environment they will rapidly aquire a charge and move back again.

Whenever the molecules of interest either is damaged at some pH between the application point and pI, or pI is outside the pH-range of the gel, good results can be achieved by stopping the process at an earlier stage. This requires a quick and even application of samples, so different migration distances is not an artefact of the application method.

Procedure:

The eluted fractions were firstly run on SDS-PAGE. Those fractions which contain protein in the range of 70-80 kDa were collected in a pool for further purification. The SDS-gel was carried out according to the discontinuous system. 5 % acrylamide stacking gel and 12 % acrylamide separating gel were used. At the end of the purification procedure, a homogenous protein was obtained. The molecular weight of this protein was detected on 4 – 20 % gradient gel. The size of this purified protein was about 75 kDa. For further identification of this purified protein, an isoelectric focussing in different pH-gradients was conducted. The IEF value of 4.9 and 5.1 was previously detected by Neumann, 1985 and Trautwein, 1993 respectively.

The IEF-gel (80 x 80 x 0.3 mm) was prepared as follows:
To 5 ml acrylamide / bisacrylamide stock-solution, the following solutions were added:
0.5 ml ampholyte, 0.25 ml glycerol, 5 µl TEMED. To polymerize the gel, 35 µl solution of 20 % APS stock solution was added. The gel was allowed to polymerise on a silanized plastic backing sheet at room temperature. The gel was then placed on the horizontal cooling platen. Before loading the samples, the system was prefocused at 1000 V (c. 15 mA) for 1 h. The samples were then loaded into the sample wells, c. 2 cm apart from the cathode. As a marker, ß-lactoglobuline A, which has an isoelectric point of 5.1 was used. Focussing of the gel was carried out at a voltage gradient ranging from 100 V to 1000 V for 1 h. The gel was then fixed and stained as described in section 4.1.5.3.

4.2 Molecular Biology Techniques

4.2.1 Preparing the host strains

The stored cells (at -80 °C) (XL1-Blue MRF strain and XLOLR strain) were revived by scraping off and streaking splinters of solid ice with a sterile wire loop on LB-tetracycline (12.5 μg/ml) plates. After the incubation overnight at 37 °C the plates were sealed with parafilm and stored at 4 °C and restreaked weekly.

4.2.2 Preparing a –80 ºC bacterial glycerol stock

Cells were streaked onto a fresh LB-tetracycline (12.5 μg/ml) plate as described in preparing the host strains. In a sterile 50-ml conical flask, 10 ml of LB-tetracycline medium was inoculated with one colony from the plate. The cells were then grown to late log phase (10 - 12 hours). 4.5 ml of a sterile glycerol-liquid medium solution (5 ml of glycerol + 5 ml of LB-medium) was added to the bacterial culture from the previous step and mixed. Aliquots in sterile microcentrifuge tubes (1ml/tube) were stored at –80 °C.

4.2.3 Construction of a genomic library

4.2.3.1 General considerations

A common starting point for cloning a specific gene is the construction of a gene library. This is a collection of recombinant clones each of which carries a different piece of DNA from the organism of interest. The number of clones needed for a complete genomic library is a function of the size of the genome and the average size of the fragment being cloned (Dale, 1994).

In general, to construct a genomic library a DNA sample would prepare, fragment randomly and the mixture of pieces should be joined with an appropriate vector. This will produce a vast number of different sorts of recombinant molecules, each having a different piece of DNA. This mixture is used for transformation or in vitro packaging in order to introduce it into a suitable host bacterium. If a plasmid or cosmid vector is used, each cell that receives the plasmid is capable of growing up into a colony, which can be isolated and purified, constituting a clone. For a phage vector, the analogous procedure of obtaining pure isolates is known as plaque purification (Dale, 1994).

4.2.3.2 Preparing of the genomic DNA of *Proteus vulgaris*

Two different types of kits namely Qiagen Genomic Tip 100/G from Qiagen and pecGold DNA Pure^{TM}BA from PeQLab were used for isolating and purifying the genomic DNA from *P. vulgaris*. The Qiagen procedure applies optimized buffer systems for careful lyses of cells and/or nuclei,

followed by binding of genomic DNA to anion-exchange resin under appropriate low-salt and pH conditions. RNA, proteins, dyes, and low molecular-weight impurities are removed by a medium-salt wash. Genomic DNA is then eluted in a high-salt buffer, concentrated and desalted by isopropanol precipitation.

The procedure from PeQLab depends on the lyses of bacterial cells with a lysozyme- containing solution. DNA as well as Plamid-DNAs can be extracted from the lysate by adding a second solution. DNA is then precipitated by adding 100 % ethanol followed by washing with 75 % ethanol. DNA elution in 1 mM EDTA, 10 mM Tris-HCl, pH 7 – 8 or in deionized H_2O is recommended.

Procedure: (According to the Qiagen's instructions with slight modifications)
Cells from 3.0 - 6.0 ml of an overnight culture of *P. vulgaris* were pelleted by centrifugation at 5000 rpm for 10 min. The supernatant was then discarded and the pellet was resuspended in 3.5 ml of buffer B1 by vortexing. Lysozyme stock solution (100 mg/ml) as well as 100 µl of proteinase K stock solution was added and the mixture was incubated at 37 °C for 1h or longer if the suspension was not homogeneous after vortexing. During incubation, the lysozyme enzymatically breaks down the bacterial cell wall, while the detergents in buffer B1 ensures complete lysis of the bacteria. RNA is degraded by RNase added to buffer B1.

Following incubation, 1.2 ml of buffer B2 was added and the mixture was mixed by inverting the tube several times. The mixture was then incubated at 50 °C for at least 30 min or till the lysate became clear. Alternatively centrifugation for 10 min at 5000 rpm, 4 °C was sometimes applied to precipitate the particulate matter. Buffer B2 and Proteinase K denature proteins such as nucleases and DNA-binding proteins and strip the genomic DNA of all bound proteins. The sample should be loaded onto the Genomic-tip promptly to prevent clogging of the Genomic-tip.

The isolation of genomic DNA from *P. vulgaris* was prepared according to Qiagen protocol as follows: (it was advisable to take aliquots for an analytical gel**)**

Genomic-tip 100/G (a midi-prep volumes) was equilibrated with 4 ml of buffer QBT and allowed to empty by gravity flow which begins automatically by reduction in surface tension due to the presence of detergent (0.15 % Triton X-100) in the equilibration buffer.

The sample was then vortexed before applying it to the equilibrated Genomic-tip and allowed to flow by gravity flow. Flow rate depends on the sample source, the number of cells and the genome size. Occasionally it should be necessary to dilute the lysate with an equal volume of buffer QBT prior to loading to Genomic-tip to prevent blocking of the Genomic-tip.

2 x 7.5 ml washes with QC buffer were sufficient to remove all contaminants that may be found in the DNA preparation. After washing the Genomic-tip with QC buffer the genomic DNA was eluted with 5 ml of buffer QF (prewarmed to 50 °C) and the Genomic-tip was allowed to drain by gravity flow.

DNA was then precipitated by adding 3.5 ml of isopropanol at room temperature. The precipitated DNA was recovered by immediate centrifugation at 13000 rpm for 20 min at 4 °C. The supernatant was then carefully removed and the DNA pellet washed by vortexing with 2 ml of ice cold 70 % ethanol, followed by centrifugation at 13000 rpm for 10 min at 4 °C. The pellet was air dried and resuspended in a suitable amount of EDTA free buffer or in deionized sterile water and allowed to dissolve on a shaker overnight. Alternatively and for immediate use the DNA was dissolved on a shaking water bath at 55 °C for 1 - 2 h.

4.2.3.3 Determining the yield, purity and length of the DNA

Yield of genomic DNA was determined spectrophotometrically by measuring the absorbance at 260 nm (absorbance readings should fall between 0.1 and 1.0); while the purity was determined by calculating the ratio of absorbance at 260 nm to absorbance at 280 nm (a pure DNA has an A_{260}/A_{280} ratio of 1.7-1.9). The length of genomic DNA was determined on 1% agarose gel.

4.2.3.4 Precipitation of nucleic acids

4.2.3.4.1 General considerations

During the course of a cloning project in many occasions, occurs the necessity to purify, concentrate nucleic acid samples or to change the solvent in which a nucleic acid is dissolved. Fulfillment of these requirements is met by nucleic acid precipitation techniques (Berger & Kimmel, 1987). Most nucleic acids may be precipitated by the addition of monovalent cations and 2 - 3 vol. of cold 95% ethanol, followed by incubation at 0 to -70 °C. The DNA or RNA may then be pelleted by centrifugation at 10000 - 13000 rpm for 15 - 20 min at 4 °C. A subsequent wash with 70% ethanol followed by brief centrifugation removes residual salt.

Procedure:

A definite concentration of DNA was mixed with 0.1 vol. of 3 M sodium acetate, pH 4.5 (sometimes 5 M ammonium acetate, pH 7.4 was applied when it was mentioned for special purposes). Two vol. of ice cold 95 % ethanol were then added and the tube incubated at -70 °C for at least 30 min, or at -20 °C overnight. The DNA was pelleted by centrifugation at 13000 rpm, 4 °C for 20 min. The supernatant was then carefully discarded and the pellet washed with 500-700 µl ice cold 70 % ethanol. The DNA was pelleted as described in the previous step and the DNA was

air dried at room temperature without letting the pellet dry completely (which if it were to happen, the DNA may become partially resistant to certain restriction enzymes) (Robertson, 1989). The pellet was finally dissolved in an appropriate vol. of deionized water or in 10 mM Tris-HCl buffer, pH 7.0-7.5.

4.2.3.5 Partial digestion of genomic DNA with restriction enzymes

Restriction enzyme digestion is performed by incubating double-stranded DNA molecules with an appropriate amount of restriction enzyme, in its respective buffer as recommended by the supplier, and at the optimal temperature for the specific enzyme. The optimal sodium concentration in the reaction varies for different enzymes. Typical digestion includes a unit of enzyme per microgram of starting DNA. One enzyme unit (depending on the supplier) is defined as ***the amount of enzyme needed to digest one microgram of double-stranded DNA completely in 1 h at the appropriate temperature*** (Roe *et al*., 1996).

The only method by which DNA can be fragmented in truly random fashion, irrespective of its base composition and sequence, is mechanical shearing. However, DNA prepared in this fashion requires several additional enzymatic manipulations (repair of termini, methylation) to generate cohesive termini compatible with those of the vectors used to generate genomic DNA libraries (Maniatis *et al*., 1978). On the other hand, partial digestion with restriction enzymes that recognize frequently occurring tetranucleotide sequences yields a population of fragments that is close to random and yet can be cloned directly (Sambrook *et al*., 1989).

Procedure:

In a 1.5 ml microcentrifuge tube the following mixture was set up as follows:
270 µl genomic DNA (20-30 ng /µl)
30 µl 10x restriction buffer

The mixture was dispensed into 10 x 25 µl equal aliquots in addition to 1 aliquot containing 49 µl. 1 µl of restriction endonuclease enzyme (*Sau* 3AI) (4 U/µl) was added to tube 1, mixed thoroughly and placed on ice. Half (25 µl) the amount of tube 1 was then transferred into tube 2 and mixed gently and replaced on ice. The similar serial dilutions performed throughout the assay tubes. Thereby, the enzyme concentration would be reduced from vial to vial to half the concentration of the previous one. Tube No. 10 served as control.

Tubes were then incubated at 37 °C for exactly 1 h and the enzyme inactivated by heating at 65 °C for 10 min in a water bath.

4.2.3.6 Size fractionation of fragments on agarose gel

4.2.3.6.1 General considerations

Agarose gel electrophoresis is employed to check the progression of a restriction enzyme digestion, to quickly determine the yield and purity of a DNA isolation or PCR reaction, and to size fractionate DNA molecules which then could be eluted from the gel. Prior to gel casting, dried agarose is dissolved in an appropriate buffer by heating and the warm gel solution then is poured into a mold which is fitted with a well-forming comb. The percentage of agarose in the gel varied. Although 0.7 % agarose gels typically is used, in cases where the accurate size fractionation of DNA molecules smaller than 1 kb is required, a 1, 1.5, or 2 % agarose gel is prepared, depending on the expected size(s) of the fragment(s).

Ethidium bromide is included in the gel matrix to enable fluorescent visualization of the DNA fragments under UV light. Agarose gels are submerged in electrophoresis buffer in a horizontal electrophoresis apparatus. The DNA samples are mixed with gel tracking dye and loaded into the sample wells. Electrophoresis usually is at 150 - 200 mA for 0.5 - 1 h at room temperature, depending on the desired separation.

When low-melting agarose is used for preparative agarose gels, electrophoresis is at 100 - 120 mA for 0.5 - 1 h, again depending on the desired separation.

Size marker is co-electrophoresed with DNA samples when appropriate for fragment size determination. After electrophoresis, the gel is placed on a UV light box and a picture of the fluorescent ethidium bromide-stained DNA is taken with a Polaroid camera (Studier, 1973).

Procedure:

1% agarose gel was prepared, by combining 0.5 gm agarose and 50 ml deionized water in a 250 ml reagent bottle, heating in a microwave for 2 - 4 min until the agarose was completely dissolved followed by cooling down to room temperature.

2.5 µl of ethidium bromide stock solution (1mg/ml) was usually added to the gel solution, swirled to mix. The gel was then poured onto a mold tray with casting comb in place. The gel was allowed to cool down for 20 - 30 min at room temperature for solidification. The gel casting comb was then removed carefully and the gel placed in a horizontal electrophoresis apparatus.

1x TAE electrophoresis buffer was then added to the reservoir until the buffer just covered the agarose gel.One-tenth volume of 10x agarose gel loading dye should to be added to each DNA sample followed by mixing and loading into the wells.

The gel was usually electrophoresed at 100 - 150 mA until the required separation had been achieved. This usually takes 0.5 - 1 h (100 - 120 mA for low gel temperature agarose),

visualization of the DNA fragments was done on a long wave UV light box and photographed with a Polaroid camera.

4.2.3.7 DNA extraction from agarose-gel

DNA elution from agarose gel was carried out using QIAquick Gel Extraction Kit from QIAGEN. This kit uses spin-column with the selective binding properties of the silica-gel membrane. DNA adsorbs to the silica-membrane in the presence of high-salt while contaminants pass through the column. Impurities are washed away and the pure DNA is eluted with Tris buffer or water. Fragments ranging from 70 bp to 10 kb were extracted and purified from standard or low-melt agarose gels in TAE buffers.

Procedure: (using a microcentrifuge protocol)

Note: all the centrifugations were carried out at 13000 rpm, 4 °C for 0.5 - 1 min unless else was recommended.

After the separation of DNA on agarose gel, the band of interest was excised with a clean, sharp scalpel and the extra gel was removed from the gel slice. The gel slice was then weighed and 3 vol. of buffer QG to 1 vol. of gel was added to 2 ml microcentrifuge tube. The tube was then incubated at 50 °C till the gel dissolved completely. For recovering DNA fragments between < 0.5 kb and > 4 kb, 1 vol. of isopropanol was added and the tube contents were mixed. 800 µl of the gel solution was applied to the QIAquick column (maximum capacity), and centrifuged.

The excess amounts of the gel solution was simply loaded again and centrifuged. After discarding the flow-through, the traces of agarose were to be removed with a wash with 500 µl buffer QG followed by centrifugation. The DNA adsorbed to the silica-membrane was washed with 750 µl of buffer PE (supplemented with absolute ethanol).

After letting it stand for 2 - 5 min, the column was centrifuged and the flow-through was discarded. Residual ethanol from buffer PE was completely removed by an additional step of centrifugation. DNA was then eluted in an adequate amount of 10 mM Tris-Cl, pH 8.5 or deionized water. The eluted DNA as stored at -20 °C to avoid degradation.

4.2.3.8 Dephosphorylation of the insert DNA

Once fractions of the desired size have been obtained, the insert DNA may be treated with alkaline phosphatase to prevent self-ligation (Sambrook *et al.* 1989), which, if it were to occur, would reduce the cloning efficiency. If the fractionated DNA contains fragments below 12 kb in size, the possibility exists of having scrambled clones containing two inserts. When lambda FIX is being used this is not necessary since the inserts are prevented from ligating to one another following

partial fill-in of the restriction overhang. When using an insertional vector to clone a specific restriction fragment, it may not be possible to dephosphorylate the insert if the vector requires dephosphorylation to prevent self-ligation (Gary & David, 1996).

4.2.3.9 Ligating the inserts

The digested DNA was separated on agarose gel and bands in the range of 5 - 10 Kbp were then cut with a sterile razor blade and extracted from the gel using Qiaquik kit from Qiagen.

The λ-ZAP Express vector is already handled by the manufacturer with *Bam* HI restriction endonuclease which generates compatible ends to those generated by *Sau* 3AI. In addition, the ends are dephosphorylated to prevent the religation of the λ-arms.

The ligation proceeds with relatively equal molar ratio of the insert to prevent multiple inserts. ZAP Express vector can accommodate inserts ranging from 0 to 12 Kb.

Procedure:

The standard ligation reaction was set in 0.5 ml tube up to a total volume of 5 µl as indicated below:

1.0 µl of the ZAP Express vector (1 µg/µl)

1.7 µl DNA insert(s) (ca. 0.6 µg)

0.5 µl of 10x ligase buffer

0.5 µl of 10 mM rATP (pH 7.5)

0.7 µl T4 DNA ligase (ca. 2 U)

0.6 µl deionized sterile H_2O

The tube(s) was then incubated at 4 °C overnight.

4.2.3.10 In vitro Packaging of the recombinant λ-ZAP Express vectors

4.2.3.10.1 General considerations

In the natural lytic cycle of λ Bcteriophage; empty phage heads are produced within the bacterial cell and lengths of DNA (cut from a multiple-length DNA molecule) are packaged into the phage heads. The cuts are made at specific sequences known as *cos* sites. It is possible to do this *in vitro*, by adding cell extracts that contain phage heads, tails and the required enzymes to the DNA. This process, known as in vitro packaging, results in infectious virus particles. These particles are capable of injecting the DNA of interest into a sensitive host cell (Dale, 1994).

Procedure: (Gigapack III Gold Packaging Extract from stratagene was used**)**

4.2.3.10.2 Preparing the Host Bacteria (VCS257 strain)

A bacterial glycerol stock was streaked on LB agar plate and incubated overnight at 37 °C. A broth LB medium supplemented with 10 mM $MgSO_4$ and 0.2 % (w/v) maltose was inoculated with a single colony and allowed to grow with constant shaking at 37 ° C to an OD_{600} of 1.0. The cells were then spun at 5000 rpm for 10 min and the supernatant was discarded. Cells were then resuspended to an OD_{600} of 0.5 with sterile 10 mM $MgSO_4$. Following dilution the bacteria should be used immediately.

An amount of 200 - 300 ng of the ligated DNA was added to the immediately thawed (on ice) packaging extract and the contents were mixed by a pipette tip, followed by a quick spin to bring all the tube contents at the bottom. The mixture was then incubated at 22 °C for 2 h 500 µl of SM buffer was then added in addition to 20 µl chloroform with gentle mixing of the contents. The tube was then spun at 5000 rpm to sediment the debris and the supernatant containing the phage was further used in titering. The supernatant containing the phage kept at 4 °C until tittering had been achieved.

4.2.3.11 Plating for Blue-White color selection

4.1.3.11.1 General considerations

Early cloning vectors had relatively few sites for subcloning DNA fragments. In this case most of the available sites were within one of the two antibiotic resistance genes, with insertion causing loss of resistance. A polycloning site or 'polylinker' facilitates the cloning of DNA fragments by providing multiple restriction enzyme sites within a short segment of DNA. Many polylinker sites are constructed so that they lie within the coding sequence of the α-domain of ß-galactosidase (*lacZ*) (Gronenborn and Messing, 1978) under the control of the inducible *lac* promoter.

In intact vectors the α fragment is produced in the presence of suitable inducers and this complements a deletion within the *lacZ* gene of the host *E.coli* strain. This produces an active LacZ protein that can be detected using colorimetric (or fluorimetric) substrates, e.g., X-Gal, giving blue colonies. Insertion of DNA fragments into the polycloning site usually disrupts the gene preventing the production of LacZα.

Colonies with inserts are white allowing the identification of plasmids carrying inserts. In most bacteriophage vectors, a polylinker within a lacZα gene has been inserted in the intergenic region (between genes II and IV) and blue/white selection for DNA insertion can still be used. This intergenic region carries essential information for regulation of gene expression and viral DNA synthesis and as such some insertions will interfere with phage function.

M13 vectors reduce the growth rate of infected *E. coli* cells giving rise to blue/white holes or 'plaques' formed in a lawn of susceptible bacteria. Bacteriophage vectors are not used for routine maintenance of DNA clones as the essential nature of the intergenic region means that inserts are unstable and/or may be tolerated in only one orientation.

Due to the problems associated with M13 vectors, plasmids that carry an M13 intergenic region (containing the origin of replication) have been developed 'phagemids' (Vieira and Messing, 1987). The presence of an M13 origin allows the production of single-stranded phagemid DNA by infection with a helper phage that carries the genes for single-stranded replication and phage packaging.

Phagemids have the advantages of giving high yields of stable plasmid DNA and of single-stranded DNA from a single vector. As with bacteriophage vectors, only one strand of DNA is produced and packaged, and the choice of strands may influence the orientation of cloned products (Jones, 1998). The color assay is used for determining the ratio of recombinants to nonrecombinants within a newly constructed library.

Procedure:

In 1.5 ml microcentrifuge the following components were pipetted and mixed:

1 µl of the packaged reaction

200 µl of XL1-Blue MRF cells at an OD_{600} of 0.5

and

1 µl of a 1 : 10 dilution of packaged reaction

200 µl of XL1-Blue MRF cells at an OD_{600} of 0.5.

The phage and bacteria were incubated at 37 °C for 20 min to allow the phage to attach to the cells with gentle shaking.

To the above mentioned-mixture, 2-3 ml Of NZY top agar (48 °C), 15 µl of 0.5 M IPTG and 50 µl of X-gal (250 mg/ml) were added.

The cells were then plated onto NZY agar plates and allowed to set for 10 min, inverted and incubated overnight at 37 °C.

4.2.3.12 Phage titer

The determination of the phage concentration takes place by preparing a series of phage dilutions followed by pipetting 1 µl from each dilution for transfection of the host cells and plating. The plates that contain individual plaques can be then selected and the plaques counted and the

concentration of the phage's suspension could be estimated and expressed in pfu/µl units on the basis of the dilution factor.

Procedure: (The used medium did not contain antibiotic)

The cells were prepared as described in (preparing the host cells). A 1: 100 dilution was prepared in LB broth supplemented with 10 mM $MgSO_4$ and 0.2 % (w/v) maltose. The cells were allowed to grow to an OD_{600} of 1.0 at 37 °C with a gentle shaking followed by dilution to an OD_{600} of 0.5 with fresh LB broth.

A serial dilution was prepared in SM buffer. To 200 µl aliquots of the host cells (at an OD_{600} of 0.5), 1 µl of each diluted phage solution was added. The mixture was then incubated at 37 °C for 15 min. To each aliquot, 4.5 ml of 48 °C top agar was added and plated on 200-mm NZY agar plate which was further incubated at 37 °C overnight. The plaques number was then counted to determine the concentration of the library (pfu/ml) based on the dilution factor.

4.2.3.13 Amplifying the ZAP Express library

It is usually recommended to amplify the primary library in Lambda vectors in order to reach a high titer of the library.

The amplification for more than one round is not advisable since slow growing clones may be significantly underrepresented.

Procedure:

The host cells were prepared as outlined in preparing the host strains, and the cells were diluted to an OD_{600} of 0.5 in 10 M $MgSO_4$. The plaques were plated as outlined in section 4.2.3.14. The plates were then incubated for 6 - 8 h at 37 °C since the plaques diameter should not exceed 1 - 2 mm.

After incubation, each plate was overlaid with 10 - 13 ml SM buffer and stored at 4 °C overnight with gentle agitation to allow the phages to diffuse into the SM buffer. The recovered suspensions were transferred into a sterile Falcon tube and the plates were rinsed with an additional 2 - 3 ml of SM buffer per plate and chloroform to a 5 % (v/v) final concentration was added. The suspension was incubated for 15 min at room temperature to precipitate cell debris followed by centrifugation at 4 °C for 10 min at 5000 rpm.

The supernatant containing the phage particles was then transferred into a new Falcon tube and the suspension appearance was visually checked. Chloroform to a 0.3 % (v/v) final concentration was

then added and the titer of the newly developed library was determined with expecting a titer of 10^9 - 10^{11} pfu/ml.

4.2.3.14 Phage plating

Procedure:

1µl of the phage solution (~5 x 10^4 pfu) was added to 600 µl of the host strain (at an OD_{600} of 0.5) per plate. The mixture was then incubated with gentle shaking at 37 °C for 15 minutes to allow the phage to attach to the cells.

To this mixture 8.5 ml of NZY top agar (48 °C) was added and plated immediately onto a warmed NZY agar plate. The plates allowed to set for 10 minutes, inverted and incubated at 37 °C for about 8 hours. The plates were then chilled for 2 h at 4 °C to prevent the NZY top agar from sticking to the nitrocellulose membrane.

4.2.3.15 Lifting the plaques

Procedure:

A nitrocellulose membrane was placed onto each NZY agar plate for 2 min to allow the transfer of the phage particles to the membrane. A waterproof ink in a syringe needle was used to prick through the membrane and agar for orientation. A second membrane was also allowed to transfer but for longer time (about 5 min).

After lifting, the membranes were denatured by spreading a 5 - 10 ml of 1.5 M NaCl and 0.5 M NaOH denaturation buffer on the membranes using a syringe for 2 – 3 min. Neutralization was performed by spreading 5 - 10 ml of a neutralization buffer (1.5 M NaCl and 0.5 M Tris-HCl pH 8.0) on each membrane using a syringe for 5 min.

A rinse buffer composed of 0.2 M Tris-HCl, pH 7.5 and 2x SSC solution buffer was applied for 30 sec by the same way as indicated in the preceding steps to rinse the membranes. The membranes were then blotted on a Whatman 3 MM papers and finally the DNA was fixed to the membranes by baking them at 80 °C for 1.5 h. The agar plates were stored at 4 °C for subsequent use.

4.2.3.16 In vivo excision of the pBK-CMV

4.2.3.16.1 General considerations

Excision is the release of a phage, insertion element, episome or any other element or DNA sequence from a nucleic acid chain (Redei, 1998). In vivo excision was performed on the selected isolates to extract the insert-containing pBK-CMV phagemid vector. The used ZAP Express vector

is designed (as mentioned by the manufacturer) to allow simple, efficient in vivo excision and recircularization of the cloned insert(s) contained within the lambda vector to form a phagemid containing the cloned insert (Terawaki *et al*., 1967 & Ishihara *et al*., 1978). This in vivo excision depends on the placement of the DNA sequences within the lambda phage genome and on the presence of a variety of proteins, including filamentous (e.g., M13) bacteriophage-derived proteins. These proteins recognize a region of DNA normally serving as the f1 bacteriophage "origin of replication". This origin of replication can be divided into two overlying parts: (1) the site of initiation and (2) the site of termination for DNA synthesis (Coetzee *et al*., 1972). These two regions are subcloned separately into the ZAP Express vector. The lambda phage (target) is made accessible to the M13-derived proteins by simultaneously infecting a strain of *E. coli* with both the lambda vector and the M13 helper phage.

Inside *E. coli*, the "helper" proteins (i.e., proteins from M13 phage) recognize the initiator DNA that is within the lambda vector. One of these proteins then nicks one of the two DNA strands. At the site of this nick, new DNA synthesis begins and duplicates whatever DNA exists in the lambda vector "downstream" (3´) of the nicking site. DNA synthesis of a new single strand of DNA continues through the cloned insert until a termination signal, positioned 3´ of the initiator signal, is encountered within the constructed lambda vector.

The ssDNA molecule is circularized by the gene II product from the M13 phage, forming a circular DNA molecule containing the DNA between the initiator and terminator. In the case of ZAP express vector, this includes all sequences of the pBK-CMV phagemid vector and the insert. This conversion is the "subcloning" step since all sequences associated with normal lambda vectors are positioned outside of the initiator and terminator signals and are not contained within the circularized DNA. In addition, the circularization of the DNA automatically recreates a functional f1 origin as found in f1 bacteriophage or phagemids.

Signals for "packaging" the newly created phagemid are linked to f1 origin sequence. The signals permit the circularized ssDNA to be packaged into phagemid particles and secreted from the *E. coli*. Following secretion of the phagemid particle, the *E. coli* cells used for in vivo excision of the cloned DNA are killed and the lambda phage is lysed by heat treatment at 70 °C. The phagemid is not affected by the heat treatment. *E. coli* is infected with the phagemid and can be plated on a selective medium to form colonies. DNA from excised colonies can be used for analysis of insert DNA, including DNA sequencing, subcloning and mapping.

The ExAssist helper phage (supplied by the manufacturer) contains an amber mutation that prevents replication of the helper phage genome in a nonsuppressing *E. coli* strain such as XLOLR cells. This allows only the excised phagemid to replicate in the host, removing the possibility of co-infection from the ExAssist helper phage.

Procedure for single clone excision:

DNA of some selected plaques that showed positive signals were excised for subsequent analysis as follows:

The plaque of interest was cored from the agar plate and transferred to a sterile microcentrifuge tube containing 0.5 ml of SM buffer and 20 µl of chloroform. The tube was then vortexed to release the phage particles into the SM buffer and incubated at 4 °C overnight.

A culture of XL1-Blue MRF´ cells was grown overnight in LB-medium supplemented with 0.2 % (w/v) maltose and 10 mM $MgSO_4$ at 30 °C with shaking. Cells were then spun down for 5 min at 5000 rpm, 4 °C. The pellet was resuspended in 10 mM $MgSO_4$ at an OD_{600} of 1.0.

In a Falcon polypropylene tube, the following components were combined:

200 µl of XL1-Blue MRF´ cells at an OD_{600} of 1.0

250 µl of phage stock solution (see above)

1 µl of the ExAssist helper phage

Note: For efficient excision the recommended concentrations of phages and *E. coli* cells were considered as stated in the manufacturer's protocol.

The tube was then incubated at 37 °C for 15 min followed by adding 3 ml of NZY broth and further incubation for 3 - 12 h at 37 °C with gentle shaking. For killing the *E. coli* cells the tube was heated at 70 °C for 15 - 20 min and the cell debris was pelleted by centrifugation at 5000 rpm, 4 °C for 15 min. The supernatant containing the excised pBK-CMV phagemid vector packaged as filamentous phage particles was carefully transferred to a sterile Falcon tube and stored at 4 °C.

4.2.3.17 Plating the excised phagemids

Procedure:

A fresh culture of XLOLR strain in NZY broth was grown overnight at 30 °C. Cells were then pelleted by centrifugation at 5000 rpm, 4 °C for 5 min. The pellet was then resuspended in 10 mM $MgSO_4$ at an OD_{600} of 1.0. 10 and 100 µl of the prepared phage stock solution were added to 2 separate sterile Falcon tubes each contains 200 µl of the freshly prepared XLOLR cells. The tubes were then incubated at 37 °C for 15 min. To each tube 300 µl of NZY broth medium was added and the tubes were further incubated for 45 min. 200 µl of the cell mixture from each Falcon tube was plated on LB-kanamycin (50 µg/ml) agar plate and the plates were incubated at 37 °C

overnight. The colonies that have been grown on the agar plates were supposed to contain the pBK-CMV double stranded vector with the cloned DNA insert.

4.2.3.18 Screening of the genomic DNA library

Rapid screening methods have been developed to enable very large numbers of colonies to be tested simultaneously. The testing commonly involves using of either a gene probe to detect the specific DNA or an antibody to detect the protein product. Screening the library using the first afore-mentioned method was highly recommended. So different types of gene probes were prepared making use of the available information on the N-terminal of HVOR from *P. vulgaris* which was determined by Trautwein and Simon, 1994 and the related genes that have been already characterized from other sources.

Slightly different gene probes (primers) were derived from the N-terminal sequence of HVOR (20 amino acids) for fishing the target gene or a part of it using polymerase chain reaction technique. Another gene probes based on the knowledge of some related sequences were synthesized to be used as heterologous probes using a low stringency of hybridization or for screening the library with PCR. The 3rd type of probes was the products of the PCR that underwent labeling and was used subsequently as homologous probes for screening the genomic library at high stringency. Table 2 shows the different types of probes that had been used for screening the genomic library of *P. vulgaris*.

4.2.3.19 Screening the genomic library by PCR

4.2.3.19.1 General considerations

PCR is an *in vitro* method for enzymatically synthesizing defined sequences of DNA. The reaction normally uses two oligonucleotide primers that hybridize to opposite strands and flank the target DNA sequence that is to be amplified. The elongation of the primers is catalyzed by a heat-stable DNA polymerase. A repetitive series of cycles involving template denaturation, primer annealing and extension of the annealed primers by the polymerase results in exponential accumulation of a specific DNA fragment. The ends of the fragments are defined by the 5' ends of the primers. Because the primer extension products synthesized in a given cycle can serve as a template in the next cycle, the number of target DNA copies approximately doubles every cycle. Thus, 20 cycles of PCR yield about a million copies (2^{20}) of the target DNA.

A number of factors influence the fidelity of the PCR reaction for example the presence of contaminants and the concentration of the $MgCl_2$. The contaminants can be amplified instead of the target template. $MgCl_2$ forms soluble complexes with dNTPs to produce the actual substrate

that the polymerase recognizes. The concentration of free Mg^{2+} depends on the concentration of the compounds that bind the ion, including dNTP, free pyrophosphate (pp_i) and EDTA.

The optimal $MgCl_2$ concentration varies from approximately 1 mM to 5 mM. The most commonly used $MgCl_2$ concentration is 1.5 mM (with dNTPs at a concentration of 200 µM each). Mg^{2+} influences enzyme activity and increases the T_m of double-stranded DNA. Excess Mg^{2+} in the reaction can increase non-specific primer binding and increase the non-specific background of the reaction.

A balanced solution of all four deoxynucleoside triphosphates (dNTPs) must be used to minimize polymerase error rate. Imbalanced dNTP mixtures will reduce Taq DNA polymerase fidelity. Increases in dNTP concentration reduce free Mg^{2+}, thus interfering with polymerase activity and decreasing primer annealing.

In most PCR applications, it is the sequence and the concentration of the primers that determine the overall success. Primer concentrations between 0.1 and 0.6 µM are generally optimal. Higher primer concentrations may promote mispriming and accumulation of non-specific products. Lower primer concentrations may be exhausted before the reaction is completed, resulting in lower yields of the desired product. A very important step in PCR is the complete denaturation of the template DNA. If the template DNA is only partially denatured, it will tend to 'snap-back' very quickly, preventing efficient primer annealing and extension, or leading to 'self-priming' which can lead to false-positive results. The choice of the primer annealing temperature is probably the most critical factor in designing a high specificity PCR. If the temperature is too low, non-specific annealing will dramatically increase.

For fragments up to 3 kb, primer extension is normally carried out at 72 °C. *Taq* DNA polymerase can add approximately 60 bases per second at 72 °C. A 45-second extension is sufficient for fragments up to 1 kb. For extension of fragments up to 3 kb, allow about 45 seconds per kb. Most PCRs should include only 25 to 35 cycles. Usually, after the last cycle, the reaction tubes are held at 72 °C for 5-15 min to promote completion of partial extension products and annealing of single-stranded complementary products (Roche, 1999).

4.2.3.19.2 Primer degeneracy

Frequently, the limiting step in detecting and/or cloning a gene is the generation of a complementary strand of nucleic acid to be used as a probe. The origin of a probe depends on what is known about the gene under investigation. Sometimes a gene cloned from another organism can be used as a probe for carrying out hybridization at low stringency conditions. Alternatively, if the protein product of a gene has been purified, probe can be designed and synthesized on the basis of its amino acid sequence and knowledge of the genetic code.

On the other hand, the necessary DNA sequence information can be obtained from sequence databases that detail the structure of millions of genes from a wide range of organisms (Nelson & Cox, 2000).

There are two types of oligonucleotide probes which can be designed from a protein sequence. One is a set of oligonucleotides which are relatively short and contain all possible nucleotide sequences that could code for a stretch of amino acids. Since the genetic code is degenerated, that is, more than one codon codes for a particular amino acid, the length of these probes is limited by the number of different sequences one desires to be included in the probe pool.

The other type of oligonucleotide that may be designed is a relatively long, low degenerated oligonucleotide which reduces the degeneracy by making guesses as to which codon to be used in order to code for a particular amino acid.

The use of a mixture of different probes for the same protein under investigation reduces the number of false signals one obtains when screening a library. A number of unmodified primers were synthesized on the bases of the available information on the N-terminal of the 80 kDa unit of the HVOR and also the N-terminal of the associated 60 kDa protein.

Due to the unavailability of other amino acid sequences from within the protein and the high degeneracy of the N-terminus which makes it inconvenient for synthesizing the right primers of low degeneracy, it would be necessary to make use of the offer provided by some oligonucleotide's manufactures for synthesizing primers revealing all the alternatives of codons degeneracy.The design of the appropriate probes (primers and oligos) was assisted by using some primer design software programs such as Primer Designer program for windows Version 3.0 from the Scientific & Educational Software 1994 - 1996 and Primer Sequence program in HUSAR Package from Heidelberg, Germany. These programs were used to ensure that the primer sequences have the following general characteristics:

1.) Are 18 - 24 bases long
2.) Contain no internal secondary structure
3.) Contain 40 – 60 % G/C
4.) Have a balanced distribution of G/C and A/T rich domains
5.) Are not complementary to each other at the 3´ ends (so primer-dimers will not form)
6.) Have a melting temperature (T_m) that allows annealing temperatures of 55 – 65 °C

Two degenerated primers based on the amino acid sequence of the N-terminus of the 80 kDa monomer of the HVOR were designed. Another two degenerated primers based on the amino acid sequence of the N-terminus of the 60 kDa associated protein were also designed as follows:

- Degenerated primers based on the HVOR N-terminus

AA seq.	M	I	N	G	W	T	G	N	I	L	R	I	N
	Met	Ile	Asn	Gly	Trp	Thr	Gly	Asn	Ile	Leu	Arg	Ile	Asn
	ATG	ATA	AAT	GGG	TGG	ACG	GGG	AAT	ATA	TTG	AGG	ATA	AAT
		ATT	AAC	GGA		ACA	GGA	AAC	ATT	TTA	AGA	ATT	AAC
		ATC		GGT		ACT	GGT		ATC	CTG	CGG	ATC	
				GGC		ACC	GGC			CTA	CGA		
										CTT	CGT		
										CTC	CGC		
D. seq.	***ATG***	***ATH***	***AAY***	***GGN***	***TGG***	***ACN***	***GGN***	***AAY***	***ATH***	***YTN***	***MGN***	***ATH***	***AAY***
M.80.$^{+}$			***AAY***	***GGN***	***TGG***	***ACN***	***GGN***	***AAY***	***AT***				
M.80.$^{-}$			***ATR***	***TTI***	***CCI***	***GTC***	***CAI***	***CCR***	***TT***				

AA seq.	L	T	T	G	A	I	S
	Leu	Thr	Thr	Gly	Ala	Ile	Ser
	TTG	ACG	ACG	GGG	GCG	ATA	TCG
	TTA	ACA	ACA	GGA	GCA	ATT	TCA
	CTG	ACT	ACT	GGT	GCT	ATC	TCT
	CTA	ACC	ACC	GGC	GCC		TCC
	CTT						AGT
	CTC						AGC
D. seq.	***YTN***	***ACN***	***ACN***	***GGN***	***GCN***	***ATH***	***WSN***

- Degenerated primers based on the N-terminus of the assoiated protein

AA seq.	M	K	K	S	L	V	A	V	G	V	I	V	A
	Met	Lys	Lys	Ser	Leu	Val	Ala	Val	Gly	Val	Ile	Val	Ala
	ATG	AAG	AAG	TCG	TTG	GTG	GCG	GTG	GGG	GTG	ATA	GTG	GCG
		AAA	AAA	TCA	TTA	GTA	GCA	GTA	GGA	GTA	ATT	GTA	GCA
				TCT	CTG	GTT	GCT	GTT	GGT	GTT	ATC	GTT	GCT
				TCC	CTA	GTC	GCC	GTC	GGC	GTC		GTC	GCC
				AGT	CTT								
				AGC	CTC								
D. Seq.	***ATG***	***AAR***	***AAR***	***WSN***	***YTN***	***GTN***	***GCN***	***GTN***	***GGN***	***GTN***	***ATH***	***GTN***	***GCN***
M.60.$^{+}$							***GCN***	***GTN***	***GGN***	***GTN***	***ATH***	***GTN***	
M.60.$^{-}$							***NAC***	***HAT***	***NAC***	***NCC***	***NAC***	***NGC***	

AA seq.	L	G	V	V	W	T	G	A
	Leu	Gly	Val	Val	Trp	Thr	Gly	Ala
	TTG	GGG	GTG	GTG	TGG	ACG	GGG	GCG
	TTA	GGA	GTA	GTA		ACA	GGA	GCA
	CTG	GGT	GTT	GTT		ACT	GGT	GCT
	CTA	GGC	GTC	GTC		ACC	GGC	GCC
	CTT							
	CTC							
D. seq.	***YTN***	***GGN***	***GTN***	***GTN***	***TGG***	***ACN***	***GGN***	***GCN***

4.2.3.19.3 General PCR procedure

For fishing the target gene (HVOR) or even part of it, a freshly prepared genomic library as well as the genomic DNA (intact or sheared) was used as templates for PCR experiments.

Before beginning, all the reagents were to be centrifuged. On ice, the following mixture was always to be set up:

Polymerase buffer, 10x conc.	10.0 µl
10 mM Nucleotide Mix	4.0 µl (final conc.: 200 µM, each dNTP)
Upstream primer	1.0 µl (0.2 – 1.2 µM)
Downstream primer	1.0 µl (0.2 – 1.2 µM)
$MgCl_2$	1.0 µl (1 – 1.5 mM)
Template DNA	10.0 µl (up to 400 µg/reaction)
Taq/*Pwo* Polymerase	1.0 µl (up to 5 units/reaction)
St. deion. water	72.0 µl
Total Vol.	100 µl

Slight modifications concerning the concentrations of $MgCl_2$, DNA template, and primers were to be considered according to the conditions and requirements of each reaction.

4.2.3.20 Cloning of the PCR products

The PCR products (fished fragments) were cloned for subsequent analysis using PCR cloning vector kit from Stratagene (PRC-Script™ Amp Electroporation-Competent Cell Cloning Kit).

4.2.3.20.1 Purifying the PCR products with the StrataPrep PCR purification kit

PCR created band(s) was/were run on 0.8 – 1 % agarose gel for verifying the integrity and quality. After the extraction of the DNA from agarose gel (see section 4.2.3.7) the sample was to be purified before proceeding with the cloning protocol.

Procedure:

An equal volume of the DNA-binding solution to the volume of the aqueous portion of the PCR product was added in a microcentrifuge tube and the components were mixed. The tube content was then transferred to a microspin cup that is seated in 2 ml receptacle tube. The microspin cup has a fiber matrix that has a binding capacity of ~10 µg of DNA. The microcentrifuge tube was then spun down at 4 °C, 13000 rpm for 0.5 – 1 min. The flow through was then discarded and the bounded DNA was washed with 750 µl of PCR wash buffer, centrifuged again as described in the preceding step and the flow through was discarded.

The rest of the wash buffer was removed by an additional centrifugation step. The microspin cup was left for 5 min at room temperature and the DNA was then eluted in adequate volume of 10 mM Tris-HCl buffer, pH 8.5.

4.2.3.20.2 Polishing the purified PCR products

The ends of PCR products generated with *Taq* DNA polymerase or other low-fidelity DNA polymerases should be polished in order to create blunt ends. This enables the polished inserts to be easily cloned into pPCR-Script Amp SK(+) cloning vector.

Procedure:

In a 0.5 ml microcentrifuge tube, a polishing reaction was set up in the order indicated below:

10.0 µl of the purified PCR product (~ 150 - 200 ng/µl)
1.0 µl of 10 mM dNTP mix (2.5 mM each)
1.3 µl of 10x polishing buffer
1.0 µl of cloned *pfu* DNA polymerase (0.5 U)

The components were then mixed and overlaid with 20 µl mineral oil to prevent the evaporation of the components of the mixture. The polishing reaction was allowed to proceed for 30 min at 72 °C in a water bath. The tube has been stored on ice till it was used in ligation reaction.

4.2.3.20.3 Ligating the insert

Procedure:

The following ligation reaction was set up in the order recommended by the manufacturer:

1 µl of the pPCR-Script Amp SK (+) cloning vector (10 ng/µl)
1 µl of PCR-Script 10x reaction buffer
0.5 µl of 10 mM rATP
2-4 µl of the blunt-ended PCR product (~ 30 – 50 ng/µl)
1 µl of *Srf* I restriction enzyme (5 U/µl)
1 µl of T4 DNA ligase (4 U/µl)

The ligation reaction was then mixed by the tip of a pipette and incubated at room temperature for 1 h. The reaction was then stopped by heating the tube for 10 min at 65 °C. The tube was kept on ice until the transformation reaction was ready.

4.2.3.20.4 Preparation of competent cells

4.2.3.20.4.1 General considerations

Competence is a physiological state of the bacterial cell when transformation is successful. It generally coincides with the second half of the generation time or its peak is near the end of the exponential growth phase (Redei, 1998). In many organisms (including *E. coli*) competent cells do not appear to occur naturally. This physiological state can be induced artificially by treating the bacterial cells with divalent cations like cold $CaCl_2$.

Procedure:

Preparation of competent cells was carried out using calcium chloride protocol according to Cohen *et al*., 1972 as follows:

Cells of XL1-Blue MRF´ *E. coli* strain were grown overnight in LB broth at 37 °C with a constant agitation. 100 ml of fresh LB broth was then inoculated with 1 ml of the overnight grown culture. The cells were allowed to grow at 37 °C to OD_{600} of 0.3. The 100 ml culture was allowed to chill on ice for 10 min. The chilled cells were then spun down at 5000 rpm, 4 °C for 10 min. The supernatant was discarded and the pellet resuspended in 40 ml of cold 0.1M $CaCl_2$. The cell suspension was then left on ice for 30 min followed by pelleting the cells by centrifugation for 10 min at 4 °C, 5000 rpm. The pellet was then resuspended in 4 ml of cold 0.1 M $CaCl_2$. The cells were then divided into equal aliquots each of 40 µl and stored at -80 °C.

4.2.3.20.5 Transformation

4.2.3.20.5.1 General considerations

Transformation was first discovered in 1928 in pathogenic strains of *Streptococcus pneumoniae* by Frederick Griffith, although he knew nothing of the actual nature of the process. The movement of donor DNA molecules across the cell membrane and into the cytoplasm of recipient bacteria is an active, energy-demanding process. It does not involve the passive diffusion of DNA molecules through permeable cell walls and membranes. Transformation is not a naturally occurring process in all species of bacteria; rather, it takes place only in those species that possess the protein and enzymatic machinery required to bind free DNA molecules in the medium and transport them to the cytoplasm. Only competent cells which secrete a competence factor (a small protein that induces the synthesis of 8 to 10 new proteins required for transformation) are capable of serving as recipients in transformation. The proportion of bacteria in a culture that are physiologically competent to be transformed depends on the growth conditions. In most bacterial species, cells that are likely to be transformed are dividing at their maximal rate. These populations of cells are growing exponentially and are fast approaching the plateau phase where nutrients in the medium become a limiting factor in the continued growth of the population (Terawaki and Rownd, 1972).

Procedure:

The electroporation chamber, cuvettes and the 1.5 ml microcentrifuge tubes were chilled on ice before proceeding with the transformation. The Epicurian Coli XL1-Blue MRF' Kan electroporation-competent cells were thawed on ice (~ 5 min). Then 40 µl of the competent cells were added to 2 µl of the experimental ligation reaction from the preceding step in a 1.5 ml microcentrifuge tube and the contents were swirled gently. The content of the tube was then transferred to a chilled electroporation cuvette with tapping the top of the cuvette until the mixture settled evenly down. The transformation was carried out at 25 µF and 1700 V (field strength of 17 kV/cm across the 0.1-cm gap in the cuvette) by sliding the cuvette into the chilled electroporation chamber until being connected with the electrical contacts. After pulsing once, 960 µl of sterile SOC medium was immediately added and the cells were resuspended. The cells were incubated at 37 °C for 1 h with constant shaking. The cells were pelleted by centrifugation at 4 °C, 5000 rpm for 5 min and then resuspended in 200 µl NZY broth medium. The cells were plated onto LB-ampicillin (50 µg/ml) agar plate. For color selection, the agar plates were prepared with 2 % X-gal and 10 mM IPTG. The plates were then inverted and incubated up to 17 h at 37 °C to allow the development of the blue color. The cells containing plasmids with inserts were white in color while the other cells that contain no insert were blue. The plates were left for 2 h at 4 °C in order to enhance the development of the blue color.

4.2.3.21 High Copy-Number Plasmid isolation Protocol

4.2.3.21.1 General considerations

Plasmids from *E. coli* consist of two types: The first group (usually referred to as high copy number plasmids), of which ColE1 is the prototype. They are relatively small (usually less than 10 kb) and unable to promote their own transfer by conjugation. They are present in multiple copies within the cell (10 - 200). Their replication is not linked to the processes of chromosomal replication and cell division. Their copy number can be increased to several thousands per cell if host protein synthesis is stopped (e.g., by treatment with chloramphenicol) (Clewell, 1972).

The second group of plasmids, exemplified by the F plasmid (usually referred to as low copy number plasmids), are larger (typically containing more than 30 kb) and able to promote their own transfer by conjugation. They are present in only one or two copies per cell. These plasmids are under stringent control (Novick *et al.*, 1976).

With the aid of NucleoTrap® Nucleic Acid purification kit from CLONTECH, the plasmid DNA containing inserts of the target clones from agarose gel was isolated. The kit is designed around a specially activated matrix (suspension) that binds nucleic acids and separates it from contaminants. This matrix binds nucleic acid in the presence of chaotropic salts which disturb the hydrate shell

surrounding the nucleic acid. During the purification procedure, the matrix is washed to remove impurities. Nucleic acid is eluted in a small volume of low-salt buffer e.g., 1 mM EDTA, 10 mM Tris-HCl, pH 8.5 in dist. H_2O.

Another Kit from Qiagen (QIAprep Miniprep®) was also used for the same purpose. The procedure is based on alkaline lyses of bacterial cells followed by adsorption of DNA onto silica in the presence of high salt using QIAprep columns. The plasmid DNA is then washed to remove the endonucleases and the purified DNA can be eluted in the above mentioned elution buffer or in deionized water.

A third kit from PeQLab with the same principle of the QIAprep Miniprep kit from Qiagen, called E.Z.N.A.® Plasmid Miniprep Kit I was also in the majority of plasmid DNA isolations.

Procedure:

(***Note: all the centrifugations were carried out at 13000 rpm, 4 °C for 0.5 - 1 min unless else was recommended)***

5 ml LB medium supplemented with antibiotic at the recommended concentration was inoculated with only one colony of the *E. coli* competent cells that contain the plasmid under investigation. The culture was then incubated up to 16 h at 37 °C with gentle shaking. The cells were then pelleted by centrifugation at 5000 rpm for 5 min and the supernatant was carefully discarded. The lysis of the bacteria was carried out by resuspending the pellet in 250 µl of buffer 1 (supplemented with RNase) followed by a brief vortexing to ensure complete homogeneity of the cell suspension. 250 µl of buffer 2 was then added and the mixture was mixed by inverting the tube 4 - 6 times. Under certain circumstances, the mixture was incubated at room temperature for 2 - 5 min untill a homogenous solution was developed to ensure complete lysis of the cells.

The neutralization of the lysate was by adding 350 µl of buffer 3 followed by inverting the tube till a white precipitate formed. The cell debris was then collected by centrifugation for 10 min at 13000 rpm, 4 °C. The lysate was then loaded onto a separation column and centrifuged. The flow through was then discarded. To ensure the complete absence of protein contamination, the plasmid DNA bounded to the *HiBind-Silikamatrix* in the separation column was washed with 500 µl of HB-buffer followed by centrifugation. The flow-through was then discarded and the plasmid DNA was washed twice with 2 x 750 µl of DNA-Washing buffer (supplemented with absolute ethanol). After centrifugation, the flow-through was discarded and the rest of DNA-Washing buffer was removed by an additional centrifugation step at 13000 rpm, 4 °C for 1 min and up to several minutes untill complete drying was ensured. Plasmid DNA was eluted in 50 - 100 µl of deionized water or in TE-buffer (10 mM Tris-HCl pH, 8.5) followed by centrifugation.

4.2.3.22 Ethanol precipitation of plasmid DNA

For plasmid sequencing, precipitation at room temperature was recommended. To the plasmid containing the DNA insert of interest, the following was added:
2.5 vol. of 95 % ethanol and 1 vol. of 3 M sodium acetate, pH 4.8 (or 1 vol. of 5 M ammonium acetate, pH 7.4). The tube contents were then mixed thoroughly followed by incubation at room temperature for 15 - 20 min. Following centrifugation at 13000 rpm, 4 °C for 10 min, the resultant pellet was washed with 500 - 700 µl of 70 % ethanol and pelleted again by centrifugation at 13000, 4 °C for 5 - 10 min. The pellet was then left to dry at room temperature.

4.2.3.23 Sequencing

The sequencing of the single- and double-stranded DNA fragments was carried out with the ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit from Perkin Elmer Applied Biosystems. This kit contains the sequencing enzyme AmpliTaq DNA polymerase, FS. This enzyme is a variant of the *Thermus aquaticus* DNA polymerase that contains a point mutation in the active site. This results in less discrimination against dideoxynucleotides. This enzyme also has a second mutation in the amino terminal domain that virtually eliminates the 5' → 3' nuclease activity of AmpliTaq DNA polymerase. This kit allows the performing of fluorescence-based cycle sequencing reactions on single-stranded or double-stranded DNA fragments.

In a microcentrifuge tube the following components were mixed on ice:

Plasmid-DNA (up to 400 ng DNA):	5.0 µl
Ready Reaction Mix and	: 4.0 µl
Primer (1 – 2 pmol)	: 1.0 µl
Total Vol.	10 µl

The Ready Reaction Mix is composed of the dye terminators, deoxynucleoside triphosphates, AmpliTaq DNA Polymerase (FS), r*Tth* pyrophosphatase (a component in AmpliTaq DNA Polymerase, FS), magnesium chloride, and the buffer. These components were premixed into a single tube and were ready to be used. The reaction mixture was then subjected to the following PCR program:

Temperature	Time
96 °C	30 sec
50 °C	15 sec
60 °C	4 min

The program was allowed to run for 25 cycles and the DNA was then analyzed with the help of a fluorescence detector by GATC, Konstanz, Germany.

4.2.3.24 Sequence analysis

All DNA sequence analysis was performed with the HUSAR-Programm package (Heidelberg Unix Sequence Analysis Resources) at the Deutschen Krebsforschungszentrums Heidelberg.

4.2.3.25 Screening the genomic library by plaque hybridization

Hybridization can be defined as the crossing of genetically different individuals; annealing DNA single strands with RNA or a single-stranded DNA of different origin (Redei, 1998). DNA hybridization is the most common sequence-based process for detecting a particular gene or segment of nucleic acid. There are many variations of the basic method, most of which making use of a labelled DNA or RNA fragment, a probe, complementary to the DNA being sought. (Nelson and Cox, 2000)

4.2.3.26 Digoxigenin labelling of templates by "Random Prime Labelling"

DIG Labelling

▼

Evaluation of Probe Labelling Efficiency

▼

Electrophoresis of Target Nucleic Acids on Agarose Gels

▼

Blot Transfer of Target Nucleic Acids to a membrane

▼

Hybridization of Probe to Target

▼

Stringent Washes of Blot

▼

Detection of Probe-Target Hybrid

Flow Chart: The different steps of Hybridization (in order)

4.2.3.26.1 General Considerations

Dig-labelling is a simple adaptation of the enzymatic labelling procedure. DIG-labelled probes and ^{32}P-labeled probes behave with similar kinetics. They may be used under similar hybridization conditions. The recommended concentration of random primed labelled DNA is around 25 ng/ml. Random primed labelling can label templates of almost any length, while for very short sequences, PCR labeling method is being highly recommended. In random primed labelling, Klenow enzyme copies the DNA template in the presence of hexameric primers and alkali-labile DIG-dUTP. On average, the enzyme inserts one DIG moiety in every stretch of 20 - 25 nucleotides.

The resulting labelled products are homogeneously labelled, very sensitive hybridization probes, able to detect as little as 0.10 – 0.03 pg target DNA. These labelled probes are especially suitable for single copy gene detection on genomic Southern blots and in screens of recombinant libraries.

4.2.3.26.2 Random Primed Labelling of DNA Probes (High Yield Method)

Different types of purified templates such as cloned inserts, isolated free of vector sequences, PCR generated templates and synthesized oligos of various lengths and amounts (ranging from 10 ng to 3 µg for single-copy gene detection on a blot) were labelled as follows:

10 ng – 3 µg template DNA was added (linear or supercoiled) to a reaction tube and autoclaved. St. bidest. water was then added to a final volume of 16 µl (up to scale). The sample was heated in a boiling water bath for 10 min and quickly chilled in an ice/ethanol water bath. The DIG-High Prime was then mixed thoroughly, and 4 µl of it was added to the denatured sample. The reaction mixture was then centrifuged briefly and incubated for at least 1 h and up to 20 h at 37 °C. The reaction was stopped by heating the sample to 65 °C for 10 min.

4.2.3.26.3 Evaluation of probe labelling efficiency

It is important to check the efficiency of each labelling reaction by determining the amount of DIG-labelled product because too much probe will lead to serious background problems and too little probe will result in little or no hybridization signal. The preferred way to roughly quantify almost all labelled nucleic acid probes (except PCR-labelled probes) was the “direct detection” method. In this method, a series of dilutions prepared from the DIG-labelled probe was spotted directly on a membrane and visualized with standard DIG detection procedures.

Another method was also applied in which case a series of dilutions of DIG-labelled DNA were applied to the marked squares of the DIG quantification test strips (included in the purchased kit “Dig high prime labelling and detection starter kit I” from Boehringer Mannheim). The test strips were already loaded with defined dilutions of a control DNA which were used as a standard. The test strips were then subjected to immunological detection with Anti-digoxigenin-AP conjugate

and the premixed stock solution of the alkaline phospatase (NBT/BCIP).The results could be recognized after approximately 30 min – 1h.

4.2.3.26.4 Electrophoresis of target nucleic acids on agarose gels

For best results, the target nucleic acids were electrophoretically separated on agarose gel and blotted under optimal conditions e.g., the concentration of the agarose gel, the length of the run and the strength of the applied electrical field. The amount of target nucleic acids (genomic DNA) loaded on an electrophoretic gel was in the range of 1.0 - 5.0 µg.

Ethidium bromide was not included in the gel, because it can cause uneven background if the gel is not run long enough. After electrophoresis, the gel was stained with 1µg/ml ethidium bromide solution and destained with water to ensure that the target nucleic acid is intact.

4.2.3.27 Screening of the *P. vulgaris* genome by blot hybridization

4.2.3.27.1 Blot transfer of target nucleic acids to a solid membrane

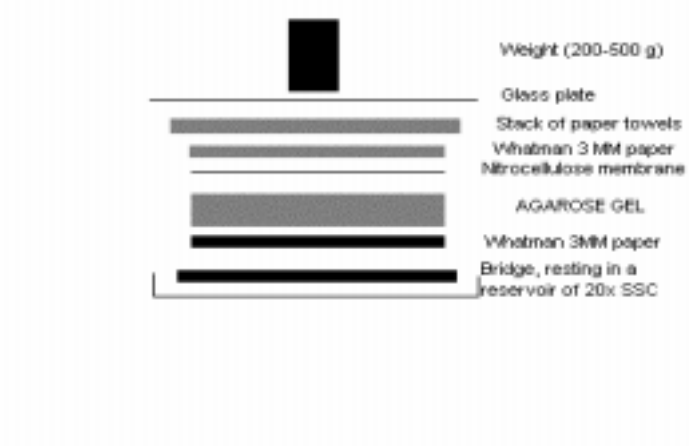
4.2.3.27.1.1 General considerations

Generally there are three methods to transfer fragments of DNA from agarose gels to solid supports (nitrocellulose filters or nylon membranes): Capillary transfer, electrophoretic transfer as well as vacuum transfer. During the course of this work, the first method was chosen because of its convenience. In capillary transfer method (Southern, 1975), DNA fragments are carried from the gel in a flow of liquid and deposited on the surface of the solid support. The liquid is drawn through the gel by capillary action that is established and maintained by a stack of dry, absorbent paper towels as shown below.

Procedure:

Following separation of a DNA sample on an agarose gel the unused areas of the gel were trimmed away with a razor blade and the bottom left-hand corner of the gel was also cut off. This helps to orient the gel during the succeeding operations. The DNA in gel was denatured by submerging the gel in several volumes of denaturation solution (0.5 N NaOH, 1.5 N NaCl) for 45 min at room temperature with gentle shaking. The gel was then rinsed briefly in sterile, deionized water for removing the excess of the denaturation buffer. The neutralization of the DNA in gel was performed by soaking the gel in several volumes of neutralization buffer (1 M Tri-HCl, pH 7.4; 1.5 M NaCl) for 30 min at room temperature with constant, gentle agitation, followed by another wash for 15 min with fresh buffer.

Equilibration of the gel was done by submerging it in 20x SSC. The blot transfer was blotted in a shallow reservoir containing transfer buffer (20x SSC) as shown below, with care taking to avoid formation of air bubbles. The gel placed upside down on a piece of Whatman 3 MM paper that had been previously soaked in 20x SSC and dangles in the transfer buffer. The air bubbles that could be formed between the gel and paper were removed by gentle rolling of a sterile pipette over the



sandwich. A piece of nitrocellulose membrane was cut exactly to the same size of the gel.

A matched corner to that of the gel was also cut. The membrane was then soaked in deionized water for rehydration until it became completely wet from beneath. It was then immersed in transfer buffer for about 7 min. The gel was then inverted so that its underside is now uppermost.

The air bubbles were also removed as described before. The inverted gel was then placed on the support so that it is centered on the wet 3 MM papers. Two pieces of Whatman 3 MM paper with the same size of the gel were wetted in 2x SSC and laid on the gel. The blot assembly was then completed by surrounding the gel with Parafilm to prevent liquid from flowing directly from the reservoir to the paper towels placed on top of the gel. This was followed by placing a glass plate and a weight of approximately 500 g on the gel. Transfer of the blot was allowed to proceed overnight in 20x SSC buffer. The paper towels were replaced as they became wet. In the following morning the gel and the nitrocellulose membrane were turned over and laid, gel side up, on a dry sheet of 3MM paper. The positions of the gel slots on the filter were marked with an appropriate pen. The gel was then peeled off and assessed for the success of transfer by staining in ethidium bromide solution (0.5 µg/ml in water) for 45 min at room temperature and visualized under UV illumination.

The membrane was then washed in 6x SSC for 6 min at room temperature to remove the excess of the agarose gel that may remain sticking to the filter. The filter was then placed flat on a paper towel till it became dry and the DNA was fixed by baking the membrane at 80 °C for 2 h.

4.2.3.28. Hybridization of DIG-labelled Probe(s) to Target

4.2.3.28.1 General considerations and types of used probes:

Different types of probes (table 4.4) were labelled and used. Firstly synthetic probes with relative codons degeneracy were synthesized by MWG-BIOTECH, Ebersberg, Germany. Also a motif's gene from *E. coli* was isolated, purified and labelled for being used as a *heterologous* probe, using a low stringency of hybridization to allow a certain degree of mismatching between the DNA sequences. A third type was also used as a *homologous* probe at high stringency. These probes were fragments already isolated from the constructed genomic library of *P. vulgaris* or from the intact genome. These probes were applied either individually or in pool(s).

S. No.	Primer Name	Direction	Length	Origin	GC%	Sequence
1	T.1^{+}	+	21	HVOR N-terminal	42	AAC GGT TGG ACA GGT AAC ATT
2	N.$^{+}$	+	20	ydhV gene	55	GGC TAA CGG TTG GAC AGG TA
3	N.$^{-}$	-	21	ydhV gene	52.4	GTA GAA CAT CGT GAG CGA AGC
4	N.1^{+}	+	17	Clone.6*	58.8	GCC TGG TGG CAG AAG TA
5	N.1^{-}	-	18	Clone.6	50	AAG CGG TGA TTA CCG TAG
6	F.1^{+}	+	23	ydhV gene	34.8	GGT TGG ACA GGT AAT ATA TTA AG
7	P.340^{+}	+	19	ydhV gene	52.6	GAA GGG AAG GCG AAA TCA C
8	P.526^{-}	-	19	ydhV gene	52.6	CTT GTT CCT CTC TCT GGC T
9	P.3^{-}	-	20	ydhV gene	55	GCG AAA CAT TGC AGC GTC TG
10	P.4^{-}	-	19	ydhV gene	47.4	CA ATC TAC TGC CTG CGT AA
11	TEST.1^{+}	+	20	Clone.2**	50	GTT ATC TTT GCA TCC AGG CC
12	TEST.1^{-}	-	20	Clone.2	50	CTG ATC CTG TTT CCT GTG TG
13	R.1^{-}	-	20	HVOR N-terminal	50	GCT CCT GTC GTC AGA TTC AT
14	R.2^{-}	-	18	HVOR N-terminal	50	GAT AGC TCC AGT CGT CAA
15	M.80.$^{+}$	+	20	HVOR N-terminal	40	AAY GGI TGG ACI GGI AAY AT
16	M.80.$^{-}$	-	20	HVOR N-terminal	40	ATR TTI CCI GTC CAI CCR TT
17	M.80.2^{+}	+	20	*Cons.seq.*	46.7	ATH AAY YTN ACN ACN GGN CC
18	M.80.2^{-}	-	20	*Cons.seq.*	46.7	GGN CCN GTN GTN AYR TTH AT

* A 349 base DNA fragment fished from *Proteus vulgaris* genome with F.$^{+}$ & P.3^{-} primers

** A 934 base DNA fragment fished from *Proteus vulgaris* genome with N.$^{+}$ & N.$^{-}$ primers

Cons. seq.: Conserved sequence

Table 4.4: The pool of synthesized primers (oligonucleotides) of different sizes that have been used in hybridization experiments.

Figure 3 (A & B) below, shows the complete sequence of the **ydhV** (EG13956) motif's gene from *E. coli* wild-type as well as the fished fragments from *Proteus vulgaris* genome that have also been used as probes for hybridization:

```
5′-
   1  ATGGCTAACG GTTGGACAGG TAATATATTA AGAGTCAATC TCACGACAGG
  51  AAATATTACC CTCGAAGATT CCAGTAAGTT TAAAAGTTTT GTCGGTGGCA
 101  TGGGCTTCGG CTACAAAATT ATGTATGACG AAGTACCGCC AGGCACGAAA
 151  CCTTTCGATG AAGCGAATAA ATTAGTCTTT GCTACCGGCC CATTAACTGG
 201  ATCTGGTGCC CCCTGTAGTT CTCGCGTAAA TATCACCTCA CTTTCTACTT
 251  TTACCAAAGG AAATTTAGTC GTCGATGCCC ATATGGGTGG CTTTTTTGCA
 301  GCGCAAATGA AATTCGCTGG ATACGACGTC ATTATTATCG AAGGGAAGGC
 351  GAAATCACCG GTATGGCTGA AGATTAAAGA TGACAAAGTT AGCCTGGAAA
 401  AAGCCGATTT CTTATGGGGA AAAGGGACGC GCGCAACGAC GGAAGAAATT
 451  TGTCGATTGA CCAGTCCGGA AACCTGTGTG GCGGCTATTG GTCAGGCTGG
 501  GGAAAACCTT GTTCCTCTCT CTGGCATGTT GAATAGCCGT AACCACAGCG
 551  GCGGTGCGGG AACTGGCGCA ATAATGGGTT CGAAAAACCT GAAAGCGATT
 601  GCGGTTGAAG GGACGAAAGG GGTCAACATT GCCGATCGTC AGGAGATGAA
 651  GCGTCTCAAT GATTACATGA TGACTGAACT TATTGGTGCG AATAACAACC
 701  ATGTCGTGCC AAGTACGCCA CAATCGTGGG CAGAGTATTC AGATCCCAAG
 751  TCACGCTGGA CAGCACGTAA AGGGCTGTTT TGGGGCGCGG CTGAAGGTGG
 801  TCCGATTGAA ACGGGTGAAA TTCCGCCAGG CAATCAGAAT ACGGTCGGCT
 851  TTCGTACCTA TAAATCCGTT TTTGACTTAG GACCGGCGGC AGAGAAATAC
 901  ACAGTAAAAA TGAGCGGCTG CCACTCTTGC CCGATCCGTT GTATGACCCA
 951  AATGAATATT CCTCGGGTGA AAGAGTTTGG CGTGCCCAGC ACAGGTGGTA
1001  ACACTTGTGT AGCAAACTTT GTCCATACCA CCATCTTCCC GAACGGGCCG
1051  AAAGATTTTG AAGATAAAGA CGATGGTCGT GTGATTGGTA ACCTGGTGGG
1101  TCTGAATTTG TTCGATGACT ACGGCCTATG GTGTAACTAC GGGCAGTTGC
1151  ATCGCGACTT TACTTATTGT TACAGCAAAG GTGTGTTCAA GCGTGTTCTG
1201  CCAGCTGAAG AGTATGCAGA AATTCGCTGG GATCAACTGG AAGCGGGTGA
1251  CGTTAACTTC ATTAAAGATT TTTACTACCG TCTGGCGCAT CGTGTGGGTG
1301  AGCTGAGTCA CCTGGCTGAT GGTTCATATG CCATCGCAGA ACGCTGGAAT
1351  TTGGGTGAAG AGTACTGGGG CTACGCGAAA AATAAACTCT GGTCGCCGTT
1401  TGGCTATCCG GTTCACCATG CCAATGAAGC GTCAGCGCAG GTCGGTTCCA
1451  TTGTTAACTG TATGTTCAAC CGTGACTGCA TGACGCATAC CCATATCAAC
1501  TTTATTGGTT CCGGCTTGCC ATTGAAACTG CAACGTGAAG TGGCGAAAGA
1551  ACTTTTTGGT TCTGAAGATG CTTACGATGA AACCAAAAAC TACACGCCAA
1601  TCAATGACGC AAAAATTAAA TATGCAAAGT GGTCGCTGTT GCGGGTCTGT
1651  TTGCATAACG CCGTCACCCT GTGCAACTGG GTCTGGCCAA TGACCGTTTC
1701  GCCGCTGAAA AGCCGTAATT ACCGGGGCGA TCTGGCGCTT GAAGCCAAAT
1751  TCTTCAAAGC GATCACCGGC GAAGAGATGA CTCAGGAAAA ATTAGATTTA
1801  GCTGCAGAGC GTATTTTTAC GTTGCATCGT GCCTACACGG TAAAACTGAT
1851  GCAAACCAAA GATATGCGTA ATGAACACGA TCTTATCTGT TCCTGGGTAT
1901  TCGACAAGGA TCCGCAGATC CCGGTCTTTA CTGAAGGTAC TGACAAAATG
1951  GATCGTGACG ATATGCATGC TTCGCTCACG ATGTTCTACA AAGAAATGGG
2001  CTGGGACCCA CAGCTTGGTT GTCCAACCCG CGAAACATTG CAGCGTCTGG
2051  GGCTGGAAGA TATTGCCGCC GATCTGGCAG CACACAATCT ACTGCCTGCG
2101  TAA
         -3′
```

Fig. 3 (A): Nucleotide sequence of ydhV gene (total length is 2103 base).

Clone 2

```
5′-
   1  TGATCGTCCA GATCGAGGCC AAGCTCAACG AGAAGAACAT TCCGCGCAAC
  51  ATGATCGGGC GCGAGAAGCG CGTGGTGGCG CTGGAGCAGT ACCTCTCGCA
 101  GGCGCGCAAT TACGACCCCG TGCTCGACGG CCTGCGCTCG GCGGTCCGCT
 151  ACGACAAGAC CTACTTCGAC AAGATCGTCG CATCGTTGCT GCCGCTGCTG
 201  GAGAAGCTCA CGAGCGGCAA GATCGCTCAG TTGTTGGCCC CGAACTATTC
```

```
     251  GGATCTCAAC GATCCCAGGC CGATCTTCGA TTGGATGCAA GTGATCAGGA
     301  AGCGCGCCGT GGTCTATGTG GGCCTGGATG CAAAGATA
                                                  -3′
```

Clone 6

```
5′-
       1  TTCGTCTTAT GGAAAGCACT GCCGGAACTT CATGACGAGA GTTACCGTTT
      51  TGTTCCGGGG GCGGTGTTGA CACTGCGCGA GGGGGAACAT GTGGCGCTGG
     101  TGGCAACTGG CTCGACAGTT CATGAAATTG TTGATGCCGC TGCGCTGTTG
     151  GCTGATGCAG GTATTCAGGC GAAAGTGGTC AGTGTACCTT CAATTCGACC
     201  ATGTGATACC AAAGCTCTGT TATCAGTATT ACAGGGCTGC AAAGCGGTGA
     251  TTACCGTAGA AGAGCACAAT ATTAATGGTG GGTTGGGAAG CCTGGTGGCA
     301  GAAGTACTGG CTGAGGGCGG AGTCGGGGCA GTGTTAAAGC GTTTAGGTAT
     351  TCCCGATGGA GAGTACCCAG CGGTTTTTTT ATCTTGGCTG GCTACCCCCA
     401  CCATCATGGT TTTGACGCCG CATCTATCGC TGCTCATGCG CACAGAAAAG
     451  ATGTGACGTT CACACTGCTG CCACTGTCAT TACCTCATTG CCCCAGAACC
     501  TTCCTTCGCA CCTGATACTT TACCATGCCA GCCCTCACCA TCATTATCGC
     551  TCTCCCTAAC CCGATCCATT AACTACTGAA AACACCTCCC CGATTGCGTC
     601  GCCAAAACCC GCACCCATTT CACCTCAAAC CCACTCTCCT TTCTTCATCC
     651  CTCACCACGC CCCTTCACAC AAACCCCAAC ACCGACACAC ACCACAACCA
     701  AAATGCCCAC TTTCCCAACT ACCCCGACCA CCCCGCCAAT TCCATACACA
     751  CTACCACTTA CACGCCCACC ACTCACCATC CCGTCACCCA CCACTACCAA
     801  CCATCCACTT CCCCAAGCAA CTACCATCAC TCGCACCCAT CGCTCGACCT
     851  TTCAACACCA CAACTCCCCT CTCCTACCCT CCCCACCCTT TCCCCCTCGC
     901  CCAATCCACC CAAAATTCAT TCCCCTCCCT CCAC
                                              -3′
```

Fig. 3 (B): Nucleotide sequences of 2 fished fragments (clone 2 & clone 6) from *Proteus vulgaris* genome.

4.2.3.29 Fishing of ydhV gene from *Escherichia coli* K-12 wild type to be used as a heterologous probe for low stringency hybridization

The ydhV gene was fished from *E. coli* to be used as a heterologous probe for hybridization at low stringency conditions. The gene was partially fished through hybridizing of some synthesized probes to the plated phage genomic library of *Proteus vulgaris*. The rest of the gene sequence was obtained from the data bank while searching for homologous protein segments matching the obtained short sequence.

The intact genome of *E. coli* K-12 wild-type was used as a template for fishing the gene using PCR technique. For achieving this, 1 sense primer in addition to 2 antisense primers were synthesized by MWG, Ebersberg, Germany. By applying different annealing temperatures and by adjusting the reaction conditions (such as the concentration of the reagents, Mg^{2+}, and primers), the fishing of the motif's gene was carried out successfully.

4.2.3.29.1 Primers used:

Sense: F.1+
Sequence : 5'-GGT TGG ACA GGT AAT ATA TTA AG-3'
Antisense: $P.3^{-}$

Sequence : 5'-GCG AAA CAT TGC AGC GTC TG/3'
Antisense: P.4^{-}
Sequence : 5'- CA ATC TAC TGC CTG CGT AA-3'

4.2.3.29.2 PCR Program:

1x Initial denaturation at 95 °C for 5 min

Denaturation, at 95 °C for 5 min

Annealing, at 55 °C for 40 sec

Elongation, at 72 °C for 1 min

Final elongation, at 72 °C for 7 min

The fished gene was then cloned in pGEM-T Easy vector purshased from Promega for sequencing and preservation. The cloning of the fished gene was as follows:

4.2.3.29.3 A-tailing:

After purification of the PCR product (ydhV gene) with StrataPrep PCR purification kit, the DNA to be cloned was A-tailed. That means a deoxyadenosine base was added to the end of the PCR generated fragments. These A-tailed fragments would have been compatible overhangs to the single 3'-T overhangs of the vector ends which were prepared by cutting with *Eco*RV followed by adding a 3' terminal thymidine to both ends.

Using this method, only one insert would be ligated into the vector as opposed to multiple insertions that could occur with blunt-ended cloning.

The A-tailing reaction was set up in 0.5 ml microcentrifuge tube as follows:

1.0 µl of *Taq* DNA polymerase 10x reaction buffer

4.0 µl of the purified PCR product

1.0 µl of dATP (final conc. of 0.2mM)

1.0 µl of *Taq* DNA polymerase (5 units)

3.0 µl st. dist. water

———

Total vol.: 10µl

The reaction was then incubated at 72 °C for 30 min. The A-tailed DNA was applied directly to ligation reaction taking into consediration the insert: vector ratio.

4.2.3.29.4 Ligation reaction:

A ligation reaction was prepared in 0.5 ml microcentrifuge type as follows:

5 µl 2x rapid ligation buffer,
1 µl pGEM-T Easy vector (50 ng)
2 µl PCR product (ydhV gene from *E. coli*)
1 µl T4 DNA ligase (3 units/µl)
1 µl st. dist. water

Total vol.: 10 µl

The contents were mixed by pipetting and the reaction was incubated overnight at 4 °C.

4.2.3.29.5 Transformation:

2 - 3 µl of the ligated product was pipetted into a chilled 1.5 ml sterile microcentrifuge tube. 50 µl of the on ice thawed JM 109 competent cells was added to the ligation product and the contents of the tube were mixed by flicking the tube. The tube was left on ice for about 20 min. The transformation was carried out by heat-shocking the cells for 45 sec in a water bath at exactly 42 °C. The tube was then kept on ice for 2 min 950 µl of SOC medium warmed to room temperature, was then added and the tube was incubated at 37 °C for 1.5 h with constant shaking.
After incubation 100 – 200 µl of the transformed cells were plated onto LB-ampicillin (0.05 mg/ml) agar plates supplemented with 0.5 mM IPTG and 80 µg/ml X-Gal.
The plates were then inverted and incubated at 37 °C overnight and up to 24 h to allow the blue colour of the non-transformed cells to develop. The development of the blue colour can be enhanced by incubating the plates for further 2 - 4 h at 4 °C. White colonies with plasmids containing the insert of interest were then selected for plasmid preparation (see 4.2.3.21).
The orientation of the cloned gene within the cloning vector as well as the sequence was determined by sequencing the insert by MWG, Ebersberg, Germany. A preliminary method was also used for determining the orientation of the insert. In this method, a restriction digestion of the gene was carried out using *Bam* HI endonuclease enzyme which recognizes the G↓GATCC in 2 sites (cuts at 1908 and 2103) within the ydhV gene, leading to the generation of 2 fragments of different sizes, 1908 and 195 bases, respectively.

4.2.3.30 Determining optimal hybridization temperature

Three factors have always to be considered when determining the optimal hybridization temperature. These factors are temperature, salt concentration and formamide concentration.

Factor	Influence
Temperature	* High temperature increases stringency * Low temperature decreases stringency
Salt conc.	* High salt decreases stringency * Low salt increases stringency
Formamide	Decrease melting point of DNA, thus lowering the temperature at which a probe-target hybrid forms (adding 1 % formamide lowers the melting Temperature by 0.72 °C)

A combination of high temperature and low salt increases the hybridization stringency.
The relative strength of different hybrids is:

RNA : RNA hybrids > RNA : DNA > DNA : DNA hybrids

Hybrid types will influence the hybridization temperature used. **RNA : RNA** and **RNA : DNA** hybrids will require higher hybridization temperature than **DNA : DNA** hybrids. **Optimum Temperature (T_{opt})** will vary according to the GC content of the probe and homology to the target. To determine the optimal hybridization temperature (T_{opt}), firstly the melting temperature T_m of the probe-target hybrid was calculated, then T_{opt} was set to a value that is 20 – 25 °C below the calculated T_m.
DIG Easy Hyb was used because it lowers the melting point by the same amount (36 °C) as a 50 % formamide (i.e. by 0.72 °C per 1% formamide) which leads to more reproducible results. It is also non-toxic and safe to use, since it does not contain formamide. In place of formamide, it contains enough urea to lower the hybridization temperature as if it would contain 50 % formamide.

4.2.3.31 Hybridization time required

Prehybridization time was ranging from 30 min up to 3 h depending upon the type of probe and the sensitivity required, while the hybridization time was about 12 - 16 h.

4.2.3.32 Stringent washes of Blots

Two-stage stringent washes were performed after hybridization to disrupt undesired hybrids:
First washes: low stringency (high salt concentrations and low temperatures), to remove nonspecifically bound probes. The membranes were washed 2 x 5 min with 2x SSC containing 0.1 % SDS at room temperature.
Second washes: High stringency (low salt concentrations and high temperatures), to remove undesired hybrids of low homology. The high stringency buffer was prewarmed to the correct

temperature (68 °C) before adding to the membrane. Otherwise, low homology hybrids might not be disrupted during the short washes. The membranes were washed 2 x 15 min in 0.1 % SSC containing 0.1 % SDS at 68 °C

General procedure: (The exact amounts of buffers differ according to the size of each membrane) According to Baas, 1998 the hybridization was carried out as follows:
The membrane was laid on a piece of nylon net and rolled up into a tube and then placed in a hybridization tube followed by adding the adequate amount (according to the size of the filter) of prehybridization buffer. The filter was then prehybridized at the recommended temperature (specific for each reaction) for 3 h. The prehybridization buffer was then replaced by hybridization buffer containing the labelled probe in the appropriate concentration and the hybridization was performed overnight at the calculated T_m. The membrane was left inside the hybridization tube and washed 2 x for 5 min at room temperature with 2x SSC / 0.1 % SDS followed by 2 washings at the hybridization temperature for 15 min with 0.2x SSC / 0.1 % SDS.

4.2.3.33 Chromogenic Detection of Probe-Target Hybrids

4.2.3.33.1 General considerations

Nitrocellulose membranes are suitable for colorimetric detection assays but give lower sensitivity than nylon membranes. DNA cannot be UV crosslinked to the membrane.
The developed signals were detected with an enzyme-linked immunoassay which is more sensitive than radioactive procedures. In this assay, the membrane was blocked to prevent non-specific interaction of the antibody with the filter.
Alkaline phosphatase-conjugated antibody, specific for digoxigenin, recognizes the DIG molecule on the labelled hybrid was added followed by addition of an alkaline phosphatase substrate (NBT/BCIP) allowing the visualization of the hybrids.

Procedure:

The detection of the positive hybrids was achieved at room temperature. The membrane (still in the hybridization tube) was washed for 1 min with DIG buffer 1 / 0.3 % Tween 20 for equilibration and 30 min with DIG buffer 2.
The detection reaction was done by immersing the membrane in freshly prepared detection buffer. The excess of detection buffer was removed by washing the membrane 2x with DIG buffer 1 followed by another wash with DIG buffer 3 for equilibration.

The membrane was then smoothly slid into a plastic bag that contains an appropriate amount of the colour developing solution. Colour developing was allowed to proceed by leaving the membrane in the dark overnight (up to 16 h).

The membrane was then soaked in buffer 4 for 10 min for stopping the color developing reaction. The membrane was air dried at room temperature and kept sandwiched between 2 pieces of 3 MM papers in the dark.

4.2.3.34 Intact DNA as a template for PCR screening experiments

Isolated and purified DNA (see 4.2.3.2) from *Proteus vulgaris* was used as a template for PCR experiments. The DNA was physically sheared either by pipetting several times with a fine tip or by vortexing for an appropriate time with fine glass beads (0.25 - 0.5 mm). The sheared DNA was then allowed to denature at 95 °C for longer time than as usual. In another experiment, the DNA was restricted with either *Bam* HI, *Eco* RI or *Not* I for 1 h at 37 °C. Following restriction digestion the DNA fragments were separated on 1 % agarose gel (4.2.3.6) and Bands in the range of 21 – 15 kbp, 14.5 – 10 kbp and 9.5 – 1 kbp were separately extracted from the gel (4.2.3.7).

5. Results

5.1 Enzyme purification

5.1.1 Ion exchange chromatography on DEAE-S

After cell lysis (4.1.1.4), the membrane fraction extracted from 20 g of wet packed cells (119.8 mg/ml) was loaded immediately on the first separation column (DEAE-S) to avoid the rapid loss of the activity since the enzyme is very sensitive to oxygen. As reported by Neumann (1985) and Trautwein (1993), the elution of the enzyme was supposed to occur at 0.2 M KCl. So the fractions were collected in a pool. The protein concentration (4.1.7) was then determined as well as the specific activity (4.1.2). The pool was then concentrated (4.1.4). The selected fractions were then separated on 12 % SDS-PAGE (4.1.5). The estimated specific activity in the supernatant of 10.5 U/mg was decreased dramatically to 3.6 U/mg after the first run of purification although the concentration of the dithionite was kept at the minimum. Dithionite has an inhibiting effect on the HVOR due to the formation of some oxidized products (Trautwein, 1993).

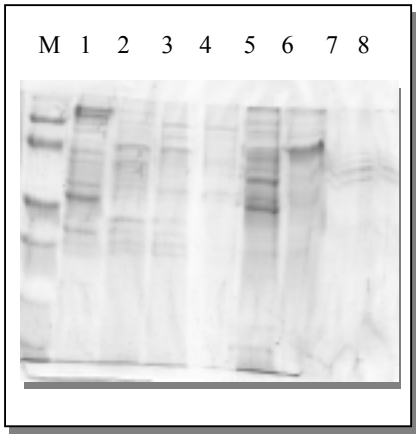


Figure 5.1: 12 % SDS-PAGE of different eluted fractions after the separation on DEAE-S column.
M: Low range marker
Lanes 1 – 8: different fractions.

5.1.2 Separation on Hydroxyapatite chromotography

A pool (protein conc. 39.5 mg/ml) of the collected fractions that showed remarked specific activity after the separation on DEAE-S was then loaded on the hydroxyapatite column. The elution of the enzyme was at a concentration of 80 - 100 mM potasium phosphate buffer (B 6). The fractions which showed activity were again collected and a pool of them was set up. The estimated protein

concentration of the pool was 17.7 mg/ml and the specific activity was dropped again to 0.8 – 1.1 U/mg.

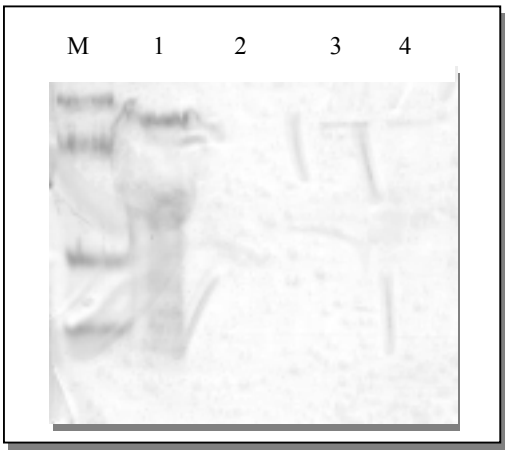


Figure 5.2: 12 % SDS-PAGE of some eluted fractions after separation of HVOR on hydroxyapatite column.
M: Low range marker
Lane 1: pool (of fractions 3, 4 and 5)
Lanes 2, 3 and 4: different fractions

5.1.3 SEC on Superdex-200

Size exclusion chromatography on Q-Sepharose or on Superdex-200 is the last step in the purification of the HVOR enzyme. The eluted fractions were assayed individually for the presence of a considerable specific activity. At the last steps of the purification of any enzyme, the detection of the highest activity is always expected. The highest specific activity that could be detected was 14 U/mg.

Figure 5.3: 8 – 16 % gradient SDS-PAGE
M: Low range marker
Lane 1: End product

This determined specific activity was not in consistence with the previously detected specific activity by Trautwein (1993) which was ≥ 1200 U/mg for the most pure preparations of the enzyme.

5.1.4 Isoelectric point

At this level of the purification process, homogenous bands on SDS/PAGE were obtained (Figures 5.2 & 5.3). Due to the rapid loss of the enzyme activity (although polidocanol was always added to all the buffers and the dithionite concentration was kept at minimum), it was hard to confirm that these bands are of the HVOR enzyme. The isoelectric point is 5.1 as it was estimated with PhastSystem IEF media pH 3 – 9 by Trautwein et al., 1994. The isoelectric point of some of the purified protein fractions were determined to confirm the isolation of the target protein (see 4.1.8). The estimated values of the isoelectric point of these fractions were lower than that reported by Trautwein *et al*., 1994. Also the molecular mass of the purified protein (70 – 75-kDa) was not completely consistent with that (80 kDa) reported by them.

5.2 Construction of a genomic library of *Proteus vulgaris*

It was always recommended to use a fresh genomic library which should have a titer of 10^8 - 10^9 to ensure the representation of the single copied sequences.

5.2.1 Preparing of the genomic DNA from *P. vulgaris*

DNA was isolated using the "Qiagen Genomic tip 100/G" kit, according to the provided protocol. The yield of DNA depends on the number of bacteria in culture and on the size of the bacterial genome. Culture volume was determined according to the manufacturer recommendations to avoid overloading the genomic tips (which if it were to occur, it would lead to reduced performance of the system accompanied with low yield). The eluted DNA was then purified and concentrated by ethanol precipitation (see 4.2.3.4). The DNA yield was estimated quantitatively (see 4.2.3.3). The length of the genomic DNA was determined on 1 % agarose gel.

The average yield and purity of genomic DNA prepared with these Genomic-tips (100/G) under optimal conditions were 85 - 95 µg and 1.71 respectively starting with 2.2 x 10^{10} bacterial cells. This bacterial density could be reached by growing the cells of *P. vulgaris* to an OD_{600} value of 2 - 4. The results were not consistent with the recommendations of the manufacturer since the yield was always between 25 - 50 µg and the purity was in the range of 1.4 – 1.56. On 1 % agarose gel (Figure 5.4), the isolated DNA matched the first band (21.226 kbp) of the λ-DNA marker which has an average length of about 50 kbp.

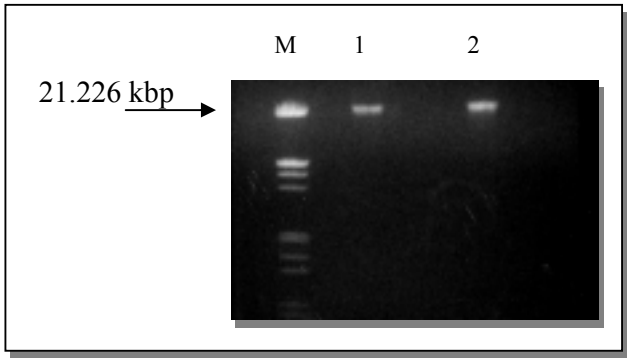


Figure 5.4: Genomic DNA from *P. vulgaris*.

M : ג-DNA marker restricted with *Hind* III and *Eco* RI.

Lanes 1 and 2 : 1 µl (c. 200 ng) genomic DNA.

5.2.2 Partial digestion of genomic DNA with restriction enzymes

5.4 - 8.1 µg of genomic DNA was digested with 4U / µl of *Sau* 3AI. This partial digestion depends on the gradual dilution of the used enzyme until reaching the optimal enzyme concentration by which the genomic DNA will be restricted efficiently under the definite parameters of temperature and digestion time (see 4.2.3.5).

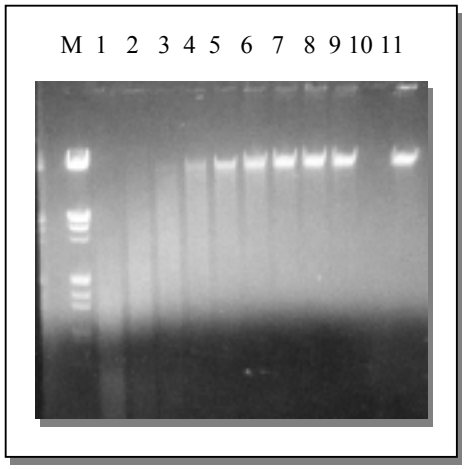


Figure 5.5: Partial digestion of the genomic DNA of *P. vulgaris*.

M: ג -DNA marker restricted with *Hind* III and *Eco* RI.

Lanes 1– 9 : A gradual dilution of *Sau* 3AI.

Lane 11 : Genomic DNA (c. 1 µg).

It is supposed that, the fragments lying between 5 – 10 kbp are representative to the genome of *P. vulgaris*. So fragments lying in this range (from lanes 3 – 5; fig. 5.5) were excised and the DNA was isolated from the agarose gel. For obtaining the adequate amount of DNA for subsequent construction of the genomic library, an additional probe was prepared.

5.2.3 Amplifying the primary genomic library

The construction of the primary genomic library of *P. vulgaris* was carried out as outlined in sections 4.2.3. A genomic library of 4.2 x 10^7 pfu/µl was obtained. It is usually recommended to amplify (section 4.2.3.13) the primary genomic library only once since slow growing clones may be significantly underrepresented. After only one round of amplification, a genomic library of 6.7 x 10^8 pfu/µl was obtained.

5.3 Screening of the genomic DNA library by plaque hybridization

A genomic library was constructed in ג-ZAP Express vector in order to be screened for a clone containing the complete sequence of the HVOR enzyme or a part of it. The library was also screened for the presence of an oxidoreductase domain especially the aldehyde oxidoreductase (AOR) domain.

The screening of the genomic library was achieved by applying two techniques. The first was the plaque hybridization and the second was the polymerase chain reaction (PCR). In plaque hybridization, the phages were plated and lifted as outlined in sections, 4.2.3.14 and 4.2.3.15. The membranes were hybridized to the labelled probes (see section 4.2.3.28).

Plaques that showed positive signals were cored out from the agar plates and transferred to a 1.5 ml microcentrifuge tube containing SM buffer.

Secondary and tertiary screening of these plaques was carried out. Single clone excision and plating of the excised phagemids were carried out according to 4.2.3.16 and 4.2.3.17. The XLOLR strain colonies that were supposed to contain pBK-CMV double stranded vector with the cloned DNA insert underwent plasmid isolation and sequencing (see 4.2.3.21 & 4.2.3.23).

A total of 13 plaques which gave positive signals had been reported. This number was reduced from 13 to 6 plaques after the second screening. After the third round of screening, only 2 plaques (Table 5.1) were selected for subsequent single clone excision and plating of the excised phagemid vectors.

At this level, one obtained colonies in XLOLR strain that contain plasmids (phagemid vectors) with the inserts of interest. The plasmids were then isolated and purified as outlined in section 4.2.3.21.

The purified plasmids were then subjected to sequencing reaction (4.2.3.23) and the sequences were read by GATC.

The used probes were designed based on the amino acid sequence of the N-terminal of HVOR. The N-terminal sequence of *P. vulgaris* HVOR is highly similar to N-terminals of two tungsten-dependent aldehyde oxidoreductases from *Clostridium thermoaceticum* and *Thermococcus litoralis* (Trautwein *et al*., 1994).

Fish name	Probe name	Sequence	Origin	Length* (bp)
Fish.1	$T.1^{+}$	5'-AAC GGT TGG ACA GGT AAC ATT-3'	*P.vulgaris*	215
Fish.2	$T.1^{+}$	5'-AAC GGT TGG ACA GGT AAC ATT-3'	*P.vulgaris*	93

Table 5.1 shows the used probe and the length of the resulted sequences.

The result of this first hybridization experiment revealed that the used probe was not specific for fishing the N-terminal of the HVOR enzyme or the above mentioned N-terminals of AOR from *C. thermoaceticum* and *T. litoralis*. Also the results showed that none of the above mentioned clones contains the complete gene sequence or the N-terminals of HVOR and the other 2 AOR enzymes.

5.4 Screening of the genomic DNA library by PCR

The second technique that had been employed for screening the genomic library of *P. vulgaris* was the polymerase chain reaction (PCR) (see 4.2.3.19). 2 sequences were fished out of a relatively large number of fragments. These 2 sequences namely clone 2 and clone 6 have received a special attention.

Clone name	Primers	Primer sequence (5' → 3')	Length* (bp)
Clone 2	$F.1^{+}$ & $P.3^{-}$	GGT TGG ACA GGT AAT ATA TTA AG GCG AAA CAT TGC AGC GTC TG	338
Clone 6	N^{+} & N^{-}	GGC TAA CGG TTG GAC AGG TA GTA GAA CAT CGT GAG CGA AGC	934

Table 5.2: shows the used primers and the length of the resulting sequences.

The above mentioned primers ($F.1^+$, $P.3^-$, N^+ and N^-) in Table 5.2 were designed based on the similarity of the N-terminal sequence of *P. vulgaris* HVOR to the N-terminal of a hypothetical protein encoding the ydhV gene (AC74743.1) which resulted during the complete sequencing of the *Escherichia coli* k-12 genome (Blattner *et al*., 1997).

This700 AA sequence has a molecular weight of about 77.876 kDa. The HVOR N-terminal is 80 % identical (and 90 % positives) with the first 20 amino acids of the N-terminal of the ydhV gene.

Due to this high similarity, the ydhV sequence was proposed to be the base of synthesizing oligonucleotides for being used as heterologous probes in hybridization experiments. Also a number of primers had been synthesized to be used in PCR experiments (see table of oligonucleotides, page No. 23).

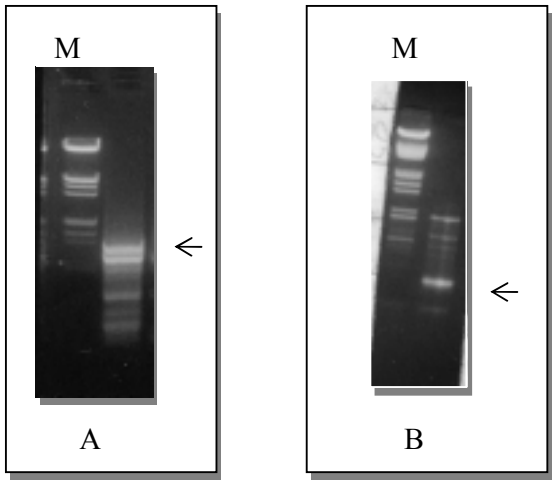


Figure 5.6: Bands fished with primers designed based on the ydhV motif's gene.

M: λ-DNA marker restricted with *Hind* III and *Eco* RI.

(A): Fragment fished with primers N^+ & N^-

(B): Fragment fished with primers $F.1^+$ & $P\text{-}3^-$

In figure 5.6 (A & B), the arrows are pointed to the DNA bands that were extracted, cloned in pGEM-T Easy vector and sequenced. A search of the genomic database with the DNA sequences revealed the similarity of clone 6 amino acid sequence to transketolase (EC 2.2.1.1) from *Yersinia pestis* (strain CO92) with a percentage of identity of 62 % and a percentage of positivity of 70 %. Another similarity was also reported to pyruvate dehydrogenase E1 (lipodoamide) beta subunit pdhB from *Bacillus halodurans* (strain C-125). The identity was 28 % and the positives were 46 %. New primers had been synthesized in a trial for fishing the rest of each sequence from the genomic library. These primers were used in combination with the previous primers. Table 5.3 summarizes the essential information of these new primers.

Primer name	Origin	Sequence (5' → 3')	Length	GC %
Test.1^{+}	Clone 2	GTT ATC TTT GCA TCC AGG CC	20	50
N.1^{-}	Clone 6	AAG CGG TGA TTA CCG TAG	18	50

Table 5.3: New synthesized primers for fishing the rest of clone 2 and clone 6.

PCR reactions were then carried out aiming at fishing the rest of the above mentioned clones (clone 2 and clone 6) using the newly synthesized primers in combination with the previously used ones as follows:

For fishing the rest of clone 2:

Primer Test.1^{+} (forward) and
Primer P.3^{-} (reverse) were used.

For fishing the rest of clone 6:

Primer N^{+} (forward) and
Primer N.1^{-} (reverse) were used.

These reactions resulted in the fishing of two new sequences named clone 2.1 and clone 6.1 respectively (Table 5.4).

Clone name	Used primers	Sequence length*
Clone 2.1	Test.1^{+} & P.3^{-}	1744 bp
Clone 6.1	N^{+} & N.1^{-}	1234bp

Table 5.4: New sequences fished with primers: Test.1^{+} & P.3^{-} → clone 2.1 and
N^{+} & N.1^{-} → cone 6.1

Figure 5.7 shows the fished fragments using the above mentioned primers (Table 5.4). Different fragments were fished with both sets of primers. The bands were extracted from the agarose-gel as outlined in section 4.2.3.7. The sequencing of the extracted DNA revealed that band 1/1 and band 3/2 were the bands of interest.

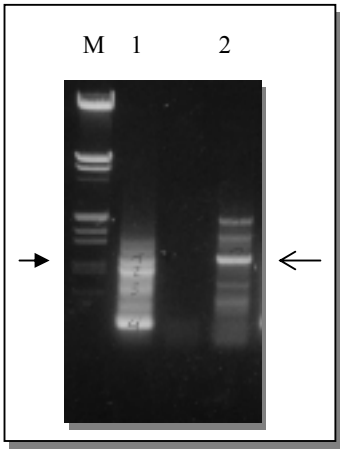


Figure 5.7: Fished bands.

M: ג -DNA marker restricted with *Hind* III and *Eco*R I.

Lane 1: Different fragments fished with primers Test.1^+ & P-3^-

Lane 2: Different fragments fished with primers N^+ & N^-

Arrows indicate the bands of interest.

By mapping the amino acid sequences obtained above in all the 3 frames, it was clear that the N-terminal of the HVOR not fished. This revealed that the used primers were not specific and other primers should be designed to fulfill this approach.

5.5 Screening of the *P. vulgaris* genome by blot hybridization

Blot hybridization was used for screening the *P. vulgaris* genome for the presence of oxidoreductase genes. The blotting was achieved as described in section 4.2.3.27.1. The hybridization and chromogenic detection of probe-target hybrids were done according to section 4.2.3.28 and section 4.2.3.33 respectively. For the types of probes that had been used see Table 4.2.

As mentioned above, the used primers were nonspecific. Due to this fact new oligonucleotides were synthesized based on the concept that, degeneracy of the primers and other hybridization probes could offer the solution for this problem. The other type of probes was the use of the ydhV isolated gene as a heterologous probe for carrying out hybridization at low stringency conditions (see 4.2.3.30 – 4.2.3.32). The motif's gene from *E coli* was isolated (Figure 5.8 A) and cloned in pGEM-T Easy vector as described in section 4.2.3.29. To be used as a probe in hybridization, the fished gene was digested with *Bam* HI restriction enzyme (Figure 5.8 B). The digestion of the ydhV gene resulted in two unequal fragments since the enzyme cuts the gene at positions 1908 and

2103 respectively. The restriction digestion resulted in two fragments with different sizes. The Lengths of them were 1908 and 195 bases, respectively.

The fragments were then separated on 1 % agarose gel. Following the extraction of the two bands from the agarose gel (see 4.2.3.7), the DNA was labelled (see 4.2.3.26) and was to be hybridized to the blotted DNA (see 4.2.3.28). For maintaining these 2 fagments for further use, they were cloned in pGEM-T Easy vector as outlied in section 4.2.3.29. Plasmids contain them were digested with *Nco* I & *Spe* I restriction enzymes (figure, 5.8 C) as they were needed.

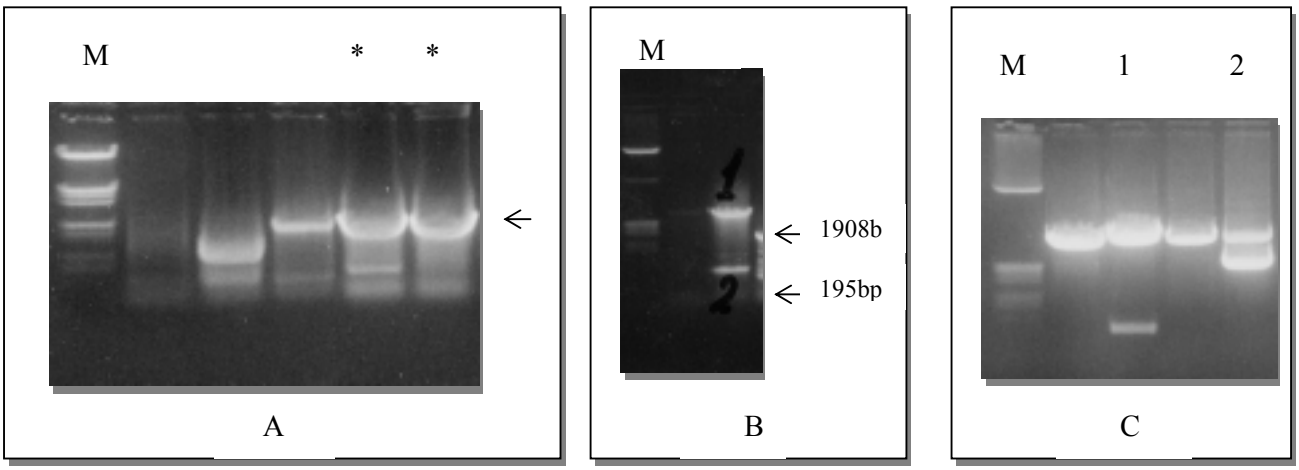


Figure 5.8: Fished ydhV motif's gene from *E .coli*

A: * Fished ydhV gene

B: Restriction digestion of the ydhV gene with *Bam* HI

C: Plasmid digestion with *Nco* I & *Spe* I shows:

Lane: I small fragment (c.195 bp) of ydhV motif's gene;

Lane: II large fragment (c.1908 bp) of ydhV motif's gene

M: ג -DNA marker restricted with *Hind* III and *Eco* R I.

Due to the expected higher mismatching percentage, the adjustment of the hybridization optimal conditions was not possible even when low stringency conditions were applied.

To overcome this problem it was much better to use relatively short oligonucleotides. These short probes (see table 4.2) were of different types and varied in length from 17 to 787 bases. Some of them were designed on the basis of the N-terminal of the HVOR enzyme and its homologous proteins. Others were already used before while trying to screen the genomic Library of *P. vulgaris*. The third type was degenerated oligonucleotides designed on the basis of the conserved

amino acid sequences available in protein databases. These probes were used either solitarily or in pools.

The hybridization was carried out as described in section 4.2.3.28. The temperature of the hybridization and the concentration of the labelled probes were the most important factors. Beginning with low temperature (50 °C) enabled the development of positive signals. The number of positive signals was markedly reduced by elevating the annealing temperature (up to 68 °C). The probes were used at first in a pool containing all the 18 oligonucleotides stated in Table 4.2. Another set of membranes were probed with the same method of hybridizing the probe to target DNA (4.2.3.28). In this second round of hybridization, the number of the used probes was reduced. The process was repeated until the appropriate conditions of hybridization and the best probe had been detected.

Other hybridization experiments were carried out using 2 fragments fished from the genomic library of *P. vulgaris*. The first one was fished with primers $F.1^+$ and $P.3^-$ (clone 2) and the other one was fished with N^+ and N^- primers (clone 6) (see Table 5.2).

Figure 5.9 shows the result of a hybridization experiment in which a pool of all the stated oligonucleotides in table 4.2 was applied. Following hybridization with this pool of oligonucleotides, the membrane was reprobed using only 2 probes ($M.80.2^+$ and $M.80.2^-$) because these two probes were highly degenerated to accommodate all the possible probabilities.

A B

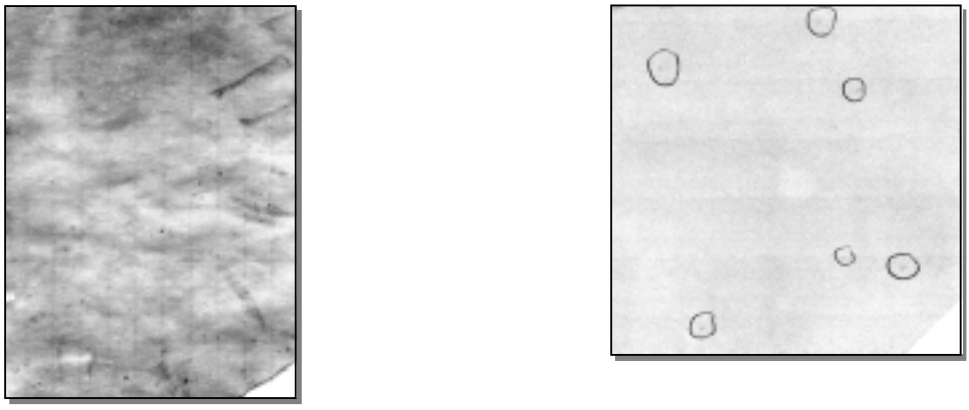


Figure 5.9 : Blotted nitrocellulose membranes hybridized to a pool of oligonucleotides (A) and to $M.80.2^+$ and $M.80.2^-$ oligonucleotides (B)

In order to place the nitrocellulose membranes evenly on the gel (alignment of the membrane to the gel), it was necessary to rewet the membranes with sterile distilled water for an adequate time until the membranes restore its original size.

The gels were not allowed to transfer the entire DNA to the nitrocellulose membrane in order to extract the DNA of the band which was expected to be matched to the positive signal. Sometimes, when it was proved that the transfer was complete and no more DNA was still present in the gel, it was expected to run a new agarose gel under the same conditions as the original one and to excise wider bands of DNA that may contain the DNA of interest.

The excised bands were then subjected to DNA extraction and purification as described in 4.2.3.7. The extracted DNAs were then used as templates for PCR reactions. The same oligonucleotides which had been previously used in hydridization were used (individually) in these PCR reactions.

The result of these PCR reactions was the fishing of different sequences (Table 5.5), which were further cloned and sequenced (see 4.2.3.20 – 4.2.3.23).

Clone name	**Primers Used**	**Primers sequence 5' → 3'**	**Origin**	**Length* (bp)**
Clone I	$N._{+}$	GGC TAA CGG TTG GAC AGG TA	ydhV*	1299
Clone II	$N._{+}$ & $N._{-}$	GGC TAA CGG TTG GAC AGG TA GTA GAA CAT CGT GAG CGA AGC	ydhV ydhV	798
Clone III	$T.1_{+}$	AAC GGT TGG ACA GGT AAC ATT	(N-ter.HVOR)	980
Clone IV	$F.1_{+}$ & $P.4_{-}$	GGT TGG ACA GGT AAT ATA TTA A(CA ATC TAC TGC CTG CGT AA	ydhV ydhV	783
CloneV	$F.1_{+}$ & $P.3_{-}$	GGT TGG ACA GGT AAT ATA TTA GCG AAA CAT TGC AGC GTC TG	ydhV ydhV	837
Clone VI	$Test.1_{+}$	GTT ATC TTT GCA TCC AGG CC	Clone 2*	170
Clone VII	$P.340_{+}$ & $P.526_{-}$	GAA GGG AAG GCG AAA TCA C CTT GTT CCT CTC TCT GGC T	ydhV ydhV	60
Clone VIII	$T.1_{+}$ & $P.3_{-}$	AAC GGT TGG ACA GGT AAC ATT GCG AAA CAT TGC AGC GTC TG	(N-ter.HVOR) ydhV	407
Clone IX	$F.1_{+}$	GGT TGG ACA GGT AAT ATA TTA	ydhV	294

* ydhV: Fished motif's gene from *E. coli* k-12 wild type

Table 5.5 shows some of these fragments and the used primers which had been used.

All the fished fragments were shared in the absence of the amino acid sequences of the HVOR N-terminal and N-terminals of the similar proteins from other organisms (see 5.2 & 5.3). This common feature reveals that the used primers were nonspecific probes or the target N-terminal of the HVOR within the genome of *P. vulgaris* is highly degenerated, the thing which leads to false priming.

5.6 Intact DNA as a template for PCR screening experiments

The DNA was handled as described in section 4.2.3.34. It was then used as a template for PCR reactions.

There were nearly no distinct differences between the use of the intact DNA or the genomic library as a template for the purpose of screening by applying PCR technique. A total number of five fragments was fished from the genome of *P. vulgaris* when sheared DNA was used (Table 5.6) as a template instead of the DNA of the genomic library.

Clone name	Primers Used	Primers sequence 5' → 3'	Origin	Length* (bp)
Clone S.1	M.80.$^{+}$ & P.3^{-}	AAY GGI TGG ACI GGI AAY AT GCG AAA CAT TGC AGC GTC TG	*P.vulgaris* (N-terminus) ydhV	888
Clone S.2	M.80.$^{-}$	ATR TTI CCI GTC CAI CCR TT	*P.vulgaris* (N-terminus)	1235
Clone S.3	M.80.2^{+} & P.4^{-}	ATH AAY YTN ACN ACN GGN CC CA ATC TAC TGC CTG CGT AA	*P.vulgaris* (N-terminus) ydhV	686
Clone S.4	M.80.2^{-}	GGN CCN GTN GTN AYR TTH AT	*P.vulgaris* (N-terminus)	1435

Table 5.6 Fished fragments when sheared DNA was used as a template for PCR experiments.

Another set of fragments was fished from the restricted DNA (see table 5.7). The fragments were named clone E, clone B and clone N referring to the restriction enzymes which had been used for restriction digestion.

Clone name	Primers Used	Primers sequence 5' → 3'	Origin	Length* (bp)
Clone E.1	C.1^{+} & P.4^{-}	CAA TCA TTG CAI CGG AAC AG CA ATC TAC TGC CTG CGT AA	Homolog ydhV*	917
Clone E.2	M.80.2^{+} & P.4^{-}	ATH AAY YTN ACN CAN GGN CC CA ATC TAC TGC CTG CGT AA	*P.vulgaris* (N-terminus) ydhV	768
Clone B	C.1^{+} & P.4^{-}	CAA TCA TTG CAI CGG AAC AG CA ATC TAC TGC CTG CGT AA	Homolog ydhV	1029
Clone N	M.80.2^{+} & P.4^{-}	ATH AAY YTN ACN CAN GGN CC CA ATC TAC TGC CTG CGT AA	*P.vulgaris* (N-terminus) ydhV	783

* ydhV: Fished motif's gene from *E. coli* k-12 wild type

Table 5.7: Fished fragments from genomic DNA partially restricted with *Eco* R I, *Bam* HI and *Not* I enzymes.

The restriction of DNA with different restriction enzymes aimed at generating fragments of different sizes and nucleotide sequences. *Bam* HI and *Eco* R I recognize a four-nucleotide sequence while *Not* I recognizes an eight-nucleotide sequence.

Assuming that the four component nucleotides (A, C, T, and G) are distributed randomly within a DNA molecule, then any four nucleotides will occur, on average, every (4^4) 256 nucleotides and an eight-nucleotide recognition site will occur every (4^8) 65536 nucleotides.

In other words, the presence of an intact gene in any of the *Not* I generated fragments, enables the fishing of it through only one PCR reaction.

The mapping analysis of the above fished fragments revealed that the fished sequences were of two types. The first one is the open frame sequences, while the second type has some nucleotide sequences representing the C-terminal of unknown proteins in the genome of the *P. vulgaris*.

A search in the genome databases with the amino acid sequences of each of these sequences showed that, these fished sequences have homologous sequences which could refer to the function of some of these fragments (Table 5.8).

Serial No.	Clone name	Length (bp)	Homologous protein	Identity
1	Clone 2	338	polydeoxyribonucleotide synthase [NAD+]	24 %
2	Clone 6	934	pyruvate dehydrogenase E1 (lipoamide) beta subunit pdhB	45 %
3	Clone I	1299	- prephenate dehydratase PheA homolog lmo1536 - shikimate kinase (EC 2.7.1.71)	36 % 35 %
4	Clone II	798	- transketolase (EC 2.2.1.1) - pyruvate dehydrogenase E1 (lipoamide) beta subunit - acetoin dehydrogenase (TPP-dependent) (EC 1.-.-.-)	66.9 % 33.3 % 33.7 %
5	Clone III	980	- hypothetical protein ECs 3990 - 3-hydroxy-3-methylglutaryl-coenzyme a reductase (EC 1.1.1.34)	71 % 34 %
6	Clone VIII	407	- cytochrome p450 71d9 (EC 1.14.-.-) - cytochrome p450 monooxygenase F3F19.10	53 % 41 %
7	Clone IX	294	probable ABC transporter, ATP binding component	32 %
8	Clone S.1	888	carboxypeptidase B (EC 3. 4. 17. 2) CPB2 precursor-human	90 %
9	Clone S.2	1235	integral membrane proteinase (EC3.4.-.-)	81 %
10	Clone S.3	686	phosphatidylcholine-sterol acyltransferase (EC 2.3.1.43)	35 %
11	Clone S.4	1435	t-complexprotein 1, theta subunit	38 %
12	Clone E	917	heme-binding protein A precursor	28 %
13	Clone B	1029	hABC transport protein homolog	38 %

Table 5.8: List of some fished fragments and their corresponding homologous protein.

5.7 Gene sequences of some homologous proteins

To explore the relationship between some of the fished fragments and the already isolated and characterized proteins, the selected fragments were aligned to their corresponding homologous with the help of Clustal program, under HUSAR package.

For the alignment, the following clones were selected on the bases of the similarity to the proteins which catalyze electron transfer reactions (Table 5.9):

Clone name	Length (bp)	Homologous protein	Identity
Clone II	798	- transketolase (EC 2.2.1.1) - pyruvate dehydrogenase E1 (lipoamide) beta subunit - acetoin dehydrogenase (TPP-dependent) (EC 1.-.-.-)	66.9 % 33.3 % 33.7 %
Clone III	980	- hypothetical protein ECs 3990 - 3-hydroxy-3-methylglutaryl-coenzyme a reductase (EC 1.1.1.34)	71 % 34 %
Clone VIII	407	- cytochrome p450 71d9 (ec 1.14.-.-) - cytochrome p450 monooxygenase F3F19.10	53 % 41 %
Clone 6	934	pyruvate dehydrogenase E1 (lipoamide) beta subunit pdhB	45 %
Clone E	917	heme-binding protein A precursor	28 %
Clone B	1029	hABC transport protein homolog	38 %

For sequence details see the Appendix

Table 5.9: List of the selected clones and their homologs.

These fished fragments were aligned to the following sequences:

acetoin dehydrogenase, (Kruger, *et al*. 1994; AC I40791),

pyruvate dehydrogenase, (Takami, *et al*. 2000; AC F83981),

AOR I, (Heilig, 1999; AC Q9V2P2),

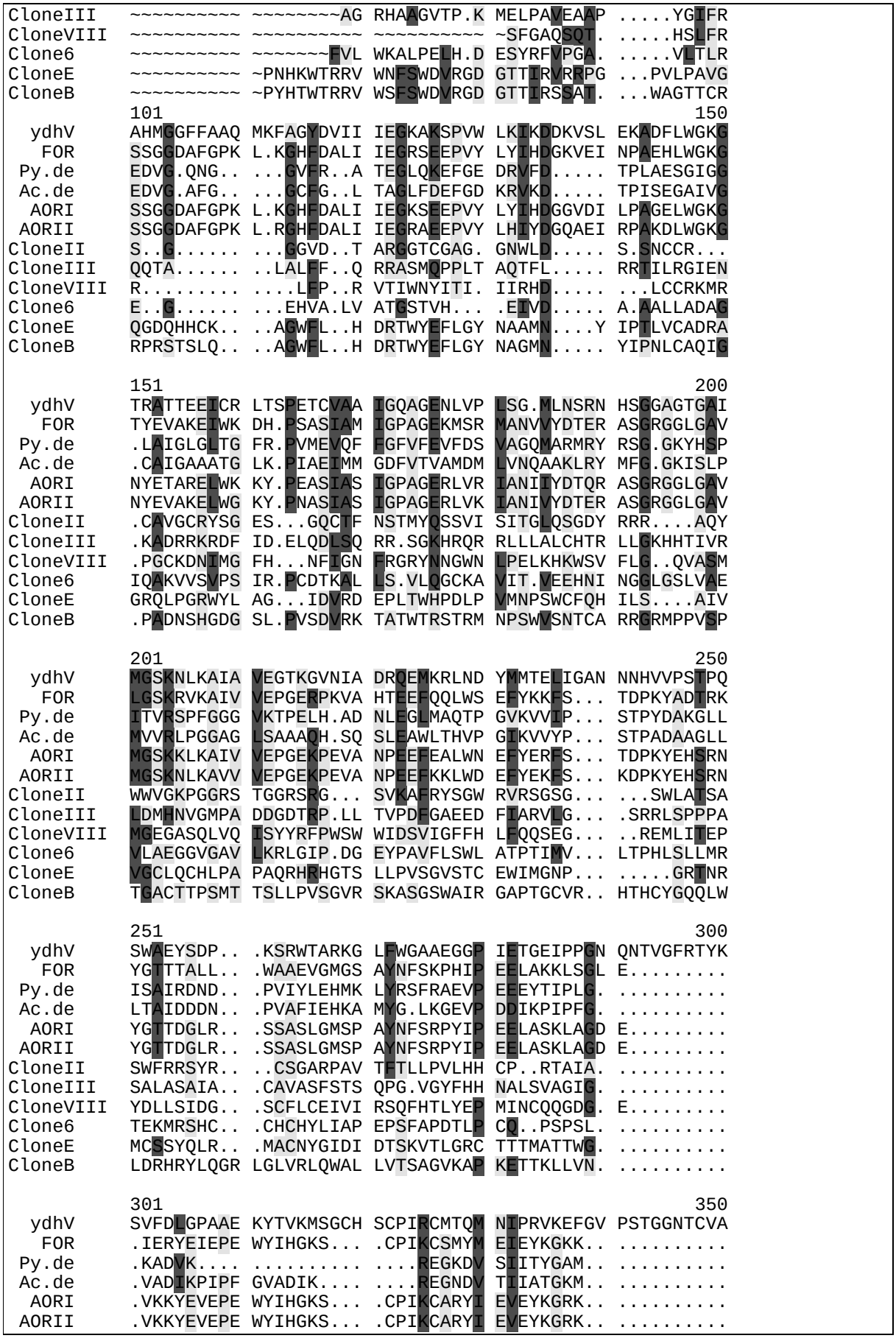
AOR II, (Kawarabayasi, *et al*. 1998; AC O57750),

FOR, (Roy, *et al*. 1999; AC O93736) and

hypothetical protein (ydhV gene), (Blattner *et al*., 1997; AC 74743.1)

The complete sequences of the acetoin dehydrogenase, pyruvate dehydrogenase, AOR, FOR and the ydhV gene, were obtained from genome database.

```
          51                                                        100
  ydhV    DEVPPGTKPF DEANKLVFAT GPLTGSGAPC SSRVNITSLS TFTKGNLVVD
   FOR    KEVPPGTDPL SPANKFVFAT GGLTGLVPGG SKVIAVSKSP ...TTRLITD
 Py.de    ~~~~~~~~~~ ~~~~~~~MAQ MTMIQAIT.D AMRNELKRD. ...ENVLVFG
 Ac.de    ~~~~~~~~~~ ~~~~~~~MKT MTYMEALR.E AMRIKMKED. ...EKVLILG
  AORI    KEVPPGTDPL SPGNKIVFAP GGLTGLIPGS SKVITVSKSP ...ETRLITD
 AORII    REVPPGTDPL SPGNKLLFVP GALTGLIPGS SKVIAVSKSP ...ETMLISD
CloneII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~IRLM ESTAGTSRE. ......LPFY
```


CloneIII ~~~~~~~~~~ ~~~~~~~~AG RHAAGVTP.K MELPAVEAAP .....YGIFR
CloneVIII ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~SFGAQSQT. .....HSLFR
Clone6 ~~~~~~~~~~ ~~~~~~~FVL WKALPELH.D ESYRFVPGA. .....VLTLR
CloneE ~~~~~~~~~~ ~PNHKWTRRV WNFSWDVRGD GTTIRVRRPG ...PVLPAVG
CloneB ~~~~~~~~~~ ~PYHTWTRRV WSFSWDVRGD GTTIRSSAT. ...WAGTTCR
101 150
ydhV AHMGGFFAAQ MKFAGYDVII IEGKAKSPVW LKIKDDKVSL EKADFLWGKG
FOR SSGGDAFGPK L.KGHFDALI IEGRSEEPVY LYIHDGKVEI NPAEHLWGKG
Py.de EDVG.QNG.. ...GVFR..A TEGLQKEFGE DRVFD..... TPLAESGIGG
Ac.de EDVG.AFG.. ...GCFG..L TAGLFDEFGD KRVKD..... TPISEGAIVG
AORI SSGGDAFGPK L.KGHFDALI IEGKSEEPVY LYIHDGGVDI LPAGELWGKG
AORII SSGGDAFGPK L.RGHFDALI IEGRAEEPVY LHIYDGQAEI RPAKDLWGKG
CloneII S..G..... ...GGVD..T ARGGTCGAG. GNWLD..... S.SNCCR...
CloneIII QQTA..... ..LALFF..Q RRASMQPPLT AQTFL..... RRTILRGIEN
CloneVIII R..... ....LFP..R VTIWNYITI. IIRHD..... ...LCCRKMR
Clone6 E..G..... ...EHVA.LV ATGSTVH... .EIVD..... A.AALLADAG
CloneE QGDQHHCK.. ..AGWFL..H DRTWYEFLGY NAAMN....Y IPTLVCADRA
CloneB RPRSTSLQ.. ..AGWFL..H DRTWYEFLGY NAGMN..... YIPNLCAQIG
151 200
ydhV TRATTEEICR LTSPETCVAA IGQAGENLVP LSG.MLNSRN HSGGAGTGAI
FOR TYEVAKEIWK DH.PSASIAM IGPAGEKMSR MANVVYDTER ASGRGGLGAV
Py.de .LAIGLGLTG FR.PVMEVQF FGFVFEVFDS VAGQMARMRY RSG.GKYHSP
Ac.de .CAIGAAATG LK.PIAEIMM GDFVTVAMDM LVNQAAKLRY MFG.GKISLP
AORI NYETARELWK KY.PEASIAS IGPAGERLVR IANIIYDTQR ASGRGGLGAV
AORII NYEVAKELWG KY.PNASIAS IGPAGERLVK IANIVYDTER ASGRGGLGAV
CloneII .CAVGCRYSG ES...GQCTF NSTMYQSSVI SITGLQSGDY RRR....AQY
CloneIII .KADRRKRDF ID.ELQDLSQ RR.SGKHRQR RLLLALCHTR LLGKHHTIVR
CloneVIII .PGCKDNIMG FH...NFIGN FRGRYNNGWN LPELKHKWSV FLG..QVASM
Clone6 IQAKVVSVPS IR.PCDTKAL LS.VLQGCKA VIT.VEEHNI NGGLGSLVAE
CloneE GRQLPGRWYL AG...IDVRD EPLTWHPDLP VMNPSWCFQH ILS....AIV
CloneB .PADNSHGDG SL.PVSDVRK TATWTRSTRM NPSWVSNTCA RRGRMPPVSP
201 250
ydhV MGSKNLKAIA VEGTKGVNIA DRQEMKRLND YMMTELIGAN NNHVVPSTPQ
FOR LGSKRVKAIV VEPGERPKVA HTEEFQQLWS EFYKKFS... TDPKYADTRK
Py.de ITVRSPFGGG VKTPELH.AD NLEGLMAQTP GVKVVIP... STPYDAKGLL
Ac.de MVVRLPGGAG LSAAAQH.SQ SLEAWLTHVP GIKVVYP... STPADAAGLL
AORI MGSKKLKAIV VEPGEKPEVA NPEEFEALWN EFYERFS... TDPKYEHSRN
AORII MGSKNLKAVV VEPGEKPEVA NPEEFKKLWD EFYEKFS... KDPKYEHSRN
CloneII WWVGKPGGRS TGGRSRG... SVKAFRYSGW RVRSGSG... ...SWLATSA
CloneIII LDMHNVGMPA DDGDTRP.LL TVPDFGAEED FIARVLG... .SRRLSPPPA
CloneVIII MGEGASQLVQ ISYYRFPWSW WIDSVIGFFH LFQQSEG... ..REMLITEP
Clone6 VLAEGGVGAV LKRLGIP.DG EYPAVFLSWL ATPTIMV... LTPHLSLLMR
CloneE VGCLQCHLPA PAQRHRHGTS LLPVSGVSTC EWIMGNP... .....GRTNR
CloneB TGACTTPSMT TSLLPVSGVR SKASGSWAIR GAPTGCVR.. HTHCYGQQLW
251 300
ydhV SWAEYSDP.. .KSRWTARKG LFWGAAEGGP IETGEIPPGN QNTVGFRTYK
FOR YGTTTALL.. .WAAEVGMGS AYNFSKPHIP EELAKKLSGL E.........
Py.de ISAIRDND.. .PVIYLEHMK LYRSFRAEVP EEEYTIPLG. ....
Ac.de LTAIDDDN.. .PVAFIEHKA MYG.LKGEVP DDIKPIPFG. ....
AORI YGTTDGLR.. .SSASLGMSP AYNFSRPYIP EELASKLAGD E.........
AORII YGTTDGLR.. .SSASLGMSP AYNFSRPYIP EELASKLAGD E.........
CloneII SWFRRSYR.. ..CSGARPAV TFTLLPVLHH CP..RTAIA. ....
CloneIII SALASAIA.. .CAVASFSTS QPG.VGYFHH NALSVAGIG. ....
CloneVIII YDLLSIDG.. .SCFLCEIVI RSQFHTLYEP MINCQQGDG. E.........
Clone6 TEKMRSHC.. .CHCHYLIAP EPSFAPDTLP CQ..PSPSL. ....
CloneE MCSSYQLR.. .MACNYGIDI DTSKVTLGRC TTTMATTWG. ....
CloneB LDRHRYLQGR LGLVRLQWAL LVTSAGVKAP KETTKLLVN. ....
301 350
ydhV SVFDLGPAAE KYTVKMSGCH SCPIRCMTQM NIPRVKEFGV PSTGGNTCVA
FOR .IERYEIEPE WYIHGKS... .CPIKCSMYM EIEYKGKK.. ....
Py.de .KADVK.... ....REGKDV SIITYGAM.. ....
Ac.de .VADIKPIPF GVADIK.... ....REGNDV TIIATGKM.. ....
AORI .VKKYEVEPE WYIHGKS... .CPIKCARYI EVEYKGRK.. ....
AORII .VKKYEVEPE WYIHGKS... .CPIKCARYI EVEYKGRK.. ....

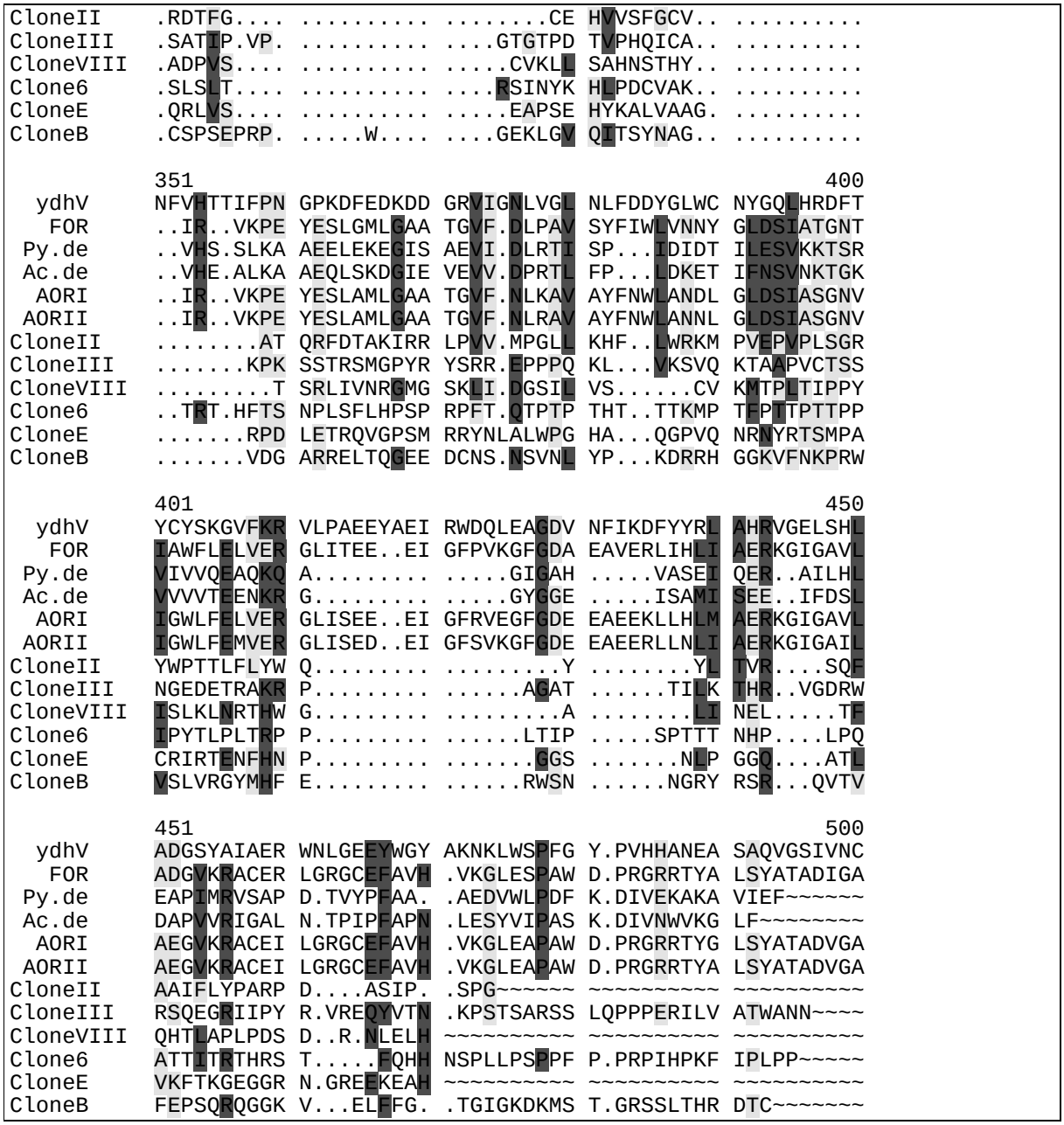


Fig 5.10: Alignment of the sequences of selected clones (cloneII, cloneIII, cloneVIII, clone6, cloneE and cloneB) and the sequences of some proteins. Abbreviations: **ydhV**, Hypothetical protein (*E. coli* k-12 wild type; Blattner *et al*., 1997; AC 74743.1); **FOR**, *Pyrococcus furiosus* (Roy, *et al*. 1999; AC O93736); **Py.de**, *Bacillus halodurans* (Takami, *et al*. 2000; AC F83981); **Ac.de**, *Clostridium magnumacetoin* (Kruger, *et al*. 1994; AC I40791); **AORI**, *Pyrococcus abyssi* (Heilig, 1999; AC Q9V2P2); **AORII**, *Pyrococcus horikoshii* (Kawarabayasi, *et al*. 1998; AC O57750). The dark grey color indicates that the amino acids are identical or they are with strong similarities. The light grey colour indicates that the amino acids have a weak similarity.

The alignment of the sequences of the clones to the complete sequences of ydhV, AOR, FOR, pyruvate dehydrogenase and acetoin dehydrogenase, revealed the presence of a similarity quite enough to relate these fished fragments to the oxidation reduction group of enzymes. Clone E (homolog to heme-binding protein-A-precursor) has one EXXH motif (at amino acid No 406 in figure 5.10), which coordinates a mononuclear metal site, most likely iron.

Another alignment of the sequences of the selected clones to AORI and AORII were also done (Figure 5.11).

```
          201                                                       250
AORI      LKAIVVEPGE KPEVANPEEF EALWNEFYER FSTDPKYEHS RNY.GTTD..
AORII     LKAVVVEPGE KPEVANPEEF KKLWDEFYEK FSKDPKYEHS RNY.GTTD..
CloneII   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~I RLMESTAG..
CloneIII  ~~~~~~~AGR HAAGVTP... .KMELPAVEA APYG....IF RQQTALALFF
CloneVIII ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Clone6    ~~~~~~~~~~ ~~~~~~~~~F .VLWKALPEL HDES.....Y RFVPGAVL..
CloneE    ~~~~~~~~~~ ~~~~~~PN.. .HKWTRRVWN FSWD.....V RGD.GTT...
CloneB    ~~~~~~~~~~ ~~~~~~P..Y .HTWTRRVWS FSWD.....V RGD.GTT...

          251                                                       300
AORI      GLRSSASLGM SPAYNFSRPY IPEELASKLA GDEVKKYEVE PEWYIHGKS.
AORII     GLRSSASLGM SPAYNFSRPY IPEELASKLA GDEVKKYEVE PEWYIHGKS.
CloneII   TSRELPFYSG GGVD...TAR GGTCGAGGNW LDSS..... ........N.
CloneIII  QRRASMQPPL TAQT...FLR RTILRGIENK ADRR...KRD ...FIDELQ.
CloneVIII ~~~SFGAQSQ THSL...FRR LFPRVTIWNY ITII..... ....IR.HD.
Clone6    TLREGEHVAL VATG...STV HEIVDAAALL ADAG..IQAK ...VVSVPSI
CloneE    .IRVRRPGPV LPAVG..QGD QHHCKAGWFL HDRT.WYEFL ..GYNAAMNY
CloneB    .IRSSATWAG TTCR...RPR STSLQAGWFL HDRT.WYEFL ..GYNAGMNY

          301                                                       350
AORI      CPIKCARYIE VEYKGRKIRV KPEYESLAML GAATGVFNLK AVAYFNWLAN
AORII     CPIKCARYIE VEYKGRKIRV KPEYESLAML GAATGVFNLR AVAYFNWLAN
CloneII   .CCRCA.... .........V GCRYSGES.. GQCT....FN STM...YQS.
CloneIII  .DLSQR.... ........RS GKHRQRRLLL ALCH.TRLLG KHH..TIVR.
CloneVIII LCCRKM.... ........RP GCKDNIMGFH NFIG...NFR GRYNNGWNL.
Clone6    RPCDTK.... ........AL LSVLQGCK.. AVIT....VE EHN.INGGLG
CloneE    IPTLVCAD.. ........RA GRQLPGRWYL AGID....VR DEP.LTWHP.
CloneB    IPNLCA.... ........QI GPADNSHG.D GSLP.VSDVR KTA..TWTR.

          351                                                       400
AORI      DLGLDSIASG NVIGWLFELV ER.GLISEEE IGFRVEGFGD EEAEE.KLLH
AORII     NLGLDSIASG NVIGWLFEMV ER.GLISEDE IGFSVKGFGD EEAEE.RLLN
CloneII   .......SVI SITGLQSGDY RR.....R.. ......AQYW WVGKP.GGRS
CloneIII  .......LDM HNVGMPADDG DT....RP.. ....LLTVPD FGAEEDFIAR
CloneVIII .........PE LKHKWSVFLG QV.....A.. ......SMMG EGASQ.LVQI
Clone6    SL....VAEV LAEGGVGAVL KRLG.IPDGE ...YPAVFLS WLATP.TIMV
CloneE    D......LPV MNPSWCFQHI LS.....A.. ......IVVG CLQCH.LPAP
CloneB    .......STR MNPSWVSNTC ARRGRMPP.. ....VSPTGA CTTPS.MTTS

          401                                                       450
AORI      LMAERKGIGA VLAEGVKRAC EILGRGCEFA VHVKGLEAPA WDPRGRRTYG
AORII     LIAERKGIGA ILAEGVKRAC EILGRGCEFA VHVKGLEAPA WDPRGRRTYA
CloneII   TGGRSRGS.. VKAFRY.... ......SGWR VR.SG..SGS WL........
CloneIII  VLGSRRLS.P PPASAL..AS ...AIACAVA SFSTS..QPG VGYFHH....
CloneVIII SYYRFPWSWW IDSVIG.... .......FFH LFQQS..EGR E.........
Clone6    LTPHLSLLMR TEKMRS.... ....HCCHCH YLIAP..EPS F.........
CloneE    AQRHRHGTSL LPVSGV.... ....STCEWI MGNPG..RTN R.........
CloneB    LLP.VSGV.R SKASGS.WAI RGAPTGCVRH THCYG..QQL WLDRHR....
```

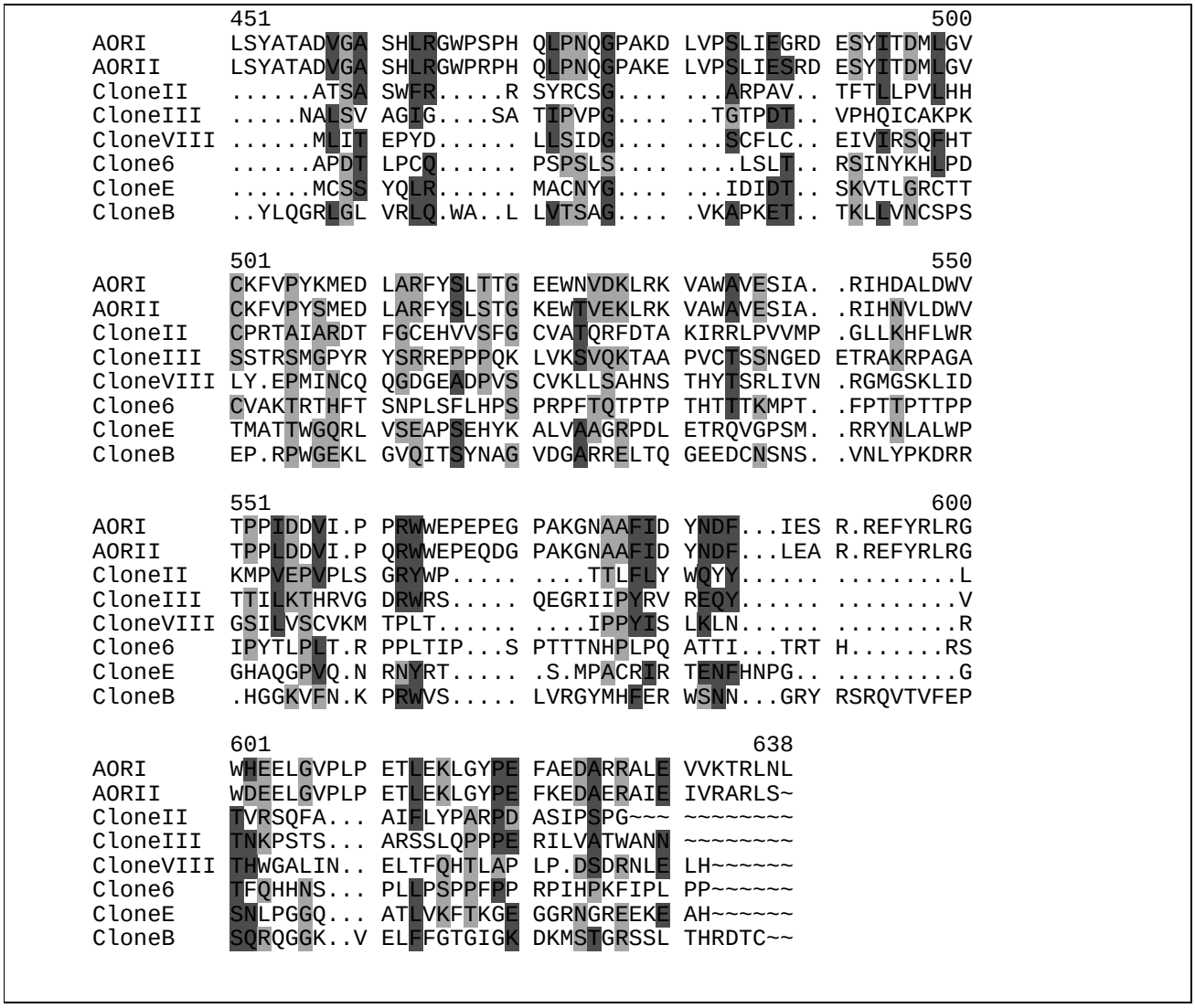


Fig. 5.11: Alignment of the sequences of selected clones (cloneII, cloneIII, cloneVIII, clone6, cloneE and cloneB) and the sequences of the 2 AORs. Abbreviations: **AORI**, *Pyrococcus abyssi* (Heilig, 1999; AC Q9V2P2); **AORII**, *Pyrococcus horikoshii* (Kawarabayasi, *et al*. 1998; AC O57750). The dark grey colour indicates that the amino acids are identical or they have a strong similarity. The light grey colour indicates that the amino acids have a weak similarity.

At position 427 in the alignment of the selected clones to the complete sequences of AORI and AORII (Figure 5.11), a Cys residue which is conserved in AORs is also conserved in all the clones except clone II. On the other hand, other Cys residue exists in both the AOR sequences, clone II and clone 6 (at position No 501). The presence of the Cys residue which is conserved in AOR and in these clones indicates the presence of a [4Fe-4S] cluster and a mononuclear tungstopterin cofactor. CloneE contains one EXXH motif (at amino acid 532 in figure 5.11).

It shoud be noted that these sequences have been placed in their respective groups based on the similarity at the amino-acid level.

A phylogenetic tree (Figure 5.12, I) has been constructed using the amino acid sequences of the fished fragments in order to confirm the results of the homology search (see Table 5.9).

This phylogenetic tree comprises the six selected sequences in addition to the following complete amino acid sequences:

ydhV, Hypothetical protein (*E. cloi* k-12 wild type; Blattner *et al*., 1997; AC 74743.1); Formaldehyde ferredoxine oxidoreductase (**FOR**) from *Pyrococcus furiosus* (Roy *et al*., 1999; AC O93736); Pyruvate dehydrogenase from *Bacillus halodurans* (Takami *et al*., 2000; AC F83981); Acetoin dehydrogenase from *Clostridium magnumacetoin* (Kruger *et al*., 1994; AC I40791); Aldehyde ferredoxine oxidoreductase (**AOR I**) from *Pyrococcus abyssi* (Heilig, 1999; AC Q9V2P2); and Aldehyde ferredoxine oxidoreductase (**AOR II**) from *Pyrococcus horikoshii* (Kawarabayasi *et al*., 1998; AC O57750).

The dendogram shows that the fished sequences fall into 2 groups. The first comprises the hydrogenases and the related fished sequences (clone III, clone VIII and clone 6), while the second group has the rest of the sequences that do not share the dehydrogenases or the AOR and FOR in the origin.

Another phylogenetic tree was constructed to test the relationship between these 6 clones and the AOR II. The six sequences and the AOR II [Isolated from *Pyrococcus horikoshii* (Kawarabayasi *et al*., 1998; AC O57750)] is depicted in Figure (5.12, II).

From the dendogram, two distinct groups are apparent, and they correspond to two different types of enzymes. Group one comprises the AOR II enzyme on one branch and clone E and clone B on the other branch. Group 2 comprises clone II and clone VIII.

The dendogram indicates that clones E and B may have the same ancestor of AOR II enzyme, that is to say there is a degree of similarity between the sequences denoted by clone E, B and the sequence of AOR II. This similarity makes it possible to relate these sequences to the aldehyde ferredoxine oxidoreductase group of enzymes.

As indicated in Figure (5.12 II), the sequences of clone III and clone 6 do not fall into any of the two groups.

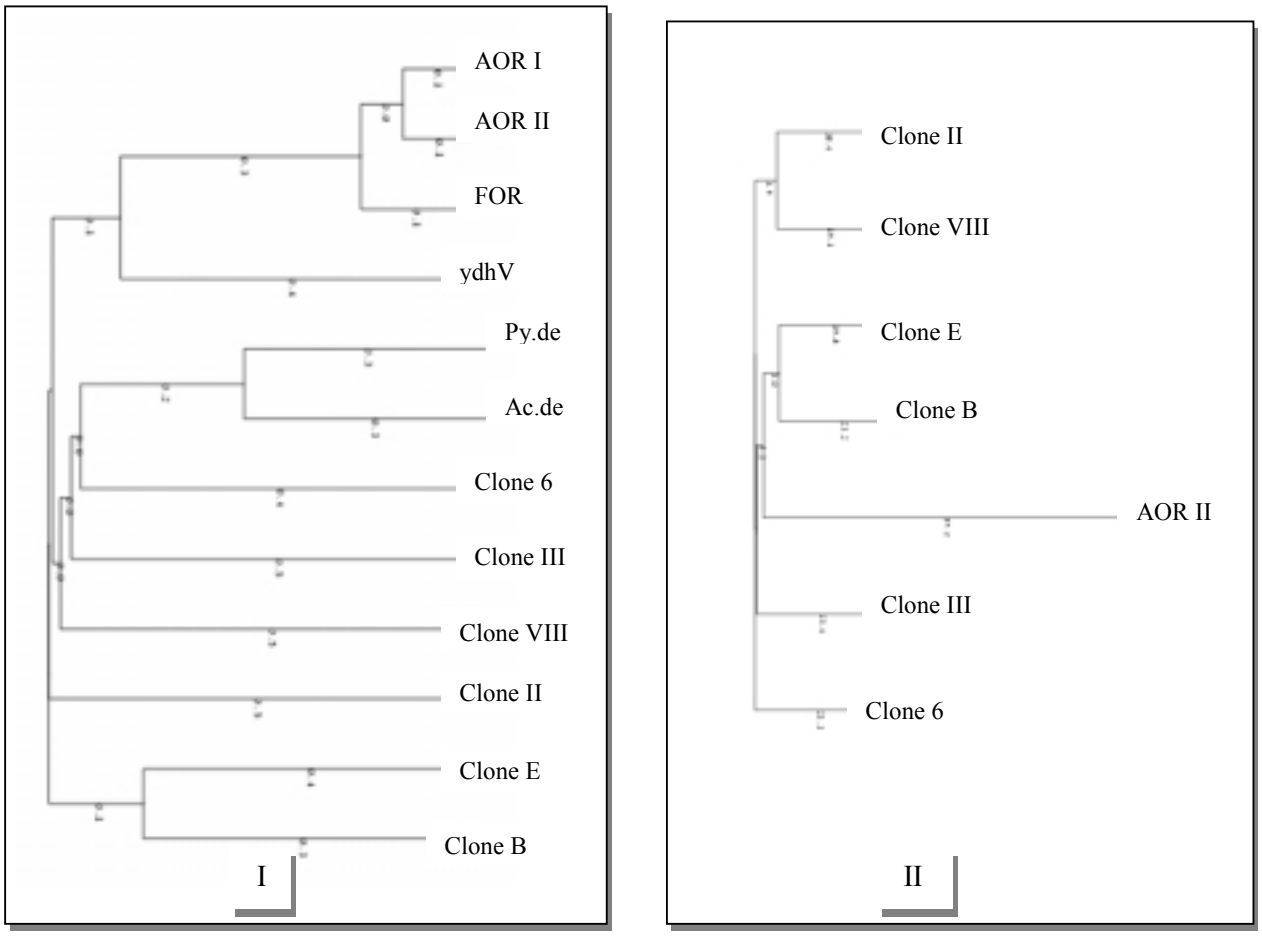


Figure: 5.12. Phylogenetic Trees

I: The relation between the amino acid sequences of the selected clones and the complete amino acid sequences of : **ydhV**, Hypothetical protein (*E. cloi* k-12 wild type; Blattner *et al*., 1997; AC 74743.1); **FOR**, *Pyrococcus furiosus* (Roy *et al*., 1999; AC O93736); **Py.de**, *Bacillus halodurans* (Takami *et al*., 2000; AC F83981); **Ac.de**, *Clostridium magnumacetoin* (Kruger *et al*., 1994; AC I40791); **AOR I**, *Pyrococcus abyssi* (Heilig, 1999; AC Q9V2P2); **AOR II**, *Pyrococcus horikoshii* (Kawarabayasi *et al*., 1998; AC O57750).

II: The relation of the amino acid sequences of the selected clones to the sequence of AOR II (*Pyrococcus horikoshii* Kawarabayasi *et al*., 1998; AC O57750).

6. Discussion

The simple way of fishing a gene is to isolate the protein that is coded by this gene. The peptide chain of the purified protein can then be sequenced totally or at least partially. The following step is the generation of the appropriate oligonucleotide probes (primers) which are to be used in the screening of the genome for fishing that target gene. On the basis of this concept, many trials have been done to purify the HVOR protein from the facultative anaerobic enterobacterium *Proteus vulgaris*. HVOR is a molybdenum-containing iron-sulphur protein. It is a membrane bound, oxidoreductase with an extremely broad substrate specificity reducing reversibly 2-oxocarboylates at the expense of reduced artificial redox mediators to (*2R*)-hydroxycarboxylates. It has been purified to a specific activity of up to 1800 μmol · min^{-1} · mg^{-1} for the reduction of phenyl pyruvate (Trautwein *et al*., 1994).

Because of the high sensitivity of the HVOR and the viologen mediators to even a trace amount of oxygen (Simon and Günther, 1998), the enzyme was losing its activity within few hours which makes it impossible to be detected through the determination of the specific activity. To avoid fast loss of activity, the enzyme should be handled in the presence of a detergent and the absence of oxygen. Efforts to precipitate HVOR still bound to membranes with ammonium sulphate were not successful. The non-ionic detergent polidocanol was always added to the buffers because it was reported by Trautwein, 1993 that this detergent plays a role in preventing the rapid loss of the activity. The substitution of the polidocanol with other detergent e.g., Triton X-100 was not successful. It was also reported that chromatography was not possible without adding polidocanol to all the buffers. Beside these advantages gained by using polidocanol, it was found that the IEF electrophoresis is greatly affected by the presence of this detergent.

The isoelectric point of the HVOR is 5.1 as determined with PhastSystem IEF media pH 3 – 9 by Trautwein *et al*., 1994. Due to the complete absence of the enzyme activity even if the enzyme preparations were preserved under nitrogen, it should be necessary to confirm the success of the purification by using another method other than enzyme assay method. Isoelectric point of the purified fractions that thought to contain the HVOR was determined. The purified protein has a molecular mass of about 70 – 75 kDa and an isoelectric point of 5.2 - 5.4. This result would be acceptable if the purified HVOR reported to have a definite molecular mass. However, different forms of the HVOR were detected (Trautwein *et al*., 1994). Monomer forms of 20, 50, 65 and 80-kDa in addition to another form of 600-kDa were also reported. By means of Superose 6 chromatography, a form of 180 - kDa was observed. Native PAGE with different gradient-gel media showed masses between 300–kDa and 1000-kDa, depending on the gradient type of the gel.

The estimated molecular mass and the isolelectric point are slightly deviated from that which were recorded by Trautwein *et al*., 1994. Such contrast does not allow the further use of the isolated fractions for subsequent blotting and sequencing.

For fishing the HVOR gene from the genome of *P. vulgaris* using other techniques, a genomic library was constructed and screened by two different methods. The first was by plaque hybridization and the second was by using the polymerase chain reaction (PCR) technique. Oligonucleotides probes for hybridization and for screening the library were designed. Another type of probes for hybridization was the use of an isolated motif's gene as a probe for hybridization at low stringency conditions. This motif's gene has N-terminal shares the N-terminal of the HVOR in 18 amino acid out of 20 (the total number of the N-terminal of the HVOR). Other probes were also used for hybridization which were fragments fished by PCR and have homologous proteins related to oxidoreductase group of enzymes. Based on the conserved amino acid sequences of the oxidoreductases available in the genome databases, an additional set of primers were designed and synthesized.

A total number of 28 different fragments was fished, cloned and sequenced. The fragments can be categorized into 3 distinct groups:

Group 1: comprises 20 open reading frame fragments

Group 2: comprises 5 C-terminus fragments

Group 3: comprises N-terminus fragment

This categorization depends on the sequence similarities of the fished fragments to the sequences found in the genome database. Open reading frame analysis of these fragments revealed that they are all fragments of genes and none of them has a complete sequence to be considered a complete gene. A search in the genome database with the amino acid sequence of each fragment revealed that the majority of them is related to the oxidation reduction group of enzymes.

The number of the fished fragments revealed also that the used probes (either for hybridization or for PCR experiments) were nonspecific to fish the HVOR gene. This is may be due to the high degeneration nature of the N-terminal sequence of HVOR which was proposed as the basis of designing a relatively large number of the used primers. On the other hand, the primers that have been designed on the basis of the homologous sequences to the N-terminal of the HVOR or the conserved amino acid sequences within the group of the oxidoreductases, led to the fishing of sequences homologs to these groups.

From these 28 fished sequences, 6 were selected for the elucidation of the function of each of them by comparing their amino acid sequences to the related amino acid sequences. These selected sequences and their homologs are showed below:

Clone name	Length (bp)	Homologous protein	Identity
Clone II	798	- transketolase (EC 2.2.1.1) - pyruvate dehydrogenase E1 (lipoamide) beta subunit - acetoin dehydrogenase (TPP-dependent) (EC 1.-.-.-)	66.9 % 33.3 % 33.7 %
Clone III	980	- Hypothetical protein ECs 3990 - 3-hydroxy-3-methylglutaryl-coenzyme a reductase (EC 1.1.1.34)	71 % 34 %
Clone VIII	407	- cytochrome p450 71d9 (ec 1.14.-.-) - cytochrome p450 monooxygenase F3F19.10	53 % 41 %
Clone 6	934	pyruvate dehydrogenase E1 (lipoamide) beta subunit pdhB	45 %
Clone E	917	heme-binding protein A precursor	28 %
Clone B	1029	ABC transport protein homolog	38 %

For sequence details see the Appendix

The alignment of these sequences (see Figure 5.10) to the complete amino acid sequences of **ydhV**, Hypothetical protein (*E. coli* k-12 wild type; Blattner *et al*., 1997; AC 74743.1); **FOR** from *Pyrococcus furiosus* (Roy *et al*., 1999; AC O93736); **Pyruvate dehydrogenase** from *Bacillus halodurans* (Takami *et al*., 2000; AC F83981); **Acetoin dehydrogenase** from *Clostridium magnumacetoin* (Kruger *et al*., 1994; AC I40791); Aldehyde ferredoxine oxidoreductase **(AORI)** from *Pyrococcus abyssi* (Heilig, 1999; AC Q9V2P2); Aldehyde ferredoxine oxidoreductase **(AORII)** from *Pyrococcus horikoshii* (Kawarabayasi *et al*., 1998; AC O57750) revealed the presence of a degree of similarity.

This similarity allows the supposition that these fished sequences are all related to the oxidoreductase group of enzymes and hence the function of these sequences could be expected and subsequently their genes in the genome of *P. vulgaris*. Clone E (homolog to heme-binding protein A precursor) has one EXXH motif (at amino acid No 406 in figure 5.10), which coordinates a mononuclear metal site, most likely iron.

Another alignment (Figure 5.11) was also constructed to explore the related sequences to the AOR group of enzymes. The alignment showed the presence of one Cys residue in all the sequences except clone II (at amino acid No. 427), while clone 6 and clone II both have the Cys residue at amino acid No 501. Cys residue (which is conserved in the AOR group of enzymes) reflects the presence of a [4Fe-4S] cluster and a mononuclear tungstopterin cofactor. Clone E contains one EXXH motif (at amino acid 532 in figure 5.11).

A phylogenetic tree (Figure 5.12 I) comprising the 6 selected sequences and the above mentioned amino acid sequences of ydhV, AOR I, AOR II, pyruvate dehydrogenase, acetoin dehydrogenase and the FOR was constructed. A phylogenetic relationship of sequences designated clone III, clone VIII, and clone 6 to the dehydrogenases can be observed. Clone 6 shares the 2 dehydrogenases the same origin, while the two other sequences (clone III and clone VIII) seem to be also related to the

dehydrogenases. It sould be noticed that the other sequences do not have a relatioship to the dehydrogenases. Another phylogenetic tree (Figure 5.12 II) was also constructed to show the relationship of the 6 sequences to the aldehyde ferredoxine oxidoreductases (AOR II). The tree shows that the sequences assigned clone E and clone B have the same origin of the AOR II enzyme and these fished sequences may reflect the presence of an AOR encoding gene in the genome of *P. vulgaris*.

It is clear that searching the genome database with the amino acid sequences of the fished fragments gave results which are relatively different from those obtained after the alignment of the fished sequences to the above mentioned amino acid sequences of the well known genes found in other organisms. The results of the phylogenetic analysis are matching the other results in only one case. This contradiction can be summarized in the following Table

Sequence name	**Homologous protein (Genome database)**	**Identity**	**Alignment results**	**Phylogenetic tree results**
Clone II	- transketolase (EC 2.2.1.1) - pyruvate dehydrogenase E1 (lipoamide) beta subunit - acetoin dehydrogenase - (TPP-dependent) (EC 1.-.-.-)	66.9 % 33.3 % 33.7 %	AOR (Cys reside at position 501)	Does not fall in any group
Clone III	- hypothetical protein EC 3990 - 3-hydroxy-3-methylglutaryl-coenzyme a reductase (EC 1.1.1.34)	71 % 34 %	AOR (Cys reside at position 501)	Dehydrogenases
Clone VIII	- cytochrome p450 71d9 (ec 1.14.-.-) - cytochrome p450 monooxygenase F3F19.10	53 % 41 %	AOR (Cys reside at position 501)	Dehydrogenases
Clone 6	- Pyruvate dehydrogenase E1 (lipoamide) beta subunit pdhB	45 %	AOR (Cys reside at positions 427 & 501 respectively)	Dehydrogenases
Clone E	- heme-binding protein A precursor	28 %	AOR (Cys reside at positions 427) + 1 EXXH motif	AOR
Clone B	- ABC transport protein homolog	38 %	AOR (Cys reside at positions 427)	AOR

From this table it could be concluded that a combination of the types of results will give rise to support the assumption that clone E is strongly related to the aldehyde ferrodoxine oxidoreductase group of enzyme since the Cys residue was detected in addition to the EXXH motif. The search of the genome database with the amino acid sequence of clone E revealed that this sequence has a homologous heme-binding protein-A-precursor. The results of the alignment and the phylogenetic relationship of this sequence support this hypothesis.

7. Summary

The aim of the present study was the screening the genome of the facultative anaerobic enterobacterium *Proteus vulgaris* for the presence of the gene encoding the (*R*)-2-hydroxycarboxylate-viologen-oxidoreductase (HVOR) protein.

HVOR is a molybdenum-containing iron-sulphur protein. It is a membrane bound, oxidoreductase with an extremely broad substrate specificity reducing reversibly 2-oxocarboylates at the expense of reduced artificial redox mediators to (*2R*)-hydroxycarboxylates. It has been purified to a specific activity of up to 1800 μmol · min^{-1} · mg^{1} for the reduction of phenyl pyruvate (Trautwein *et al*., 1994).

For screening the genome of *P. vulgaris* for the HVOR gene, a genomic library was constructed and screened by two different methods. The first was by plaque hybridization and the second was by using the polymerase chain reaction (PCR) technique. The blot hybridization technique was also used to screen the partially digested DNA.

Five different types of oligonucleotide probes for hybridization and for screening the library were designed.The first type was those probes designed based on the available information on the N-terminal of the HVOR as well as its homolgos (available in the genome databases). The second type of probes was the use of an isolated motif's gene (**ydhV,** isolated from *E. coli* k-12 wild type; Blattner *et al*., 1997; AC 74743.1)) as a probe for hybridization at low stringency conditions.

The third type of probes was the 2 unequal fragments (1908 and 190bp respectively) resulted from the restriction digestion of the ydhV gene with Bam HI. The fourth type of probes, which were also used for hybridization, was those fragments which have been fished by PCR and have homologous proteins related to oxidoreductase group of enzymes. The fifth type was the probes designed based on the conserved amino acid sequences within the group of the oxidoreductases.

A total number of 28 different fragments were fished, cloned and sequenced. The fragments were categorized into 3 distinct groups: Group 1 comprises 20 open reading frame fragments; Group 2 comprises 5 C-terminus fragments and Group 3 contains one N-terminus fragment.This categorization is based on the sequence similarities of the fished fragments to the sequences found in the genome database.

Open reading frame analysis of these fragments revealed that they are all fragments of genes. A search in the genome database with the amino acid sequence of each fragment disclosed that the majority of them are related to the oxidoreductases.

The number of the fished sequences (28 fragments) indicated also that the used probes (either for hybridization or for PCR experiments) were not specific enough to fish the target gene (HVOR).

This is may be due to the high degeneration nature of the N-terminal sequence of HVOR. The primers that have been designed on the basis of the homologous sequences to the N-terminal of the HVOR or the conserved amino acid sequences within the group of the oxidoreductases, led to the fishing of sequences homologs to this group.

From these 28 fished sequences, 6 were selected, on the basis of their similarity to the proteins which catalyze electron transfer reactions, for the elucidation of their function by comparing their amino acid sequences to those of the related amino acid sequences. One fragment assigned as clone E was found to have one EXXH motif, which coordinates a mononuclear metal site, most likely iron.

The presence of one Cys residue (which is conserved in the AOR enzymes) in all the fragments reflects the presence of a [4Fe-4S] cluster and a mononuclear tungstopterin cofactor.

Two phylogenetic trees were constructed to elucidate the phylogenetic relationship of the selected 6 sequences to some of the well known sequences of the oxidoreductases as a whole and to the AOR group in specific.

The first tree comprises the 6 selected sequences and the complete amino acid sequences of ydhV gene, aldehyde ferredoxine oxidoreductases AOR I [from *Pyrococcus abyssi* (Heilig, 1999)] and AOR II [from *Pyrococcus horikoshii* (Kawarabayasi *et al*., 1998)], pyruvate dehydrogenase, acetoin dehydrogenase and the FOR revealed the relationship of the sequences designated clone III, clone VIII, and clone 6 to the dehydrogenases.

The second phylogenetic tree indicated a relative relationship of these 6 sequences to the (AOR II). The sequences assigned clone E and clone B were found to share the same origin of the (AOR I) and these fished sequences reflect the presence of an AOR encoding gene in the genome of *P. vulgaris*.

It can be concluded that **clone E** is strongly related to the aldehyde ferrodoxine oxidoreductase group of enzyme since the Cys residue was detected in addition to the EXXH motif. The results of the alignment and the phylogenetic relationship of this sequence support this hypothesis.

8. References

- Adams, M. W. W. and Mortenson, L. E. (1985). Chapter 10: Molybdenum reductases; Nitrate reductase and formate dehydrogenase: In molybdenum enzymes. Ed. by Spiro T. G., J. Wiley & Sons. UK. 519-593.
- Anderson, G. L., Williams, J. and Hille, R. (1992). The purification and characterization of arsenite oxidase from *Alcaligenes faecalis*, a molybdenum containing hydroxylase. *J. Biol. Chem.* 267, 23674-23682.
- Baas, D. (1998). Die Isolierung der für Transhydroxylase aus *Pleobacter acidigallici* codierenden Gene sowie deren Sequenzierung, Analyse und Expression. Ph.D Thesis, Karlsruhe Uni. Germany.
- Bauder, R., Tshisuaka, B. and Lingens, F. (1990). VII. Quinoline oxidoreductase from *Pseudomonas putida* a molybdenum-containing enzyme. Biol. Chem. Hoppe-Seyler 371: 1137-1144.
- Berger, L. S. and Kimmel, A. R. (1987). Guide to Molecular Cloning Techniques: *Methods Enzymol*. 152.813 pp.
- Bilous, P. T. and Weiner, J. H. (1985). Dimethylsulfoxide reductase activity by anaerobically grown *Escherichia coli* HB101. J. Bacteriol. 162:1151-1155.
- Blattner, F. R., Plunkett, G. III, Bloch, C. A., Perna, N. T., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., MaGregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Mau, B. and Shao, Y. (1997). The complete genome sequence of *Escherichia coli* k-12. Science 277:1453-1474.
- Börner, G., Karrasch, M. and Thauer, R. K. (1991). Molybdopterin adenine dinucleiotide and molybdopterin hypoxanthine dinucleotide in formylmethanofuran dehydrogenase from Methanobacterium thermoautotrophicum (Marburg). FEBS Lett. 290:31-34.
- 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.
- Campillo-Campbell, A. D. and Campbell, A. (1982). Molybdenum cofactor requirement for biotin sulfoxide reduction in *Escherichia coli*. J. Bacteriol. 149:469-478.
- Clewell, D. B. (1972). Nature of Col E: plasmid replication in *Escherichia coli* in the presence of chloramphenicol. J. Bacteriol. 110:667.
- Coetzee, J. N., Datta, N. and Hedges, R. W. (1972). R factors from *Proteus rettgeri*. J. Gen. Microbiol. 72:543-552.

- Cohen, S. N., Annie, Chang C. Y., and Leslie, Hsu. (1972). Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. 69:2110.
- Coughlan, M. P. (1980). Chapter 4: Aldehyde oxidase, xanthin oxidase, xanthin dehydrogenase: hydroxylases containing molybdenum, iron, sulfur and flavin. 119-185.
- Courtright, J. B. (1976). Drosophila gene-enzyme systems. Adv. Genet. 18:249-314.
- Dale, W. J. (1994). Molecular genetics of bacteria. J. Wiley & Sons. UK. 287 pp.
- Davis, B. J. (1964). Disc electrophoresis-II: Method and application to human serum proteins. Annale. N. Y. Acad. Sci. 121:404-427.
- DiJoseph, C. G., Bayer, M. E. and Kaji, A. (1973). Host cell growth in the presence of the thermosensitive drug resistance factor, Rst1. J. Bacteriol. 115:399-410.
- DiJoseph, C. G. and Kaji, A. (1974). The thermosensitive lesion in the replication of the drug resistance factor Rts1. Proc. Natl. Acad. Sci. USA 71:2515-2519.
- Dilworth, G. L. (1983). Occurrence of molybdenum in the nicotinic acid hydroxylase from *Clostridium barkeri*. Arch. Biochem. Biophys. 221:565-569.
- Edward I. S., Dimitri, C. and William, E. N. (1993). Molybdenum enzymes, cofactors, and model systems. ACS symposium Series: 535. Am.Chem.Soc.pp 387.
- Gary, D., F. and David T. (1996). Plant gene isolation principles and practice. J. Wiley & Sons. UK. pp 426.
- Gersten, D. M. (1996). Gel electrophoresis: Proteins: Essential techniques techniques series. J. Wiley & Sons. UK. 177 pp.
- Gronenborn, B. and Messing, J. (1978). Nature. 272, 375.
- Heider, J., Ma, K. and Adams, M. W. W. (1995). Purification, chracterization and metabolic function of aldehyde ferredoxin oxidoreductase from the hyperthermophilic and proteolytic archaeon *Thermococcus* strain ES-1. J. Bacteriol. 177:4757-4764.
- Heilig R. "*Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure and evolution."; Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases
- Hille, R. (1996). The mononuclear molybdenum enzymes. Chem.Rev. 96:2757-2816
- Ishaq, M. and Kaji, A. (1980). Mechanism of T4 phage restriction by plasmid Rts1. Cleavage of T4 phage DNA by Rts1-specific enzyme. J. Biol. Chem. 255:4040-4047.
- Ishihara, M., Kamio, Y. and Y. Terawaki. (1978). Cupric ion resistance as a new genetic marker of a temperature-sensitive R plasmid, Rts1 in *Escherichia coli*. Biochem. Biophys. Res. Commun. 82:74-80.

- Janosi, L., Yonemitsu, H., Hong, H. and Kaji A. (1994). Molecular cloning and expression of a novel hydroxymethylcytosine-specific restriction enzyme (Ovu Rts1 I) modulated by glucosylation of DNA. J. Mol. Biol. 242:45-61.
- Jones, P. (1998). Vectors: Cloning Applications: Essential techniques series. J. Wiley & Sons. UK. 158 pp.
- Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino Y., Yamamoto, S., Sekine, M., Baba, S. I., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K. I., Yoshizawa, T., Nakamura, Y., Robb, F.T., Horikoshi, K., Masuchi, Y., Shizuya, H. and Kikuchi, H. (1998). "Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. DNA Res. 5:55-76.
- Kruger, R. N., Oppermann, F. B., Lorenzl, H. and Steinbuchel, A. (1994). Biochemical and molecular characterization of the *Clostridium magnumacetoin* dehydrogenase enzyme system. J. Bacteriol. 176, 3614-3630
- Kunst, A., Draeger, B. and Ziegendorn, J. (1984). D-glucose. In: Methods of Enzymatic Analysis. 3rd Ed. Bergmayer, H. U., Weinheim, Verlag Chemie. 6:163-172.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.
- Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978). The isolation of structural genes from libraries of eucaryotic DNA. *Cell* 15:687.
- Mehra, R. K. and Coughlan, M. P. (1984). Purification and properties of purine hydroxylase II from *Aspergillus nidulans*. Arch. Biochem. Biophys. 229:585-595.
- Meyer, O. and Rajagopalan, K. V. (1984). Molybdopterin in carbon monoxide oxidase from carboxydotrophic bacteria. J. Bacteriol. 157: 643-648.
- Mukund, S. and Adams, M. W. W. (1991). The novel tungsten-iron-sulfur protein of the hyperthermophilic archaebacterium, *Pyrococcus furiosus*, is an aldehyde ferredoxin oxidoreductase: evidence for its participation in a unique glycolytic pathway. J. Biol. Chem. 266:14208-14216.
- Murata, T., Ohnishi, M., Ara, T., Kaneko, J., Han, C., Fang, Y, Li, Takashima, K., Nojima, H., Nakayama, K., Kaji, A., Kamio, Y., Miki, T., Mori, H., Ohtsubo, E., Terawaki, Y. and Hayashi, T. (2002). Complete nucleotide sequence of plasmid Rts1: Implications for Evolution of large plasmid genomes. J. Bacteriol. 184:3194-3202.
- Nelson, L. D. and Cox, M. M. (2000). Principles of biochemistry. 3rd ed. Worth Publishers. New York. 1152pp.

- Neuhoff, V., Arold, N., Taube, D. and Erhardt, W. (1988). *Electrophoresis*. 9:255-262.
- Neuhoff, V., Stamm, R. and Eibl, H. (1985). Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: A systematic analysis. *Electrophoresis*. 6:427-448.
- Neumann, S. (1985). Über eine nicht Pyridinnukleotid-abhängige 2-Oxosäure-Reduktase aus zwei *Proteus* species. Dissertation TU München.
- Novick, R. P., Clowes, R. C., Cohen, S. N., Curtiss, R. III, Datta, N., and Falkow, S. (1976). Uniform nomenclature for bacterial plasmids: a proposal. Bacteriol. Rev. 40:168.
- Oltmann, L. F., Claasen, V. P., Kastelein, P., Reijnders, W. N. M. and Stouthamer, A. H. (1979). Influence of tungstate on the formation and activities of four reductases in *Proteus mirabilis*. FEBS Letters. 106:43-46.
- Oltmann, L. F. and Stouthamer, A. H. (1975). Reduction of tetrathionate, trithionate and thiosulphate, and oxidation of sulphide in *Proteus mirabilis*. Arch. Microbiol. 105:135-142.
- Ornstein, L. (1964). Disc electrophoresis-I: Background and theory. Annals. N. Y. Acad. Sci. 121:321-349.
- Peferoen, M., Huybrchts, R. and DeLoff, A. (1979). FEBS Lettars. 145:369-372.
- Rajagopalan, K. V. (1980). Chapter 7: Sulfite oxidase: In Molybdenum and molybdenum-containing enzymes. Ed. by M. P. Coughlan, Pergamon Press Oxford-Frankfurt. 241-271.
- Read, S. M. and Northcote, D. H. (1981). Minimization of variation in the response to different proteins of the Coomassie blue G dye-binding assay for proteins. Anal. Biochem. 116:53-64.
- Redei, P. G. (1998). Genetics Manual: current theory, concepts, terms. Word scientific.New York.
- Renart, J., Reiser, J. and Stark, G. R. (1979). Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. Proc. Natl Acad. Sci. USA, 76:3116-3120.
- Robertson, M. S. (1989). Isolation and characterization of SCR_related sequences from the platyfish *Xiphophorus maculates* (Poeciliidae; Teleostei) and an evolutionary analysis of the SRC gene-family. PhD Thesis. 143 pp.
- Roche Molecular Biochemicals: PCR Application Manual 2nd ed . (1999). Germany
- Roe, B. A., Crabtree, J. and Kahn, A. (1996). DNA isolation and sequencing. Essential techniques series, J. Wiley & Sons. UK. 164 pp.
- Roy, R., Mukund, S., Schut, J. G., Dunn, M. D., Weiss, R. and Adams, W. W. (1999). Purification and molecular characterization of the tungsten-containing formaldehyde ferredoxin oxidoreductase from the hyperthermophilic archaeon *Pyrococcus furiosus*: the third of a putative five-member tungstoenzyme family. J. Bacteriol. 181:1171-1180

- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd Edition, Cold Springer Harbor Laboratory Press. USA.
- Schinschel, C. and Simon, H. (1993). Effect of carbon sources and electron acceptors in the growth medium of *Proteus* spp. On the formation of (*R*)-2-hydroxycarboxylate-viologen-oxidoreductase and dimethylsulphoxide reductase. Appl. Microbiol. Biotechnol. 38:531-536.
- Schauer, N.L., and J.G. Ferry. 1986. Composition of the coenzyme F_{420}-dependent formate dehydrogenase from *Methanobacterium formicicum*. Journal of Bacteriology 165:405-411
- Schinschel, C. and Simon, H. (1993). Preparation of pyruvate from (R)-lactate with *Proteus* species. J. Biotechnol. 31:191-203.
- Shah, V. k. and Brill, W. J. (1977). Isolation of an iron-molybdenum cofactor from nitrogenase. Proc. Natl. Acad. Sci. 74:3249-3253. USA
- Schummer, A., Yu, H. and Simon, H. (1991). Polyfunctional (2*R*)-hydroxycarboxylic acids by reduction of 2-oxo acids with hydrogen gas or formate and resting cells of *Proteus vulgaris*. Tetrahedron. 43:9019-9034.
- Simon, H. and Günther, H. (1998). Chiral synthons by selective redox reactions catalysed by hitherto unknown enzymes present in resting microbial cells. Chemistry. 20:817-885; Elsevier Science *B. V*., Amsterdam.
- Skopan, H. (1986). Kinetische Untersuchungen zur Dehydrierung von 2- (R)-Hydroxycarboxylaten mit Carboxamidomethylviologen durch eine Oxidoreduktase aus *Proteus vulgaris*. Diplomarbeit TU München.
- Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503.
- Studier, F.W. (1973). Analysis of bacteriophage T7 early RNAs and proteins on slab gels, J. Mol. Biol. 79:237-248.
- Takami, R, Nakasone, H., Takaki, K., Maeno, Y., Sasaki, G., Masui, R., Fuji, N., Hirama, F., Nakamura, C., Ogasawara, Y., Kuhara, N. and Horikoshi, S. K. (2000). Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.Nucl. Acids Res. 28:4317-4331
- Terawaki, Y., Kakizawa, Y., Takayasu H. and Yoshikawa, M. (1968). Temperature sensitivity of cell growth in *Escherichia coli* associated with the temperature-sensitive R (KM) factor. Nature 219:284-285.
- Terawaki, Y. and Rownd, R. (1972). Replication of the R factor Rts1 in *Proteus mirabilis*. J. Bacteriol. 109:492-498.
- Terawaki, Y., Takayasu, H. and Akiba, T. (1967). Thermosensitive replication of a kanamycin resistance factor. J. Bacteriol. 94:687-690.

- Thanos, L., Bader, J., Günther, H., Krauss, F, F. and Simon, H. (1987). Electroenzymatic and electrochemical reduction: Preparation of chiral compounds. *Methods. Enzymol*. 136:302-317.
- Trautwein, E. T. (1993). Weitere charakterisierung der molybdänhaltigen 2-(R)-Hydroxycarboxylat-Viologen-Oxidoreduktase aus *Proteus vulgaris*. Dissertation TU München.
- Trautwein, T., Krauss, F., Lottspeich, F. and Simon, H. (1994). (2*R*)-hydroxycarboxylate-viologen-oxidoreductase from *Proteus vulgaris* is a molybdenum-containing iron-sulphur protein. Eur. J. Biochem. 222:1025-1032.
- Vieira, J. and Messing, J. (1987). Production of single stranded DNA. *Methods. Ezymol*. 153: 3-11.
- Walker, M. J. (1984). Methods in molecular biology: Proteins. Humana Press. Cifton, New Jersey. 1:365pp.
- Warburg, O. and Christian, W. (1942). Isolierung und Kristallisation des Gärungsfermentes Enolase. Biochem. Z. 310:348-421.
- Waterborg, H. J. and Matthews, R. H. (1984). The lowery method for protein quantitation. In: Methods in molecular biology. Proteins: Ed. By Walker M. J. Humana Press Clifton, New Jersey. 1:365 pp.
- White, H. and Simon, H. (1992). The role of tungstate and/or molybdate in the formation of aldehyde oxidoreductase in *C. thermoaceticum* and other acetogens; immunological distances of such enzymes. Arch. Microbiol. 158:81-84.
- Williams, J. W., Rinderle, S. J., Schrier, J. A., Alvey, L. J. and Tseng, K. (1986). Arsenite oxidase: A molybdenum-containing iron-sulfur protein. Fed. Proc. (Metalloenzymes) 45:1660.
- Yamamoto, I., Okubo, N. and Ishimoto, M. (1986). Further characterization of trimethylamine N-oxide reductase from *Escherichia coli*, a molybdoprotein. J. Biochem. 99:1773-1779.
- Yokota, T., Kanamaru, Y., Mori, R., and Akiba, T. (1969). Recombination between a thermosensitive kanamycin resistance factor and a nonthermosensitive multiple-drug resistance factor. J. Bacteriol. 98:863-873.

9. Appendix

Nucleotidesequence of sequences designated clone II, clone III, clone VIII, clone 6, clone E and clone B.

Clone II

```
5'-
  1  ATTCGTCTTA TGGAAAGCAC TGCCGGAACT TCATGACGAG AGTTACCGTT

 51  TTATTCCGGG GGCGGTGTTG ACACTGCGCG AGGGGGAACA TGTGGCGCTG

101  GTGGCAACTG GCTCGACAGT TCATGAAATT GTTGATGCCG CTGCGCTGTT

151  GGCTGATGCA GGTATTCAGG CGAAAGTGGT CAGTGTACCT TCAATTCGAC

201  CATGTGATAC CAAAGCTCTG TTATCAGTAT TACAGGGCTG CAAAGCGGTG

251  ATTACCGTAG AAGAGCACAA TATTAATGGT GGGTTGGGAA GCCTGGTGGC

301  AGAAGTACTG GCTGAGGGCG GAGTCGGGGC AGTGTTAAAG CGTTTAGGTA

351  TTCCGGATGG AGAGTACGCA GCGGCAGCGG ATCGTGGCTG GCTACGTCAG

401  CATCATGGTT TTGACGCCGC AGCTATCGCT GCTCAGGCGC GAGACCTGCT

451  GTGACGTTCA CACTGTTGCC AGTGTGATTA CATCATTGCC CGAGAACCGC

501  CATTGCACGC GATACATTTG GCTGTGAGCA CGTAGTATCA TTTGGTTGTG

551  TTGCAACTCA ACGATTTGAT ACTGCGAAAA TCCGGCGTTT GCCTGTTGTG

601  ATGCCAGGGT TACTTTAAAA GCACTTTCTT TGGCGGAAAA TGCCAGTGTG

651  AGAGCCTGTT CCACTGTCAG GCCGCTATTG GCCAACCACT CTTTTTCTGT

701  ATTGGTGACA ATACTACTTA ACAGTTCGTT CGCAGTTTGC AGCGTAAATA

751  TAGTTTCTAT ATCCAGCCCG ACCGGATGCG TCGATACCGT CGCCAGGC
                                                        -3'
```

Clone III

```
5'-
  1  GCGTAAGGGC GGCACGCCGC AGGGGTGACA CCAAAAATGG AACTCCCCGC

 51  GGTGGAGGCC GCTCCCTACT AGGGAATCTT TCGCCAGCAA ACCGCCCTCG

101  CACTGTTTTT CCAGCGTCGC GCCAGTATGC AGCCCCCACT TACCGCTCAA

151  ACCTTCCTGC GATAACGCAC AATTTAGCTT CGCGGAATCG AGAATAAAGC

201  GGATCGCCGC AAACGGGACT TCATTGACGA ACTTCAGGAT CTCAGCCAGC

251  GTCGTTCTGG AAAGCACCGT CAGCGGAGAC TCTTGCTCGC CCTCTGCCAC

301  ACACGCCTGC TGGGTAAACA CCACACCATC GTGCGTCTCG ATATGCACAA

351  TGTTGGTATG CCCGCCGACG ATGGTGACAC ACGCCCACTT CTCACCGTTC
```

```
401  CAGACTTTGG CGCGTGAGAA GAGGATTTCA TCGCAAGGGT CCTGGGATCT

451  TAACGGAGAC TTTCCCCGCC GCCAGCCAGT GCTTTGGCAT CGGCAATTGC

501  CTGCGCTGTT GCGTCTTTCA GCACTTCCCA GCCCGGCGTT GGCTATTTCC

551  ACCATAACGC CCTCAGCGTC GCCGGAATCG GCAGCGCCAC CATTCCCGTG

601  CCGGGAACGG GGACACCCGA CACCGTTCCT CATCAAATTT GCGCGAAACC

651  CAAGTCTTCT ACACGATCGA TGGGACCTTA CAGATATAGC CGCCGCGAAC

701  CGCCACCGCA AAAGCTCGTG AAATCGGTTC AGAAAACCGC AGCGCCCGTC

751  TGTACGTCGT CCTGAAACGG GGAGGATGAA ACGCGAGCGA AAAGGCCTGC

801  AGGTGCAACC ACAATACTAA AAACCCATCG TGTAGGGGAC CGCTGGAGAA

851  GCCAGGAAGG GAGAATCATA CCGTACCGAG TCAGGGAGCA ATACGTGACA

901  AACAAACCCT CGACCAGTGC ACGCTCGTCT TTGCAACCGC CCCCCGAACG

951  AATCCTCGTG GCGACATGGG CCAACAACGA
                                      -3'
```

Clone VIII

```
5'-
  1  TCTTTTGGCG CTCAATCTCA GACACACAGC CTTTTTCGCA GATTGTTTCC

 51  AAGAGTCACT ATCTGGTGAT AAAATTATAT CACGATAATT ATAAGACATG

101  ATTTGTGCTG CAGGAAAATG AGGCCTGGAT GCAAAGATAA CATCATGGGC

151  TTTCATAACT TCATTGGCAA CTTCAGGGGA AGATACAATA ACGGTTGGAA

201  CTTGCCCGAG TTGTAGAAGC ATAAGTGGTC CGTATTTCTT GGACAGGTCG

251  CGTAGTCCAT GATGGGGGAG GGAGCCAGCC AGCTGGTGCA AATTTCCTAT

301  TACAGGTAAT TTCCATGGTC CTGGTGGTAG ATTTGAGATT CTGTTATTGG

351  TTTCTGATTT CATCTCTAGT TTCAGCAAAG TGAAGGCAGA GAGATGCTAA

401  TAACAGA
            -3'
```

Clone 6

```
5'-
  1  TTCGTCTTAT GGAAAGCACT GCCGGAACTT CATGACGAGA GTTACCGTTT

 51  TGTTCCGGGG GCGGTGTTGA CACTGCGCGA GGGGGAACAT GTGGCGCTGG

101  TGGCAACTGG CTCGACAGTT CATGAAATTG TTGATGCCGC TGCGCTGTTG

151  GCTGATGCAG GTATTCAGGC GAAAGTGGTC AGTGTACCTT CAATTCGACC

201  ATGTGATACC AAAGCTCTGT TATCAGTATT ACAGGGCTGC AAAGCGGTGA

251  TTACCGTAGA AGAGCACAAT ATTAATGGTG GGTTGGGAAG CCTGGTGGCA
```

```
301  GAAGTACTGG CTGAGGGCGG AGTCGGGGCA GTGTTAAAGC GTTTAGGTAT

351  TCCCGATGGA GAGTACCCAG CGGTTTTTTT ATCTTGGCTG GCTACCCCCA

401  CCATCATGGT TTTGACGCCG CATCTATCGC TGCTCATGCG CACAGAAAAG

451  ATGTGACGTT CACACTGCTG CCACTGTCAT TACCTCATTG CCCCAGAACC

501  TTCCTTCGCA CCTGATACTT TACCATGCCA GCCCTCACCA TCATTATCGC

551  TCTCCCTAAC CCGATCCATT AACTACTGAA AACACCTCCC CGATTGCGTC

601  GCCAAAACCC GCACCCATTT CACCTCAAAC CCACTCTCCT TTCTTCATCC

651  CTCACCACGC CCCTTCACAC AAACCCCAAC ACCGACACAC ACCACAACCA

701  AAATGCCCAC TTTCCCAACT ACCCCGACCA CCCCGCCAAT TCCATACACA

751  CTACCACTTA CACGCCCACC ACTCACCATC CCGTCACCCA CCACTACCAA

801  CCATCCACTT CCCCAAGCAA CTACCATCAC TCGCACCCAT CGCTCGACCT

851  TTCAACACCA CAACTCCCCT CTCCTACCCT CCCCACCCTT TCCCCCTCGC

901  CCAATCCACC CAAAATTCAT TCCCCTCCCT CCAC
                                          -3′
```

Clone E

```
5′-
  1  CCAAATCACA AATGGACGCG CCGCGTGTGG AACTTCTCGT GGGATGTGCG

 51  TGGGGATGGC ACCACCATAC GAGTTCGGCG GCCTGGGCCG GTACTACCTG

101  CCGTAGGCCA GGGTGATCAA CATCATTGCA AGGCCGGCTG GTTCCTCCAT

151  GACCGCACCT GGTATGAGTT CCTGGGCTAC AACGCAGCCA TGAACTACAT

201  TCCCACGCTG GTGTGCGCAG ATCGGGCCGG CCGACAACTC CCAGGGCGAT

251  GGTACCTTGC CGGTATTGAT GTGCGATAAG ATGAACCGCT TACATGGCAC

301  CCGGATCTAC CAGTGATGAA CCCATCCTGG TGTTTCCAAC ACATACTGAG

351  CGCGATCGTG GTCGGATGCC TCCAGTGTCA CCTACCGGCG CCTGCACAAC

401  GCCATCGATG ACATGGAACA TCACTGCTAC CGGTCAGTGG GGTGTGATCG

451  ACATGCGAGT GGATCATGGG CAATCCGGGG CGCACCAACA GGATGTGTTC

501  CTCATACCAA CTGCGGATGG CATGTAACTA TGGCTGAATC GACATCGATA

551  CCTCCAAGGT AACGCTGGGC CGTTGTACGA CTACAATGGC CACTACTTGG

601  GGACAGCGCT TGGTGAGTGA AGCCCCAAGC GAGCACTACA AAGCTCTGGT

651  AGCTGCAGGT CGACCGGACC TGGAGACAAG GCAAGTGGGG CCTTCAATGC

701  GTCGATACAA CCTGGCCTTA TGGCCGGGGT AACACGCACA AGGACCGGTG

751  CAAAATCGTA ATTACCGTAC CTCTATGCCC GCATGCCGGA TCCGCACCGA
```

```
 801  GAACTTTCAC AACCCCGGGG GGTCCAACTT ACCCGGGGGC TAACAGGCTA

 851  CCCTTGTCAA ATTTACCAAA GGGGAGGGGG GAAGGAATGG ACGAGAAGAG

 901  AAGGAAGCAT GACACGC
                        -3'
```

Clone B

```
     5'-
   1  CCATATCACA CATGGACGCG CCGCGTGTGG AGCTTCTCGT GGGATGTGCG

  51  TGGCGATGGC ACCACCATAC GGAGTTCGGC GACCTGGGCC GGTACTACCT

 101  GCCGTAGGCC AAGGTGATCA ACATCATTGC AGGCCGGCTG GTTCCTCCAT

 151  GACCGCACCT GGTATGAGTT CCTGGGCTAC AACGCTGGCA TGAACTACAT

 201  TCCCAATCTG TGTGCGCAGA TCGGGCCGGC CGACAACTCC CATGGCGATG

 251  GTAGCTTGCC GGTATCTGAT GTGCGGAAGT GAACCGCTAC ATGGACCCGG

 301  TCTACCAGGA TGAACCCATC CTGGGTTTCC AACACATGCT GAGCGCGTCG

 351  TGGTCGGATG CCGCCAGTGT CACCTACCGG CGCCTGCACA ACGCCATCGA

 401  TGACATGAAC ATCACTGCTA CCGGTCAGTG GGGTGCGATC GAAAGCGAGT

 451  GGATCATGGG CAATCCGGGG CGCACCAACA GGGTGTGTTC GCCATACCCA

 501  CTGCTATGGG CAGCAACTAT GGCTGGATCG ACATCGATAC CTCCAAGGAA

 551  GGCTGGGCCT TGTACGATTA CAATGGGCAC TACTTGTGAC ATCCGCTGGG

 601  TGAGTGAAAG CCCCAAAAGA GACTACAAAG CTGCTTGTGA ACTGCTCACC

 651  TTCCGAACCG AGGCCTTGGG GAGAGAAATA ATTGGGGGTG CAAATAACGT

 701  CATACAACGC TGGTGTTGAT GGCGCAAGGC GGGAACTAAC CCAGGGAGAA

 751  GAGGATTGCA ATTCCAATTC CGTGAACTTA TACCCCAAAG ACAGGCGCCA

 801  CGGAGGAAAA GTATTCAACA AACCACGCTG GGTGTGAAGC TTAGTCCGAG

 851  GGTATATGCA TTTTGAGCGT TGGTCCAATA ACGGAAGGTA TAGATCTAGA

 901  CAATAAGTGA CTGTGTTTGA GCCCTCGCAA CGACAAGGAT GAGGGAAAGT

 951  AGAGCTATTC TTCGGGACAG GAATAGGCAA AGACAAGATG TCGACGGGGC

1001  GCTCGTCGTT AACGCATAGA GACACGTGT
                                     -3'
```