Physical Characterization of *katG* Encoding Catalase HPI of *Escherichia coli*

by

Barbara Lynne Triggs-Raine

A thesis
submitted to the Faculty of Graduate Studies
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Department of Microbiology
University of Manitoba
Winnipeg, Manitoba
1987

ISBN 0-315-86132-0

Canada

PHYSICAL CHARACTERIZATION OF katG ENCODING CATALASE HP1 OF ESCHERICHIA COLI

BY

BARBARA LYNNE TRIGGS-RAINE

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

TO DARREN

AKNOWLEDGEMENTS

Throughout my research program many people supported me with their guidance, friendship, and technical assistance. I am especially grateful to my supervisor, Dr. Peter Loewen who provided all of these things as well as the encouragement I needed to get through some major obstacles during my research.

I am also grateful to Dr. Harry Duckworth and to Dr. Lynda Donald for many helpful discussions, the gifts of HB101/pBR322 and M13mp18 DNA, and an invaluable crash course in DNA sequencing. Furthermore, I am indebted to Bill Taylor and Dr. Duckworth for doing the amino acid sequence, and to Jack Switala and Dr. Loewen for isolating HPI and the cyanogen bromide fragments for amino acid sequencing. Discussions and moral support provided by Dr. Glen Klassen and by Dr. Charles Chow were appreciated. Thank you also to Tom Hobman for the gift of pAT153 DNA.

Thank you to my many friends in the department who provided an environment that made coming to work a pleasure, in particular Dave Boyd, Jack Switala, Mike Mulvey, Nancy Mohr, Pamela Sorby, Robert Hurta, Sally McNabb, Sharon Berg, and Teo DeVos. These people were true friends, constantly providing support and assistance when I needed it. I could not have asked for a better group of friends to work with. I would also like to thank some friends from outside of the university who patiently put up with the excuse, "I have to work in the lab", Cindy Nordick, the late Arlene Nightingale, Tulip and Henry Schnerch, Heather and Gerald Baete, and Katherine Newton.

To my family, my parents, and my sisters, Penny and Karen, I thank you for understanding when I had to say I was too busy because I had to study or I had to work in the lab and for providing me with unending moral support.

Finally, to Darren, for his patience during the long hours at the lab, for rides back and forth to the lab at all hours, for giving me support in times of stress, for helping me with my thesis, and for encouraging me to strive to reach my career goals, I am eternally grateful.

ABSTRACT

1.0 ABSTRACT

Three loci affecting catalase activity in *E. coli* were identified by transposon Tn*10* mutagenesis, then mapped by interrupted mating and P1 transduction. These loci were the same three that had been previously identified by nitrosoguanidine mutagenesis: *katE* at 37.8 min, *katF* at 59 min, and *katG* at 89.2 min.

The Clarke and Carbon plasmid pLC36-19 had been isolated and shown to encode the structural gene for HPI catalase. In an attempt to localize the HPI gene within the large 19.2 kb insert of pLC36-19, transposon Tn*5* insertions were made. Although none of the transposon insertions disrupted catalase activity, deletion mutagenesis of one transposon carrying plasmid, pC1, localized the gene on a 3.8 kb *Hin*dIII fragment. The *Hin*dIII fragment was subcloned into pAT153 generating two catalase positive plasmids, pBT22 and pBT54, both containing the same insert but in opposite orientations. The ends of the gene were defined by a specific *Eco*R1 deletion at the 0 kb end and by BAL31 deletions at the 3.8 kb end. Maxicell analysis of the deletion plasmids suggested the promoter was at the 0 kb end and this was confirmed by subcloning a 320 bp *Bgl*II fragment from the 0 kb end into a promoter cloning vector. The gene on pBT22 coding for the HPI protein was shown by hybridization to be the same gene in which a transposon Tn*10* had been inserted and mapped as *katG*.

Fragments of *katG* were subcloned into M13mp18/19 and sequenced by dideoxy chain termination sequencing. Approximately 3 kb of the 3.8 kb *Hin*dIII insert was sequenced and found to contain an open reading frame

2181 bp in length. Potential Shine-Dalgarno and Pribnow box sequences were found appropriately placed upstream from the open reading frame, and a potential terminator was found downstream. The open reading frame predicted a polypeptide of 726 amino acids having a molecular weight of 80 049 daltons and an amino acid composition similar to those previously determined on the HPI protein. Several cyanogen bromide peptides of the HPI protein were partially purified and sequenced in an Edman sequenator. All of the peptide sequences were identified in the predicted amino acid sequence. The codon usage of the gene suggested it was highly expressed. No obvious homology with other catalase protein or DNA sequences was found and it was proposed that the HPI catalase was a unique hydroperoxidase unlike any others which have been previously sequenced.

The promoters of three mutant *katG* genes, two of which were suspected promoter mutants, were cloned and sequenced. The sequences were identical to the wild type gene suggesting that the two suspected promoter mutants may actually be mutants for the positive regulatory gene, *oxyR*.

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES

LIST OF ABBREVIATIONS

LIST OF ABBREVIATIONS

A	adenosine
A_{260}	absorbance at 260 nm
A_{280}	absorbance at 280 nm
A_{600}	absorbance at 600 nm
Ap ($Ap^{R/S}$)	ampicillin (ampicillin resistant/sensitive)
AT	3-amino-1,2,4-triazole
ATP	adenosine 5'-triphosphate
β_{280}	extinction coefficient of yeast enolase
BLC	bovine liver catalase
bp	base pair(s)
BSA	bovine serum albumin
BMC	Boehringer Mannheim Canada
BRL	Bethesda Research Labs
C	cytidine
C-terminus	carboxyl terminus
cAMP	cyclic adenosine 5'monophosphate
cDNA	complementary DNA
Cm ($Cm^{R/S}$)	chloramphenicol (chloramphenicol resistant/sensitive)
ColE1	colicin E1
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
ddATP	2'3'-dideoxyadenosine 5'-triphosphate
ddCTP	2'3'-dideoxycytidine 5'-triphosphate
ddGTP	2'3'-dideoxyguanosine 5'-triphosphate
ddNTP	2'3'-dideoxynucleotide 5'-triphosphate
ddTTP	2'3'-dideoxythymidine 5'-triphosphate
DEAE	diethylaminoethyl
dGTP	2'-deoxyguanosine 5'-triphosphate
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	2'-deoxynucleotide 5'-triphosphate
ds	double stranded
DTT	dithiothreitol
dTTP	2'-deoxythymidine 5'-triphosphate
EDTA	ethylenediaminetetraacetic acid
EtBr	ethidium bromide
G	guanosine
HPI	hydroperoxidase I
HPII	hydroperoxidase II
HPLC	high pressure liquid chromatography
IPTG	isopropyl β-D-thiogalactoside
Kb	Kilobase pairs
Klenow	DNA Polymerase I (Klenow fragment)
Km ($Km^{R/S}$)	kanamycin (kanamycin resistant/sensitive)
LBA	LB + ampicillin
LBC	LB + chloramphenicol
LBK	LB + kanamycin
LBS	LB + streptomycin
LBSK	LB + streptomycin + kanamycin

LBST	LB + streptomycin + tetracycline
LBT	LB + tetracycline
MOI	multiplicity of infection
MOPS	3-[N-Morpholino]propanesulfonic acid
mRNA	messenger RNA
MSH	2-mercaptoethanol
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NEN	New England Nuclear
NG	nitrosoguanidine
N-terminus	amino terminus
NUV	near-UV light
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
PPO	2,5-diphenyloxazole
PVC	*Penicillium vitale* catalase
RNA	ribonucleic acid
RNase	ribonuclease
RF	replicative form
SOD	superoxide dismutase
SDS	sodium dodecyl sulfate
ss	single stranded
SSC	Sodium citrate and NaCl solution
T	thymidine
Tc ($Tc^R/^S$)	tetracycline (tetracycline resistant/sensitive)
TCA	tricarboxylic acid cycle
TE	Tris-EDTA
TEN_{250}	TE + 250 mM NaCl
TEMED	N,N,N',N'-tetramethylethylenediamine
Tn	transposon
Tris	Tris(hydroxymethy)aminomethane
U	uracil
U	unit(s)
UV	ultraviolet
vir	virulent
wt	weight
w/v	weight per volume
Xgal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

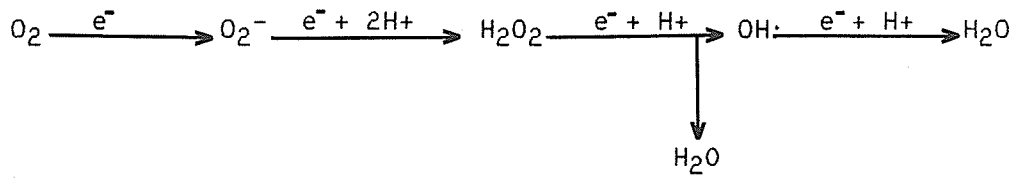
HISTORICAL

2.0 HISTORICAL

2.1 Oxygen: Both Friend and Foe

2.1.1 Introduction

Organisms benefiting from the efficiency of aerobic respiration must also deal with the consequences - toxic oxygen species. During aerobic respiration, the sequential reduction of oxygen to water leads to the production of the superoxide anion radical (O_2^-), hydrogen peroxide (H_2O_2), and the hydroxyl radical ($OH\cdot$) (Fridovich, 1978), the same intermediates thought to mediate the oxygen dependent toxicity of ionizing radiation (Moody and Hassan, 1982; Sammartano and Tuveson, 1984).



(Fridovich, 1978)

Recent evidence showing that these species may play a significant role in aging and cancer (reviewed in Halliwell and Gutteridge, 1985) has renewed interest in the study of the mechanisms of oxygen toxicity and cellular protection from toxic oxygen species. In *Escherichia coli* and *Salmonella typhimurium*, oxygen is now recognized as a serious environmental stress capable of eliciting a global protective response (Christman et al., 1985). The generation of dangerous oxygen species in biological systems and the protective mechanisms which have evolved in response to them are discussed in subsequent sections. This is followed by a closer look at

catalase, its regulation, and potential function in *Escherichia coli*. The physical characterization of catalase genes and proteins, and finally, *E. coli* gene expression, are reviewed as an introduction to the physical characterization of the *E. coli* catalase gene, *katG*.

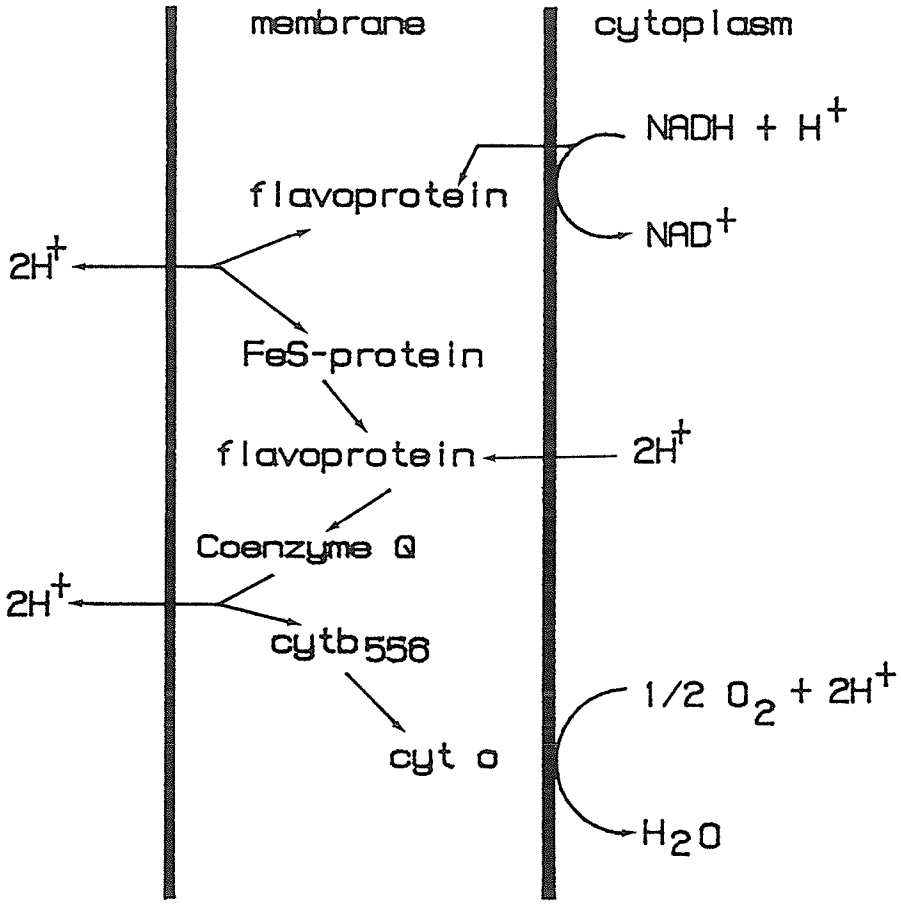
2.1.2 The Production of Toxic Oxygen Species

Many cellular components and processes are responsible for the generation of toxic oxygen species. Although aerobic respiration accounts for a significant portion of intracellular toxic oxygen species (Fridovich, 1977), other factors such as the presence of certain transition metals (reviewed by Halliwell and Gutteridge, 1985), γ-rays (Van Hemmen and Meuling, 1975), x-rays (Sammartano and Tuveson, 1984), near-UV light (Sammartano and Tuveson, 1984), radiomimetic drugs (Breimer and Lindahl, 1985), and certain enzymes (Halliwell and Gutteridge, 1985) can directly contribute to or augment the production of toxic oxygen species.

The efficiency of aerobic respiration is based on the ability of the organism to carry out oxidative phosphorylation. The generation of ATP by oxidative phosphorylation requires the passage of electrons from a reduced electron carrier, commonly NADH, through an electron transport chain to the terminal electron acceptor, oxygen. The electron transport chain typical of *E. coli* is illustrated in Figure 1. The passage of electrons through this chain is necessary to allow the passage of protons across the membrane, capturing the energy as the electrons are passed from compounds of very negative E_0''s to compounds with less negative E_0''s in the form of a proton motive force which can then be coupled to the formation of ATP according to the chemiosmotic theory (Mitchell, 1979).

The chemistry of oxygen dictates its sequential reduction.

Figure 1. The electron transport chain typical of *E. coli* grown under aerobic conditions. (Haddock and Jones, 1977)


membrane
cytoplasm
NADH + H+
flavoprotein
NAD+
2H+
FeS-protein
flavoprotein
2H+
Coenzyme Q
2H+
cytb556
1/2 O2 + 2H+
cyt o
H2O

Oxygen's unpaired electrons in the π^* orbitals are of parallel spin and thus require two electrons of antiparallel spin to reduce them (Fridovich, 1977). The Pauli exclusion principle states that electrons occupying the same orbital must be of equal and opposite spin (Dickerson et al., 1979), making it impossible for two electrons of the same spin to be passed to oxygen simultaneously. These restrictions make O_2 relatively unreactive, but the loss of single electrons during electron transport can lead to the formation of the toxic intermediates O_2^-, H_2O_2, and OH·. Cytochrome oxidase, the complex involved in the transfer of electrons to the terminal acceptor, usually oxygen, is very efficient and loss of electrons rarely occurs at this point. The most common points of loss are in the transfers leading to coenzyme Q and from coenzyme Q itself (Halliwell and Gutteridge, 1985).

Although O_2^- is a product in biological systems that directly attacks many cellular targets (Bilinski, 1985; Fridovich, 1986), it is generally accepted that the secondary products of O_2^- are the major mediators of oxygen toxicity. The production of O_2^- always leads to the production of H_2O_2, and these two species, O_2^- and H_2O_2, can then react to form the highly reactive product OH· (Fridovich, 1977). Singlet O_2 ($O_2[^1\Delta g]$) is also a dangerous oxygen species produced in biological systems by the photosensitization of certain molecules, but under normal conditions, $O_2[^1\Delta g]$ is not produced in significant quantities.

The hydroxyl radical is considered the most dangerous of the toxic oxygen species (Beauchamp and Fridovich, 1970). As stated above, OH· can result from the reaction of H_2O_2 and O_2^-. The likely mechanism involved is the Haber-Weiss reaction, a specific example of the Fenton reaction

where O_2^- acts as the reductant for the metal (Winterbourn, 1983). Other compounds such as ascorbate can also reduce Fe(III).

$$(1) \quad O_2^- + Fe(III) \longrightarrow O_2 + Fe(II)$$

$$(2) \quad Fe(II) + H_2O_2 \longrightarrow Fe(III) + H_2O + OH\cdot$$

$$(1) + (2) \quad O_2^- + H_2O_2 + H^+ \longrightarrow O_2 + H_2O + OH\cdot$$

(Halliwell, 1982)

The accessibility of the metal ions in biological systems is controversial (Fee, 1982), but it has been suggested that these metals are widely available (Halliwell et al., 1983; Fridovich, 1986).

Ionizing radiation (x-rays, γ-rays, and UV light) has also been linked to OH· mediated oxygen dependent killing (Halliwell and Gutteridge, 1985). In *E. coli*, studies on the oxygen dependent effects of ionizing radiation have focused on near-UV (300-400 nm) radiation (Tyrrell, 1985; Sammartano and Tuveson, 1984; Eisenstark and Perrot, 1987). This will be discussed in more detail in section 2.7.

2.1.3 The Targets of Toxic Oxygen Species

Reactive oxygen species attack many biological molecules, including membrane fatty acids, proteins, and DNA. Imlay and Linn (1986) have proposed that the respiration rates of higher organisms may be related to the rates of age related deterioration. The reactions of toxic oxygen species in biological systems have been reviewed (Halliwell and Gutteridge, 1985).

The extremely reactive hydroxyl radical is capable of attacking almost any biological molecule. It can abstract protons from molecules such as membrane lipids, forming H_2O, and generating a new free radical which can go on to react with other molecules. Singlet oxygen is also known to cause free radical chain reactions among membrane fatty acids

(Patterson, 1981).

Cellular proteins are another target of toxic oxygen species. The oxidation of methionine, tryptophan, histidine, and cysteine eventually leads to enzyme inactivation. Even H_2O_2 which is less reactive than OH· or $O_2[^1\Delta g]$ can attack and oxidize thiol groups of proteins, leading to enzyme inactivation.

There is a large accumulation of evidence demonstrating that DNA is a target of toxic oxygen species. Oxygen is known to cause genetic damage (Joenge et al., 1983). Paraquat, for example, which produces O_2^- is mutagenic, causing both basepair substitutions and frameshift mutations (Moody and Hassan, 1982). Plumbagin, a quinone that diverts electrons from the electron transport chain to oxygen is also mutagenic (Farr et al., 1985).

The ability of H_2O_2 to cause DNA strand scission was demonstrated by Ananthaswamy and Eisenstark (1977). Other studies have also shown that H_2O_2 in combination with O_2^- causes DNA strand scission (Farr et al., 1985), probably by the formation of OH·, a major species involved in oxygen toxicity. Ionizing radiation, radiomimetic drugs, and normal aerobic respiration all lead to the formation of OH· which can cause ss and ds DNA breaks (Breimer and Lindahl, 1985). Armel et al. (1977) found that 90% of the DNA damage due to ionizing radiation was the result of the hydroxyl radical, OH·.

2.2 Protection From Toxic Oxygen Species

The accumulation of oxygen in the earth's atmosphere during the precambrian era forced organisms to evolve defense mechanisms for protection from toxic oxygen species. The resulting defense mechanisms

were diverse, and the various levels of oxygen tolerance exhibited by bacteria have formed one of the bases of procaryotic classification. Obligate anaerobes, which cannot survive in the presence of oxygen lack protective enzymes and must sequester themselves in anaerobic environments for survival. All other organisms are aerobes containing various complements of protective enzymes which permit them to tolerate or even thrive in the presence of normal oxygen levels.

The Gram negative, facultative anaerobe, *Escherichia coli* has been the subject of a great deal of study on protection from oxidative damage. *E. coli* has two levels of defense, the predamage defense system which removes toxic oxygen species before they harm cellular components, and a postdamage defense system which repairs damage resulting from toxic oxygen species (Yonei et al., 1987). The enzymes traditionally recognized as part of the predamage defense system include superoxide dismutases, catalases, and peroxidases. These enzymes protect cells by removing H_2O_2 and O_2^- before they can form toxic secondary oxygen species. The reactions catalyzed by the protective enzymes superoxide dismutase (SOD), catalase, and peroxidase are given below.

Superoxide dismutase $O_2^- + O_2^- + 2H^+ \longrightarrow H_2O_2 + O_2$

Catalase $H_2O_2 + H_2O_2 \longrightarrow 2H_2O + O_2$

Peroxidase $H_2O_2 + RH_2 \longrightarrow 2H_2O + R$

(Fridovich, 1978)

The presence of SOD in nearly all aerobes confirms its role as a major oxidative defense. Elevated SOD levels correlated with increased resistance to O_2 and O_2- producing compounds (Fridovich, 1986). However, SOD is not essential because *E. coli sodA sodB* mutants can survive aerobically (Touati, in press) and there are aerobes that contain no SOD

such as *Lactobacillus plantarum, Neisseria gonorrhoeae*, and *Mycoplasma* (Fee, 1982). However, the absence of SOD is usually compensated for in some way such as the accumulation of Mn^{++} (Halliwell, 1982) or an elevated catalase level. The importance of SOD is further exemplified by the conservation of SOD protein sequences (Sato et al., 1987).

The ability of *E. coli* catalase mutants to grow normally at standard oxygen concentrations (Loewen, 1984), and the existence of many aerobes which lack catalase, show that catalase is also not essential to aerobes. Like SOD deficient cells, cells lacking catalase usually have a compensatory enzyme such as NADH peroxidase (Yousten et al., 1975; Padgett and King, 1986). The importance of catalase to the organism is demonstrated by the sensitivity of catalase mutants to H_2O_2 (Loewen, 1984) and the resistance to H_2O_2 exhibited by catalase overproducers (Christman et al., 1985).

There are a number of enzymes which have been implicated in postdamage repair of oxidatively damaged DNA. In *E. coli, polA* or *polC* (DNA Polymerases I and III; Hagensee and Moses, 1986), *recA* (Carlsson and Carpenter, 1980), and *xthA* (exonuclease III; Demple et al., 1983) mutants are more sensitive to H_2O_2 than wild type cells. Imlay and Linn (1986) have found that some *polA recB* strains are not even aerotolerant. These H_2O_2 and O_2 sensitivities infer a role for these genes in the repair of oxidatively damaged DNA.

2.3 Catalase: The Enzyme

2.3.1 Introduction

Catalase was one of the first bacterial enzymes to be characterized and later crystallized (Herbert and Pinsent, 1948). Catalase (H_2O_2:H_2O_2

oxidoreductase EC1.11.1.6) is responsible for the decomposition of H_2O_2 to oxygen and water. Some catalases have an associated peroxidase activity (donor:H_2O_2 oxidoreductase EC1.11.1.7) that reduces H_2O_2 to water using various organic compounds as electron donors.

The dismutation of H_2O_2 by catalase is a two stage reaction. In the first step, H_2O_2 reacts with the heme group of the enzyme to form compound I:

$$H_2O_2 + \text{Fe-E} \longrightarrow H_2O + O{=}\ \text{Fe-E} \quad \text{(Compound I)}$$

Compound I can then react with H_2O_2 in the catalatic mode or with other organic donors in the peroxidatic mode:

$$H_2O_2 + O{=}\ \text{Fe-E} \longrightarrow H_2O + O_2 + \text{Fe-E} \quad \text{(Compound II)}$$

The peroxidase reaction usually occurs when H_2O_2 is at a low concentration (Vainshtein et al., 1986). Compound I is the site of aminotriazole inhibition which will be referred to in subsequent sections.

The catalase enzymes isolated from various sources, such as animals, plants, and bacteria usually have many common features. They contain four identical subunits that form a tetramer with a molecular weight between 225 000 and 270 000. Each subunit is usually associated with one protoheme IX group and acts as an independent unit (Schonbaum and Chance, 1976). Typical catalases have a broad pH optimum and are sensitive to inhibition by 3-amino-1,2,4-triazole (AT).

Bovine liver catalase, human catalase, and rat liver catalase are nearly identical (see section 2.9) and have the structure of a typical catalase. Catalases isolated from maize (Skadsen and Scandalios, 1986) and *Drosophila melanogaster* (Nahmias and Bewley, 1984) also have typical catalase structures. A number of fungal and bacterial catalases,

including those from *Neurospora crassa* (Jacob and Orme-Johnson, 1979), *Saccharomyces cerevisiae* (Seah et al., 1973; Seah and Kaplan, 1973), *Penicillium vitale* (Vainshtein et al., 1986), *Escherichia coli* (see section 2.3.2), *Proteus mirabilis* (Jouve et al., 1984), *Rhodospirillum rubrum* (Nadler et al., 1986), *Comamonas compransoris* (Nies and Schlegel, 1982), *Rhodopseudomonas capsulata* (Nadler et al., 1986), *Lactobacillus plantarum* (Kono and Fridovich, 1983a; 1983b), *Mycobacterium avium* (Mayer and Kalkinbam III, 1986), *Thermoleophilum album* (Allgood and Perry, 1986), *Bacillus larvae* (Dingman and Stahly, 1984), *Rhodopseudomonas spheroides* (Clayton, 1959), and *Micrococcus lysodeikticus* (Herbert and Pinsent, 1948) have been partially or in some cases fully purified and characterized. Many of the bacterial and fungal catalases are atypical and some unique features of these enzymes will be discussed in section 2.3.3 with comparison to the *E. coli* catalases.

2.3.2 *E. coli* Catalases HPI and HPII

The *E.coli* catalases hydroperoxidase I (HPI) and hydroperoxidase II (HPII) have been purified and characterized (Claiborne and Fridovich, 1979; Claiborne et al.,1979; Loewen and Switala, 1986).

HPI, one of two electrophoretically distinct *E. coli* catalases, was first identified as an activity which copurified with the *E. coli* dianisidine peroxidase (Claiborne, 1978). The copurification led to the realization that HPI was a bifunctional enzyme having both catalase and broad spectrum peroxidase activity. Further physical characterization (Claiborne and Fridovich, 1979) revealed that HPI was composed of four identical 84 000 dalton subunits (78 000 by SDS-PAGE) which combined to form a 337 000 dalton tetramer containing two protoheme IX groups. Gel

filtration chromotography and SDS-PAGE by Loewen and Switala (1986) gave apparent subunit molecular weights of 81 000 and 84 000 respectively, agreeing well with previous HPI studies. On native polyacrylamide gels run in Tris-glycine buffer, HPI separated into two isoenzyme forms labelled HPI-A and HPI-B, the significance of which is presently unclear (Loewen et al., 1985a; Meir and Yagil, 1985).

Claiborne and Fridovich (1979) found HPI to have optimum catalase activity at pH 7.5, optimum peroxidase activity at pH 6.5, and a Km (at pH 7.5) of 3.9 mM H_2O_2. Some other interesting properties of HPI were its resistance to inhibition by AT, and its heat lability at 70°C (Meir and Yagil, 1985). Analysis of the HPI protein showed that the protein contained no cysteine and it was concluded that there were no disulfide linkages between subunits (Claiborne and Fridovich, 1979). Cyanogen bromide cleavage of HPI suggested that it contained 13-14 methionine residues (Claiborne et al, 1979).

HPII catalase migrated much more slowly than HPI on native polyacrylamide gels and had no associated peroxidase activity (Loewen and Triggs, 1984). The characterization of this enzyme (Claiborne et al., 1979) showed that it was composed of four 78 000 dalton subunits, having a combined molecular weight of 312 000. The spectrum of the enzyme and its reduced hemochromagen revealed that the HPII associated prosthetic group was not protoheme IX. Loewen and Switala (1986) found HPII to be a hexamer composed of six identical 90-92 000 dalton subunits and having a combined molecular weight of 532 000. Each subunit was associated with an unusual heme group which had the properties of a dihydroporphyrin or chlorin type heme, similar to those of heme d (a_2). The reasons for the differences between the HPII characterizations of these two groups are

presently unknown.

Loewen and Switala (1986) found HPII to have optimal catalase activity over a broad pH range (pH 4-11). Using an indirect approach where HPI and HPII were separately inactivated, Meir and Yagil (1985) found HPII to have two pH optima, one at pH 6.8 and one at pH 10.5. In contrast to the HPI protein, HPII was sensitive to AT, resistant to 70°C, had a Km of 18.2 mM H_2O_2, and was resistant to 0.1% SDS or 7M urea (Meir and Yagil, 1985), the latter property also having been confirmed by Loewen and Switala (1986).

Characterization of HPI and HPII showed that these enzymes were unrelated, contradicting a previous suggestion that HPII was derived from HPI (Hassan and Fridovich, 1978). HPII did not cross react with antisera prepared against HPI (Loewen et al., 1985b) nor did HPI cross react with antisera prepared against HPII (Claiborne et al., 1979). There were no similarities between the pattern of cyanogen bromide peptides produced from HPI and HPII (Claiborne et al., 1979).

2.3.3 Comparison of *E. coli* Catalases HPI and HPII to Other Catalases

The properties of the *E. coli* catalases HPI and HPII are not typical of the previously characterized catalases. However, many of the atypical features have been observed in other bacterial and fungal catalases. A discussion of these features and the catalases of other organisms exhibiting these properties follows.

HPI catalase has a tetrameric structure with a larger molecular weight than most tetrameric catalases. *Neurospora crassa* (Jacob and Orme-Johnson, 1979) and *Penicillium vitale* (Vainshtein et al., 1986) catalases are large in size as well, and the dimeric catalase of *Comamonas compransoris* has a similar subunit molecular weight (Nies and

Shlegel, 1982). Another unusual feature of HPI is its associated broad range peroxidase activity. Most catalase associated peroxidase activity allows the oxidation of only short chain organic alcohols and acids such as ethanol and formate, while broad spectrum peroxidases can oxidize a variety of compounds including pyrogallol and dianisidine. Broad spectrum peroxidase activities have also been associated with catalases from *Thermoleophilum ablum* (Allgood and Perry, 1986), *Mycobacterium avium* type-T (Mayer and Kalkinbam, 1986), and *Rhodopseudomonas capsulata* (Nadler et al., 1986). Unlike other catalases, HPI was not sensitive to inhibition by AT, a property shared by *R. capsulata* catalase (Nadler et al., 1986), *M. avium* type-T catalase (Mayer and Kalkinbam, 1986), and *T. album* catalase (Allgood and Perry, 1986). Although most catalases have a broad pH optimum, HPI had a specific optimum of pH 6.8 like *R. capsulata* (Nadler et al., 1986) and *C. compransoris* (Nies and Schlegel, 1982) catalases. Nadler et al. (1986) recommended the formation of a new group of catalases containing *E. coli* HPI, *R. capsulata* catalase, *M. avium* type-T catalase, and *C. compransoris* catalase. It was surprising that the *Proteus mirabilis* catalase had no associated peroxidase activity but resembled the mammalian enzyme very closely (Jouve et al., 1984), since both *P. mirabilis* and *E. coli* are members of the Enterobacteriaceae family.

HPII also shows a variety of atypical properties. It is a monofunctional enzyme with a hexameric structure. Kono and Fridovich (1983b) described a pseudocatalase (non-heme) from *Lactobacillus plantarum* which was hexameric. Another notable difference between HPII and other catalases is that it contains a heme d like prosthetic group very similar to that of *Neurospora crassa* (Jacob and Orme-Johnson,

1979). In many other ways HPII conforms to the characteristics of a typical catalase similar to the type-M catalase of *M. avium* (Mayer and Kalkinbam, 1986). They are both heat stable and sensitive to inhibition by AT (Mayer and Kalkinbam, 1986; Meir and Yagil, 1985). The *Rhodospirillum rubrum* catalase (Nadler et al., 1986) is also sensitive to AT, and pH independent between pH 5.5 and 10.

Like *E. coli*, both *Mycobacterium avium* (discussed above) and *Saccharomyces* have two types of catalase. Catalase A and T in yeast have pH optima at pH 6-8 and pH 9.5, (Seah and Kaplan, 1973) similar to HPI and HPII of *E. coli* (Meir and Yagil, 1984). Catalase A (Cohen et al., 1985) and T (Hartig and Ruis, 1986) from yeast have been cloned, and do not cross hybridize with each other.

Although this is not an exhaustive comparison of the catalases characterized to date, it does illustrate the diversity of oxidation protective enzymes which have evolved in procaryotic and lower eucaryotic organisms.

2.4 The Regulation of Catalase

2.4.1 Induction of Catalase

Catalase levels in *E. coli* are affected by the composition of the growth medium, the growth phase of the culture, and other factors (Yosphe-Purer and Henis, 1976). Catalase can be induced in *E. coli* and *S. typhimurium* by H_2O_2 or compounds that generate H_2O_2 as a secondary product.

The induction of catalase by H_2O_2 was first described in *S. typhimurium* (Finn and Condon, 1975) then later demonstrated in *E. coli* (Yosphe-Purer et al., 1977; Hassan and Fridovich, 1978; Richter and

Loewen, 1981). The addition of compounds that generated H_2O_2 to cultures also resulted in the induction of catalase. Pyocyanine which produced H_2O_2 was shown to induce catalase (Hassan and Fridovich, 1980) and later ascorbic acid which reacts with oxygen in the presence of metal ions to produce dehydroascorbate and H_2O_2 (Morgan et al., 1976), was shown to induce catalase in mid-log phase *E. coli* cultures (Richter and Loewen, 1981). Richter and Loewen (1981) found that the addition of ascorbate at concentrations between 0.57 mM and 5.7 mM to aerated cultures of *E. coli* caused an eight fold increase in catalase activity. Furthermore, chloramphenicol interfered with the catalase induction, demonstrating the requirement of protein synthesis for catalase induction. It has been suggested that H_2O_2 may activate a positive regulator or inactivate a negative regulator of catalase synthesis (Loewen et al., 1985a) and recently a H_2O_2 inducible positive regulator, *oxyR* has been described (Christman et al., 1985). The *oxyR* function will be discussed in detail in section 2.4.3.

Catalase induction by cysteine, an H_2O_2 generating compound, has been demonstrated in *Pseudomonas fluorescens* although 1.5 mM H_2O_2 itself did not act as an inducer (Himelbloom and Hassan, 1986).

Catalase activity has been observed to increase as *E. coli* enters stationary phase (Loewen and Triggs, 1984; Meir and Yagil, 1984). Finn and Condon (1975) observed that the specific activity of catalase decreased during log phase growth of *Enterobacter aerogenes*, *Salmonella typhimurium*, *Serratia marcescens*, *Haemophilus influenzae*, *Rhodopseudomonas spheroides*, and *Saccharomyces cerevisiae*, then increased as the cells entered stationary phase. The increase into stationary phase coincided with a decrease in the pH of the medium. *Bacillus larvae*

had no catalase during log phase growth but activity increased as it entered stationary phase and decreased once again during sporulation (Dingman and Stahly, 1984).

In *E. coli* it was known that there were two distinct catalase activities, HPI and HPII, but the relationship of these two activities to the regulation of catalase was difficult to interpret since the activities of the two catalases could not be completely separated in vivo. Loewen and Triggs (1984) were partially successful in differentiating HPI and HPII by assaying with two different H_2O_2 concentrations, HPI having a much lower Km than HPII; there was still a great deal of "spillage" of one activity to the other. Loewen et al. (1985a) later differentiated the two activities by isolating transposon mutants deficient in either HPI or HPII activity. The HPI activity was inducible by ascorbate or H_2O_2 while HPII was responsible for the increase in catalase activity as cells entered stationary phase. HPII showed no response to ascorbate and HPI did not increase during stationary phase. HPII levels were also found to be 5-10 fold higher during mid-log phase growth on TCA cycle intermediates and there was an apparent correlation between the increased synthesis of catalase and the levels of the TCA cycle enzymes. HPII levels were not influenced by whether or not active electron transport was occurring, contradicting a previous suggestion that components of the *E. coli* respiratory chain were linked to catalase levels (Hassan and Fridovich, 1978). Meir and Yagil (1984) found that H_2O_2 induced the pH 6.8 catalase activity (corresponding to HPI) but that the pH 10.5 activity (corresponding to HPII) remained constitutive, indirectly confirming the independent induction of HPI and HPII.

2.4.2 Is Catabolite Repression Involved In Catalase Synthesis?

The presence of glucose in the medium lowers the levels of catalase. Catabolite repression has been suggested to be involved in catalase expression in yeast (Sulebele and Rege, 1968), *Bacteriodes fragilis* (Gregory et al., 1977), and *E. coli* (Hassan and Fridovich, 1978; Yoshpe-Purer et al., 1977) Richter and Loewen (1982) showed that glucose did not inhibit catalase nor did cAMP enhance basal or induced catalase levels, refuting the involvement of catabolite repression in catalase expression.

Morgan et al.(1986) confirmed that glucose affects catalase by demonstrating lower HPI mRNA levels in the presence of glucose. Although glucose may lower the catalase levels, this was not a case of classical catabolite repression as glucose still affected *cpd cya* mutants (lacking the ability to generate cAMP) of *S. typhimurium* (Eisenstark and Perrot, 1987). The role of glucose in the inhibition of catalase remains unresolved.

2.4.3 The *oxyR* Regulon

Global regulons responding to phosphate starvation, DNA damage, heat shock and aerobic-anaerobic shift have been identified in *E.coli* (Gottesman, 1984). A regulon responding to oxidative stress recently identified in *S. typhimurium* and *E. coli* will be discussed in this section.

The presence of a regulon responding to oxidative stress was implicated by the ability of *E. coli* (Demple and Halbrook, 1983) and *S.typhimurium* (Winquist et al., 1984) to adapt to H_2O_2 by a process different from the SOS or alkylation response. Christman et al. (1985)

induced resistance to 10 mM H_2O_2 by pretreating *S. typhimurium* with 60 μM H_2O_2 for 60 min. This was a chloramphenicol sensitive response, demonstrating the requirement for protein synthesis in the development of resistance. Cells pretreated with H_2O_2 were also resistant to other chemical oxidants such as N-ethylmaleimide and heating at 50°C. Two dimensional (2-D) electrophoresis of the pretreated cells revealed the induction of 30 proteins, twelve of which reached maximal synthesis 10 minutes after H_2O_2 treatment, and eighteen of which reached maximal synthesis 10-20 min from the time of H_2O_2 addition. Catalase HPI was identified as one of the twelve proteins induced early.

Christman et al. (1985) isolated an H_2O_2 resistant mutant, *oxyR1* which constitutively overexpressed nine of the twelve early proteins. It was also resistant to several chemical oxidants, certain organic hydroperoxides, and heating at 50°C. Using 2-D electrophoresis five of the overexpressed proteins were identified as HPI-A (D69), HPI-B (D71), a novel NAD(P)H alkylhydroperoxidase (F52 and C22), and the dnaK (C69) protein. HPII was not induced by H_2O_2 or overexpressed in *oxyR1* mutants. A number of the proteins induced by oxidative stress could be induced by other stresses such as naladixic acid and ethanol. The overlap of stress responses was further substantiated by the fact that 5 of the 30 oxidative stress proteins are also heat shock proteins, three of which were identified as the *dnaK* protein, *groEL* protein, and NAD(P)H alkylhydroperoxidase (Morgan et al., 1986). The *oxyR* gene (mapping at 88 min in *S. typhimurium*) product was found to be a diffusible protein which was a positive regulator of *katG* (HPI; 88 min), and *ahp* (alkyl-hydroperoxidase; 13.8 min) in *S. typhimurium* and possibly *sodA* (superoxide dismutase; 89.5 min) and *gsh* (glutathione reductase; 13.8

min) (Christman et al., 1985). The induction of HPI protein was shown to occur at the transcriptional level (Morgan et al., 1986) with *oxyR1* mutants containing 50x the normal level of HPI mRNA.

Morgan et al. (1986) found that the proteins induced by H_2O_2 are not the same proteins affected by an aerobic-anaerobic shift. None of the *oxyR* regulated proteins, including HPI were inducible by O_2. HPI-A and HPI-B were in fact found to be more abundant during anoxic growth.

Some interesting features regarding the induction of the *oxyR* regulon were observed by Christman et al. (1985). H_2O_2 induced the synthesis of adenylated nucleotides previously shown to be induced by oxidative stress, heat shock, and amino acid deficiency. The main adenylated nucleotide shown to accumulate was AppppA (Bochner et al., 1984; Lee et al., 1983). These dinucleotides have been termed alarmones (Stephens et al., 1975) and are believed to lead to the redirection of metabolism to cope with cell stress. Furthermore, the alkylhydroperoxidase overexpressed in *oxyR1* was a heat shock protein that required *oxyR* in addition to the positive heat shock regulator, *htpR* for heat shock induction in *S. typhimurium* but not in *E. coli*.

The mechanism of induction of the *oxyR* regulon is not well understood but a number of ideas have been proposed. The *oxyR* protein may be analagous to htpR of the heat shock response which acts as the sigma subunit of RNA polymerase allowing recognition of the heat shock promoters of the genes encoding 17 heat shock proteins (Gottesman, 1984). It has been suggested that abnormally folded proteins may induce the heat shock response (Goff and Goldberg, 1985). Another analogous protein is the ada protein, responsible for the induction of DNA repair by alkylating agents. The methylation of this protein allows it to activate

a number of genes (Teo et al.,1986). Morgan et al. (1986) have suggested that the direct oxidation of a regulatory protein by H_2O_2 may be the inducer of the oxidative stress response or adenylated nucleotides may serve as a primary signal.

2.5 The Function of Catalase in *E. coli*

Catalase is known to break down H_2O_2, but the ability of catalase deficient mutants of *E. coli* to grow normally (Loewen, 1984) has obscured the functional significance of catalase. However, the presence of two independently regulated catalases in *E. coli* suggests that they serve an important function. Some characteristics relating to the function of bacterial catalases are presented.

Barbado et al. (1983) isolated *E. coli* mutants sensitive to H_2O_2 and found they were deficient in catalase and/or peroxidase. *E. coli* catalase mutants are always sensitive to H_2O_2 (Loewen, 1984; Meir and Yagil. 1984). O_2^- sensitive mutants of *Bacillus coagulans* have lower catalase levels (Vassilyadi and Archibald, 1985).

Transposon Tn*10* mutants of *S. typhimurium* unable to survive in macrophage were isolated and found to be less virulent in vivo as well. Two of these isolates were *oxyR* mutants, but none of them were catalase deficient. A transposon Tn*10* insertion removing HPII activity did not affect the virulence of the strain (Fields et al., 1986).

Catalase is important in the protection of *Lactobacillus plantarum* from the toxic accumulation of H_2O_2 during stationary phase (Kono and Fridovich, 1983a). Catalase is also needed for yeast to sporulate. During sporulation, catalase levels increase, (Ota, 1986) and mutants deficient in catalase are unable to sporulate inferring a role for

catalase in the protection of enzymes required for sporulation (Ota, 1986).

2.6 Loci Affecting Catalase Synthesis

Four loci affecting catalase levels in *E. coli* have been identified and mapped. Mutants of two of the loci, *katE* and *katF*, were deficient in HPII and had very low catalase levels when tested on agar plates. Three factor cotransductional crosses mapped *katE* at 37.8 min in the order *pps-pheS-pfkB-katE-xthA* (Loewen, 1984) and *katF* at 59 min in the order *srl-mutS-katF-iap-cysH-pyrG* (Loewen and Triggs, 1984) on the *E. coli* chromosome. Mutants of the *katG* locus were deficient in HPI catalase but appeared to have relatively normal catalase levels when tested on agar plates. HPI mutants could only be identified when the HPII catalase was already absent. The *katG* locus was mapped by three factor cotransductional crosses among several markers, *pfkA-metB-katG-ppc-argH-metA* at 89.2 min on the *E. coli* chromosome (Loewen et al., 1985b). Christman et al., (1985), identified a fourth locus, *oxyR*, which is a positive regulator of *katG*, and mapped it to 88 min on the *S. typhimurium* chromosome.

Three catalase loci *katA-D* have been identified in *S. typhimurium* but of these, only *katC* was mapped by P1 transduction to near the *proAB* region at 7 min (Levine, 1977). None of these loci have been identified in *E. coli*.

2.7 Catalase and Near-UV(NUV) Light

Near-UV (NUV) light (300-400 nm) is one of many forms of ionizing radiation that exhibits oxygen dependent killing (McCormick et al.,

1976). The passage of radiation through water generates OH·, hydrogen atoms, and hydrated electrons. The hydrogen atoms and hydrated electrons produce O_2^-, the OH· form H_2O_2, and these species, O_2-, OH·, and H_2O_2, further react to form more of the secondary product, OH· (Fridovich, 1978).

Evidence has accumulated implicating H_2O_2 as an important mediator of NUV killing. The lethal photooxidation product which resulted from the NUV exposure of tryptophan was characterized and found to be H_2O_2 (McCormick et al., 1976). Sammartano et al. (1986) showed that cells could be protected from the toxic effects of NUV if catalase was incorporated in the media. Mutants of the repair enzyme exonuclease III were sensitive to both H_2O_2 and NUV suggesting that H_2O_2 was playing an important role in NUV mutagenesis (Sammartano and Tuveson, 1984). Pretreatment of *E. coli* and *S. typhimurium* with H_2O_2 (an inducer of HPI catalase) induced resistance to NUV (Sammartano et al., 1986). Far-UV light did not kill in an oxygen dependent fashion (Sammartano and Tuveson, 1984).

E. coli mutants hypersensitive to NUV were isolated (Tuveson, 1981) and later identified as *katF* (HPII) mutants (Sammartano et al., 1986). This same group examined *katE* (HPII) and *katG* (HPI) mutants and found that these mutants were not NUV sensitive, contradicting the implicated role of catalase in NUV protection. However pretreatment of *E. coli* cells with H_2O_2 (inducing all *oxyR* regulated proteins including HPI catalase) did increase NUV resistance (Sammartano et al., 1986; Tyrrell, 1985).

Kramer and Ames (1987) have confirmed the involvement of the *oxyR* regulon in NUV resistance but they have demonstrated that HPI catalase is

not a mediator of this resistance. While *katG* mutants were not NUV sensitive, *ahp* (alkylhydroperoxidase) and *gsh* (glutathione reductase) mutants were NUV sensitive. NUV was able to induce H_2O_2 sensitivity, because the absorption of NUV light by catalase, caused its inactivation. High levels of catalase actually acted as a sensitizer of cells to NUV. Eisenstark and Perrot (1987) have confirmed this using the catalase mutant UM228 and UM228 carrying a *katG*-containing multicopy plasmid (pBT22). Although the cells carrying the plasmid were slightly more resistant to H_2O_2, they had increased NUV sensitivity. *OxyR1* mutants which overproduce HPI also show increased sensitivity to NUV. The role of catalase as a sensitizer is supported by the fact that *hemA* mutants that are defective in the synthesis of δ-aminolevulinic acid, a precursor to porphyrin synthesis, are more resistant to NUV light (Tuveson and Sammartano, 1986).

2.8 Catalase and Spontaneous Mutagenesis

The identification of DNA as one of the targets of toxic oxygen species was further supported by the knowledge that the rate of spontaneous mutagenesis in *S. typhimurium* is influenced by the *oxyR* regulon. Cells with an *oxyR* deletion have higher rates of spontaneous mutation than their wild type parent, while cells carrying the *oxyR1* mutation, which results in the overproduction of oxidative defense products have lower rates of spontaneous mutation (Sies et al., in press). The largest increase in mutation rates of *oxyR* deletion involves T:A to A:T transversions, the predominant oxygen related mutation (Levin et al., 1982).

Higher mutation rates have been observed during the aerobic growth

of cells lacking both the manganese and iron superoxide dismutases (Farr et al., 1986). Sies et al. (in press) introduced various multicopy plasmids bearing oxidative defense products into *oxyR* mutants and observed the effect on spontaneous mutation rates. Plasmids bearing *sodA* (superoxide dismutase) lowered the mutation rates slightly and *katG* (HPI) and *ahp* (alkylhydroperoxidase) bearing plasmids decreased the spontaneous mutation rate to near that of the wild type while extracellular addition of catalase or superoxide dismutase did not lower the mutation rates. Sies et al. (in press) suggest that these results confirm the role of reactive oxygen species from respiration in mutagenesis.

2.9 Catalase Gene and Protein Sequences

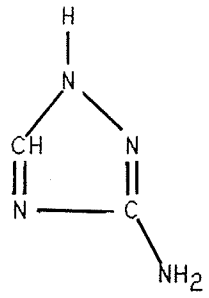
The DNA sequences from rat liver catalase ([cDNA]; Osumi et al., 1984; Furuta et al., 1986), human catalase ([cDNA]; Quan et al., 1986; Bell et al., 1986) and yeast T-catalase (Hartig and Ruis, 1986) have been determined and the amino acid sequences predicted. A comparison of these sequences to the chemically determined amino acid sequences of bovine liver catalase (BLC) and bovine erythrocyte catalase (Schroeder et al., 1982a; 1982b) revealed a 90% homology between BLC and rat liver or human catalase, and a 40.7% homology between BLC and the most conserved portion of yeast catalase T. The three dimensional structures of *Penicillium vitale* catalase (Vainshtein et al., 1986) and BLC (Fita and Rossmann, 1985b) were compared (Melik-Adamyan et al., 1986) and found to be highly conserved at the tertiary and secondary levels. The x-ray crystal structures of BLC identified the amino acids important in catalysis or heme binding (Fita and Rossmann, 1985b) and the conservation of these

same amino acids in other catalases substantiated this.

The x-ray structure and amino acid sequence of BLC were in complete agreement (Fita and Rossmann, 1985b). Although the original sequence of BLC (Schroeder et al., 1982a) was only 506 amino acids in length, predicted sequences from other sources are 526 amino acids in length (Furuta et al., 1986). BLC appeared to be in a processed form when it was isolated. The 3-D structure of catalase has been described as a N-terminal arm followed by a β-barrel and then an α-helical domain (Fita and Rossmann, 1985b). The β-barrel, made up of eight antiparallel β-pleated sheets, contains the heme binding residues and consequently shows the highest levels of conservation. As the sequence proceeds toward the carboxyl terminus, the proteins diverge into what has been identified as a high affinity NAD(P)H binding region in human catalase (Kirkman et al., 1984) and BLC (Fita and Rossmann, 1985a) and a flavodoxin binding domain in *P. vitale* catalase (Vainshtein et al., 1986). In yeast T catalase, although the carboxyl terminus is divergent, no particular function has been assigned to it (Hartig and Ruis 1986). The functional significance of the NAD(P)H and flavodoxin binding domains remains unclear since they do not appear essential to catalase activity.

The amino acids in the β-barrel domain which are essential to catalysis include His-74, Ser-113, and Asn-147. These residues are conserved in all of the aforementioned catalases. There are many residues with heme contacts; notably, the phenol group of Tyr-347 occupies the 5th ligand position of the Fe in protoporphyrin IX. Tyr-357 is conserved in yeast T catalase and only 5/25 heme contacts shorter than 0.4 nm in BLC are replaced in yeast catalase T (Hartig and Ruis, 1986). His-74 has been identified as the probable active site since it is the site of

action of aminotriazole in BLC (Margoliash et al., 1960). AT is active on compound I of the catalase reaction and thus H_2O_2 is required to form it. The structure of AT is shown below.



(Halliwell and Gutteridge, 1985)

The heme groups in *P. vitale* and BLC are buried deep within the protein forcing substrates to enter a polar channel which becomes hydrophobic as it approaches the heme group (Murthy et al., 1981).

2.10 *E. coli* Control Sequences

Gene expression in *E. coli* is controlled by the DNA sequence and a complex network of proteins capable of interacting with nucleic acids. Although there are many types of regulation in *E. coli* only a few elements have been well characterized. The recent accumulation of DNA sequences has aided in the identification of consensus sequences or structural elements typical of *E. coli* regulatory elements. The sequences have also been useful in the analysis of codon usage as a mechanism of regulation of gene expression.

The promoters of RNA transcription are very important in the regulation of gene expression. In a recent study by Harley and Reynolds (1987), 263 promoter regions were aligned and the promoter sequences were compared. They found the controlling elements upstream of the transcriptional start site, (-35 and -10 regions) to have the following

consensus sequence:

c.a..t.....TTGACA..t.....ggTATAATg

-35 -10

The hexameric sequences at -35 and -10, and the 17 bp distance between them were highly conserved. A few nucleotides occurring with frequencies greater than random variation are also indicated. This group found a correlation between promoter homology and strength in many instances, and a poor consensus to the -35 sequence in positively regulated genes. Initiation of the transcript commonly occurred 6-8 bp downstream from the 3' end of the -10 region at a purine residue. The purine was commonly an A residue which has been weakly associated with the sequence CAT at the startsite. The promoter may also be capable of binding positive or negative regulatory proteins that can further modify gene expression.

Although termination of transcription has not been as well characterized as promotion, certain features important in termination have been identified. All rho independent, or simple terminators contain a region of hyphenated dyad symmetry followed by a series of U residues. More complex terminators (Platt and Bear, 1983) have been characterized and found to show extreme variation. It is thought however that many rho dependent terminators still show a region of hyphenated dyad symmetry rich in G/C residues, but this was not followed by the poly U sequence (Rosenberg and Court, 1979).

Translational control is not as clearly understood. Proteins can initiate with a methionine, valine, or leucine codon, indicating the need for other factors to control recognition. In most promoters, this appears to be a sequence called the Shine-Dalgarno sequence (AGGAGG) which is complementary to the 3'-end of the 16s rRNA. However, this

sequence in not always present (Ganoza et al., 1987).

The codon choice of organisms is representative of the availability of tRNA molecules in the cell (Ikemura and Ozeki, 1982). The codon choices also vary depending of the anticodon modifications which occur in the cell. Organisms have been shown to bias the codon usage in highly expressed genes to achieve optimum codon-anticodon interaction energies that are neither too strong nor too weak. Thus in highly expressed genes the codon usage will be clearly biased to avoid the usage of low abundance tRNAs and codon-anticodon interactions which are energetically unfavourable (Grosjean and Fiers, 1982). The codon usage acts as a form of translational control, modulating the expression of genes at the level of translation. Codon context has also been implicated as a factor influencing gene expression (Shpaer, 1986).

2.11 Characterization of a Plasmid Encoding the HPI Catalase

The Carbon and Clarke plasmid bank (Carbon and Clarke, 1976) was screened for potential catalase carrying plasmids by dropping 30% H_2O_2 on the edge of the colonies and looking for above average oxygen evolution. Of 88 colonies that appeared to have above average oxygen evolution, only one, pLC36-19, had higher catalase when assayed in the oxygraph. Characterization of pLC36-19 revealed it to be a 25.5 kb plasmid made up of a 19.1 kb *E. coli* genomic DNA insert and a 6.4 kb ColE1 vector. The plasmid complemented catalase deficient mutants and produced the HPI protein in these mutant cells. Minicell preparations of pLC36-19 containing cells produced an 84 000 dalton polypeptide, the same size as the HPI subunit (Loewen et al., 1983). The characterization of the HPI encoding gene on this plasmid is the subject of the study presented in this thesis.

MATERIALS AND METHODS

3.0 MATERIALS AND METHODS

3.1 Bacterial Strains and Plasmids

All bacterial strains used in this study were derivatives of *Escherichia coli* K-12. Table 1 lists the strains with their genotype and their source or method of construction. Stationary phase cultures of new strains were stored in 50% glycerol at -20°C and in 8% DMSO at -60°C. The plasmids used and/or constructed during this study are described in Table 2.

3.2 Standard Bacterial Growth Conditions, Media, and Solutions

3.2.1 Bacterial Growth Conditions

Unless otherwise stated, broth cultures of *E. coli* were grown with aeration at 37°C. Precultures were prepared by inoculating 10 mL of LB broth with a loop full of the glycerol stock culture and incubating overnight. Log phase cultures were prepared by subculturing 0.1 mL of the preculture into a fresh tube of LB broth and shaking for 2-4 h. Cell densities of liquid cultures were measured using a Klett-Summerson colorimeter equipped with a blue filter.

Agar plates were usually inverted and incubated one to two days at 37°C in a Fisher Isotemp Incubator.

3.2.2 Media and Commonly Used Solutions

A. LB Medium (Miller, 1972)
per liter:

10.0 g	tryptone (Difco)
5.0 g	yeast extract (Difco)
5.0 g	NaCl

for agar plates: add 10.0 g agar (Difco)

Table 1. List of *E. coli* strains

Strain	Genotype	Reference or Source
JA200	F$^+$ *recA trp thr leu*	Clarke & Carbon(1976)
HB101	*pro leu rpsL hsdM hsdR endI recA lacY*	Boyer & Roulland-Dussoix(1969)
JM101	*supE thi* (*lac proAB*)/F' *traD36 proAB*	Yanisch-Perron et al. (1985)
MP180	*thi-1* HfrH	Pearson(1972)
CSR603	*thr-1 leuB6 proA2 phr-1 recA1 argE3 thi-2 uvrA6 ara14 lacY1 galK2 xyl-5 mtl-1 gyrA98 tsx33 ppsL31 λ$^-$ supE44*	Sancar et al.(1979)
RR1	*pro leu rpsL hsdM hsdR endI lacY*	Bolivar et al.(1977)
NK5012	*sulI$^+$ thr leu lacY1 tonA*	H. Goldie
CSH7	F$^-$ *lacY rpsL thi-1*	C.S.H.C.[a]
CSH57a	F$^-$ *leuB6 proC83 purE42 trpE28 his-208 argG77 ilvA681 metA160 thi-1 ara-14 lacY1 galK2 xyl-5 mtl-1 azi-6 rpsL109 tonA23 tsx-67 supE44 malA38 xthA*	C.S.H.C.[a]
JM96	*cysH thr leu trp his argH thi lac xyl gal mal rpsL*	Jones-Mortimer (1968,1973)
JF627	*thi-1 pyrE60 pyrG edd relA1 argE3 his-4 proA2 thr-1 leu-6 mtl-1 xyl-5 ara-14 galK2 lacY1 rpsL31 supE44*	Friesen et al.(1976)
4206	Hfr *thi recA relA deoB*	C.G.S.C.[b]
4913	Hfr *pheS5 relA1 tonA22* T2^R *pit10 spoT1*	C.G.S.C.[b]
5128	Hfr(R4) *deoA upp udp metB argE relA*	C.G.S.C.[b]
5236	*aceB6 ppc-2 glc-1 thi-1 relA1 lacZ43 λ$^-$ spoT1*	C.G.S.C.[b]
6374	F$^-$ *hisS glyA relA argH thi lacYorZ xyl mtl rpsL supE*	C.G.S.C.[b]
PA309	*trp his argH leu rpsL*	Jones-Mortimer

UM1	as CSH7 but *katE1 katG14*	Nitrosoguanidine
UM2	as CSH57a but *katE2 katG15*	Nitrosoguanidine
UM5	as UM1 but *his*	Nitrosoguanidine
UM53	as UM2 but *recA*	4206 x UM50 ⟶ *thyA*$^+$ *recA*
UM56-64	as CSH57a but *katF3 katG16*	Nitrosoguanidine
UM68	as HB101 but *kat-26*	Nitrosoguanidine
UM69	as HB101 but *kat-27*	Nitrosoguanidine
UM120	as MP180 but *katE*::Tn*10*	MP180 x λ561 ⟶ Tc^R
UM122	as MP180 but *katF*::Tn*10*	MP180 x λ561 ⟶ Tc^R
UM123	as UM69 but *kat-18*	Nitrosoguanidine
UM178	as UM5 but *lac*$^+$ *katG*$^+$	CSH64 x UM5 ⟶ *Lac*$^+$
UM196	as UM178 but *katG17*::Tn*10*	UM178 x λ561 ⟶ Tc^R
UM197	as CSH57A but *katG17*::Tn*10*	P1(UM196)xCSH57a ⟶ Tc^R
UM202	as MP180 but *katG17*::Tn*10*	P1(UM196)xMP180 ⟶ Tc^R
UM228	as HB101 but *kat-19*	Nitrosoguanidine
UM229	as HB101 but *kat-28*	Nitrosoguanidine
UM230	as HB101 but *kat-29*	Nitrosoguanidine
UM262	as RR1 but *recA katG17*:Tn*10*	4206 x UM263 ⟶ *thyA*$^+$ *recA*
UM263	as RR1 but *thyA katG17*:Tn*10*	P1(UM196)xRR1 ⟶ Tc^R Trimethoprim ⟶ *thyA*

[a] Cold Spring Harbour Collection
[b] Coli Genetic Stock Center, B. Bachmann, Curator

Table 2. Bacterial plasmids

Plasmid	Characteristics	Reference of Source
pAT153	Ap^R Tc^R	Twigg & Sherratt(1980)
pBR322	Ap^R Tc^R	Sutcliffe(1978)
pKK232-8	Ap^R Cm^S	Brosius(1984)
ColE1	$colE1^R$	Hershfield et al.(1974)
pLC36-19	$colE1^R$ *katG*$^+$	Clarke & Carbon(1976)
pA30	pLC36-19::Tn*5*; Km^R	λNK467 x JA200/pLC36-19
pE6	pLC36-19::Tn*5*; Km^R	λNK467 x JA200/pLC36-19
pC1	pLC36-19::Tn*5*; Km^R	λNK467 x JA200/pLC36-19
pD31	pLC36-19::Tn*5*; Km^R	λNK467 x JA200/pLC36-19
pBT1	Km^R *katG20*	*Hin*dIII deletion of pC1
pBT2	Km^R *katG*$^+$	*Hin*dIII deletion of pC1
pBT5	Ap^R Tc^S *katG21*	pLC36-19·*Sph*1+pAT153·*Sph*1
pBT22	Ap^R Tc^S *katG*$^+$	pLC36-19·*Hin*dIII+pAT153·*Hin*dIII
pBT54	Ap^R Tc^S *katG*$^+$	pLC36-19·*Hin*dIII+pAT153·*Hin*dIII
pBT24	Ap^R Tc^S *katG22*	*Eco*R1 deletion from pBT22
pBT25	Ap^R Tc^S *katG25*	BAL31 deletion from pBT22
pBT28	Ap^R Tc^S *katG*$^+$	BAL31 deletion from pBT22
pBT29	Ap^R Tc^S *katG23*	BAL31 deletion from pBT22
pBT30	Ap^R Tc^S *katG24*	BAL31 deletion from pBT22
pGprm1	Ap^R Cm^R	pBT22·*Bgl*II+pKK232-8·*Bam*HI
pMSKatG14	Ap^R Tc^S *katG14*	as pBT22 but *katG14*
pMSKatG15	Ap^R Tc^S *katG15*	as pBT22 but *katG15*

pMSKatG16	Ap^R Tc^S *katG16*	as pBT22 but *katG16*

· indicates a restriction enzyme digest
+ represents a ligation
x represents an infection
/ indicates that the host contains the plasmid succeeding the symbol

for LB top agar: add 8.0 g agar

for R top agar: add 8.0 g agar; The agar was supplemented after autoclaving with 6.0 mM glucose and 2.0 mM $CaCl_2$.

B. M9 Medium (Miller, 1972)
per liter:

0.5 g	NaCl
6.0 g	Na_2HPO_4
1.0 g	NH_4Cl
3.0 g	KH_2PO_4

for agar plates: add 10.0 g agar

The medium was supplemented after autoclaving with 3 µM vitamin B1, 1.0 mM $MgSO_4$, and 1.0 mL of trace elements.

Trace Elements
per liter:

2.5 g	$FeSO_4.H_2O$
2.9 g	H_3BO_3
1.2 g	$CoSO_4.7H_2O$
0.1 g	$CuSO_4.5H_2O$
0.09 g	$MnCl_2.4H_2O$
2.5 g	$Na_2MoO_4.2H_2O$
2.1 g	$ZnSO_4.7H_2O$
5.0 mL	conc. H_2SO_4

Other supplements, including 16.7 mM glucose, 90.8 mM pyruvate, 0.16 mM amino acids, 94 µM adenosine, 0.22 mM uracil, and 0.1 mM cytidine were added as required.

C. TB or λ Medium (Silhavy et al., 1984)
per liter:

5.0 g	NaCl
10.0 g	tryptone

TBYM broth was prepared by supplementing TB broth with 5.8 mM maltose and 0.01% (w/v) yeast extract.

for agar plates: add 10.0 g agar

for top agar: add 7.0 g agar

D. F-top agar (Miller, 1972)
per liter:

8.0 g	NaCl
8.0 g	agar

E. 2xYT Medium (Miller, 1972)
per liter:

5.0 g	NaCl
16.0 g	tryptone
10.0 g	yeast extract

This was adjusted to pH 7 before autoclaving.

F. K Medium (Rupp et al., 1971)
M9 Medium was supplemented before autoclaving with 1% casamino acids and after autoclaving with 0.3 µM vitamin B1 and 16.6 mM glucose.

G. Hershey's Salts (Worcel and Burgi, 1974)
per liter:

5.4 g	NaCl
3.0 g	KCl
1.1 g	NH_4Cl
0.015 g	$CaCl_2.2H_2O$
0.02 g	$MgCl_2.6H_2O$
0.02 mg	$FeCl_3.6H_2O$
0.087 g	KH_2PO_4
0.028 g	Na_2SO_4
12.1 g	Trizma base
HCl to pH 7.4	

H. Hershey's Medium (Worcel and Burgi, 1974)
Hershey's Salts were supplemented after autoclaving with 3 mM vitamin B1, 22.2 mM glucose and 0.16 mM amino acids as required by the strains.

I. Maloy Medium (Maloy and Nunn, 1981)
per liter:

5.0 g	tryptone
5.0 g	yeast extract
10.0 g	NaCl
50.0 mg	chlortetracycline hydrochloride

After autoclaving, the medium was cooled to 45°C, then 20 mL of sterile 3.5 M $NaH_2PO_4.H_2O$, 6.0 mL of sterile 10 mM fusaric acid, and 5.0 mL of sterile 20 mM $ZnCl_2$ were added.

The following antibiotics ampicillin (100 µg/mL), tetracycline (15 µg/mL), streptomycin (80 µg/mL), and kanamycin (25 µg/mL) were added as required to the above media. Antibiotic containing agar plates are represented by adding the first letter of the antibiotic name onto the media name, for example, LBA are ampicillin containing LB plates.

J. SM Buffer (Miller, 1972)

0.05 M	Tris-HCL pH 7.5
0.01 M	$MgSO_4$
0.1%	gelatin

K. Phenol
Redistilled phenol was buffered with

0.1 M	NaCl
0.1 M	Tris-HCl pH 7.6
1.0 mM	EDTA

L. TE (Maniatis et al., 1982)

10.0 mM	Tris-HCl pH 8.0

1.0 mM EDTA pH 8.0

M. TEN_{250}
TE containing 250 mM NaCl

N. Ethanol
Unless otherwise stated all ethanol was 95%.

3.3 Measurement of DNA and Protein Concentrations

DNA concentrations were estimated by electrophoresing a specific volume of a sample of unknown concentration beside a sample containing a known quantity of DNA. After staining the gel in EtBr, the two samples were photographed under an indirect UV light source. The concentration was then estimated on a comparative basis. DNA concentrations were also determined spectrophotometrically by taking the A_{260} of the sample. Since 1 A_{260} unit of bacterial DNA was known to contain approximately 50 µg of DNA, the quantity of DNA was calculated on this basis. The ratio of the A_{260} to the A_{280} was then used as an indicator of purity, pure DNA having an A_{260}/A_{280} ratio of 1.8 (Maniatis et al., 1982).

Protein concentrations were estimated spectrophotometrically by taking the ratio of the A_{280} to the A_{260}. This ratio was compared to a table of values for known protein concentrations to get the value F, where:

$$F = \frac{2.303}{\beta_{280}} \times \frac{\text{per cent protein}}{100}$$

F was then substituted into the following formula to give the protein concentration in mg/mL (Layne, 1957).

$$[\text{protein}] = F \times 1/d \times A_{280}$$

where d = cuvette width in cm

3.4 Isolation of Catalase Mutants by Nitrosoguanidine (NG) Mutagenesis

HB101 was grown to early exponential phase (2 x 10^8 cells/mL) in LB broth and the cells were collected by centrifugation at 12 000 x g for 5 min. The pellet was resuspended in 10 mL of 1 M sodium citrate pH 5.5 and 0.5 mL of NG (1 mg/mL prepared in citrate buffer) was added. This was incubated with aeration for 30 min at 37°C. The cells were pelleted by centrifugation at 12 000 x g for 10 min and after washing the pellet once in M9 medium, it was resuspended in 10 mL of M9 medium supplemented with glucose, proline, and leucine. The culture was shaken overnight at 37°C, then serially diluted in SM and plated on LBS plates. The dilution which gave approximately 200 colonies per plate was used to spread 11 plates, which after incubation overnight were screened for catalase mutants by the colony H_2O_2 drop test described in section 3.7.1. Colonies which evolved little or no oxygen were quickly picked and transferred to another LBS plate for rescreening.

3.5 Transposon Mutagenesis

3.5.1 Preparation of Transposon Carrying Phage Lysates

The *E. coli* strain NK5012 was grown to mid-log phase in TBYM broth. A phage lysate of λ561(*b221 C1857 rex*::Tn*10 Oam29 Pam80*) (Foster et al., 1981) or λNK467(*b221 C1857 rex*::Tn*5 Oam29 Pam80*) (DeBruijn and Lupski, 1984) was serially diluted and 0.1 mL aliquots were mixed with 0.1 mL aliquots of NK5012. The phage were allowed to adsorb at room temperature for 20 min, then 2.5 mL of molten TB top agar was added, and each mixture was plated on a fresh TB agar plate. After overnight incubation at 42°C, 1-3 plaques were picked into 0.5 mL of an overnight NK5012 culture grown in TBYM broth. The mixture was incubated

20 min at room temperature, mixed with 7.5 mL of TB top agar, and plated on 3 TB agar plates. The plates were then kept at 42°C for no more than 16 h before scraping the top agar off and vortexing it for 30 sec with 2.0 mL of SM and 5 drops of chloroform. The cell debris was removed by centrifugation at 12 000 x g for 5 min, and the supernatant was stored at 4°C. Lysates were titered in the same way as single plaques were isolated at the start of this procedure, and the ability of the lysates to introduce transposons was tested by placing a drop of lysate onto recipient cells spread on an appropriate selective medium. If a lysate was good, several colonies would grow after a 24 h incubation at 42°C.

3.5.2 Transposon Tn*10* Mutagenesis of MP180 and UM178

MP180 or UM178 was grown to late mid-log phase in TBYM broth. The cells were pelleted by centrifugation at 12 000 x g for 5 min and the pellet was resuspended in 0.5 mL (1/20 volume) of TB broth. Phage λ561 was added to give a MOI of 0.2 and this was incubated at 37°C for 45 to 60 min. Aliquots were plated on 4 to 10 LBT plates and incubated overnight at 42°C. Any colonies which grew were judged to contain a Tn*10* insertion in the chromosome because the Tn*10* confers tetracycline resistance, while the λ phage was unable to survive in MP180 or UM178 at 42°C. The temperature sensitive phage repressor, *c1857*, forced the phage to grow lytically at 42°C, but the phage grew poorly in the *su-* host because of the amber mutations in essential phage genes. The following day the Tn*10* containing colonies were screened for catalase mutants by the H_2O_2 drop test.

The stability of the transposon insertions was tested by looking for revertants in the population. Colonies grown on LB plates were picked onto LBT plates to screen for tetracycline sensitive (Tc^S)

revertants. Aliquots of the cultures were spread on Maloy medium, a medium selective for Tc^S cells (Maloy and Nunn, 1981). Since the complete loss of the transposon making the cells Tc^S should result in the cell concomitantly regaining catalase activity, 10 mM H_2O_2 was added to an early log phase culture to select for catalase positive cells. After 15 min, 10^{-1} and 10^{-2} dilutions of the culture were plated on LB plates, incubated overnight at 37°C, then scored for catalase the following day.

3.5.3 Transposon Tn*5* Mutagenesis of pLC36-19

λNK467 was used to isolate Tn*5* inserts in JA200/pLC36-19 as described in section 3.5.2 for MP180, except that colonies were selected on LBK plates. Approximately 1000 Km^R colonies were scraped from the agar plates, resuspended in 100 mL of LB broth, and grown to mid-log phase. One mL of this culture was mixed with 1.0 mL of mid-log phase UM53 and the mixture was shaken at 37°C for 1 h.

After the mating, 0.1 mL aliquots of the mixture were spread on LBSK plates. Small scale alkaline-SDS preparations were done on the resulting colonies, and plasmids larger than pLC36-19 were isolated and restriction enzyme digested to identify the location of the Tn*5* insert.

3.6 Genetic Mapping of Chromosomal Genes

3.6.1 Interrupted Mating

Interrupted matings were carried out as described by Miller, (1972). The donor and recipient strains were grown to early exponential phase in LB broth. Ten mL of the recipient culture were placed in a prewarmed 125 mL Erlenmeyer flask, 0.1 mL of the donor cells was added, and the incubation was continued with aeration at 37°C. Each of the

samples taken at specific time intervals was diluted 1:10 in SM, vortexed vigorously for 60 sec, then 0.1 mL aliquots were plated on an appropriate selective medium.

3.6.2 P1 Transduction

To prepare the P1 lysate needed for the transduction the strain was grown to mid-log phase in LB broth. A 0.1 mL aliquot of the culture was mixed with 20 µl of $P1_{(vir)}$ and the mixture was incubated at 37°C for 20 min. The phage adsorbed bacteria were mixed with 2.5 mL of molten R top agar and plated on an LB plate. After overnight incubation at 37°C, the top agar was scraped off and vortexed with 1.0 mL of LB broth and 1 drop of chloroform per plate. After allowing the vortexed mixture to sit at room temperature for 30 min the debris was pelleted at 12 000 x g for 10 min and the supernatant was collected to provide a P1 lysate.

The strain to be transduced was grown to late log phase (1×10^9 cells/mL) and collected by centrifugation at 12 000 x g for 5 min. The pellet was resuspended in 1.0 mL of 1 M $CaCl_2$ and 0.2 mL aliquots of the cell suspension were mixed with 0.1 mL of three different dilutions of the P1 lysate. As a control, 0.2 mL of the cell suspension was added to a fourth tube containing no P1 lysate. After incubation for 60 min at 37°C, 0.2 mL of 1 M sodium citrate was added. This was mixed with 2.5 mL of molten F top agar and plated on the appropriate selective medium. Following a 24-48 h incubation, colonies were picked onto other media to score for the presence or absence of particular genetic markers. Specific genetic markers were scored as follows: *pyrG* (CTP synthetase) strains for their requirement for 0.1 mM cytidine; *cysH* (adenylsulfate reductase) strains for their requirement of

0.16 mM cysteine; *pheS* (phenylalanyl-tRNA synthetase) strains for their inability to grow at 37°C; *metB* (cystathionine γ-synthetase) strains for their requirement of 0.16 mM methionine; *argH* (argininosuccinate lyase) strains for their requirement of 0.16 mM arginine, and *ppc* (phosphoenolpyruvate carboxylase) strains for their inability to grow on 0.2% glucose (Bachmann, 1983). The resulting genetic linkage data were used to calculate map distances by the method of Wu (1966). The cotransduction frequency of any two markers, x, was substituted into the following formula to give the map distance between markers in minutes:

$$2(1-\sqrt[3]{x}) = \text{distance in minutes.}$$

3.7 Assay of Catalase Activity

3.7.1 Colony H_2O_2 Drop Test

A drop of 30% H_2O_2 (Fisher) was placed on the edge of each colony. Catalase positive colonies evolved oxygen which appeared as bubbles, while catalase negative colonies bubbled little or not at all.

3.7.2 Quantitation of Catalase by Oxygraph

Catalase activity was quantitated in a Gilson Model 5/6H oxygraph equipped with a Clark electrode using the method of Rørth and Jensen (1967). Assays were done using a final H_2O_2 concentration in the reaction chamber of 1.5 mM or 60 mM. One unit of catalase was defined as the amount of enzyme dismutating 1.0 μM of H_2O_2 per min at 37°C.

3.7.3 Catalase and Peroxidase Visualization on Polyacrylamide Gels

Cell free crude extracts electrophoresed on native polyacrylamide gels as described by Davis (1964), were stained for catalase or

peroxidase activities using the method of Gregory and Fridovich (1974).

Cells were collected from 250 mL cultures of log phase or stationary phase cells and resuspended in 1.0 mL of SM. The cells were sonicated for four-30 sec periods with a 30 sec incubation on ice between each sonication. Cell debris was removed from the suspension by centrifugation at 12 000 x g for 10 min, leaving a cell free extract which could then be used for electrophoresis. Appropriate volumes of the sample containing 50% glycerol and 0.025% bromophenol blue dye were loaded on a 9.5% polyacrylamide slab gel prepared in 67 mM Tris-HCl pH 8.1 and electrophoresed at 25 mA for 3 h or until the blue dye reached the bottom. The running buffer was Tris-Glycine, (10 mM Tris-HCl, 70 mM glycine).

The gel was removed from the apparatus and soaked in 50 mM potassium phosphate pH 7.0 containing 0.4 mg/mL of diaminobenzidine hydrochloride (Sigma) for 45 min in the dark. If the gel was to be stained for catalase the solution also contained 0.05 mg/mL of horseradish peroxidase (Sigma). The gel was rinsed in water, then soaked for approximately 2 h in 20 mM H_2O_2 also prepared in 50 mM potassium phosphate pH 7.6.

Catalase bands appeared as white bands on a brown background while peroxidase bands appeared as brown bands on a light background.

3.7.4 Ascorbate Induction of Catalase

Cultures were grown with aeration at 37°C in a Klett flask containing 25 mL of LB broth. At a Klett of approximately 50 (time=0), 1.0 mL of freshly prepared 50 mM ascorbate (J. T. Baker Chemical Company) was added. Samples were taken before and after the ascorbate was added and the catalase activity was measured in the oxygraph using 1.5 mM H_2O_2.

3.8 Plasmid DNA Isolation

3.8.1 Plasmid Screening Preparations

Isolation of plasmid DNA from bacterial colonies grown on a plasmid selecting medium was done as described by Barnes (1977). Most of the colony was picked with a toothpick and resuspended in 25 µL of 50 mM NaOH, 0.5% SDS, 5 mM EDTA, and 0.025% bromocresol green. After a 45 to 60 min incubation at 68°C, 2.5 µL of 25% Ficoll was added and the sample was loaded on an agarose gel for electrophoresis.

3.8.2 Preparation of Plasmid DNA by the Alkaline Extraction Method

Materials

Media: LB broth + the appropriate plasmid selecting antibiotic

Lysis Buffer: 50 mM glucose
10 mM EDTA
25 mM Tris-HCl pH 8

Alkaline-SDS: 1% SDS /0.2 N NaOH prepared fresh

High Salt A: 3.0 M potassium acetate
1.8 M formic acid

High Salt B: 3 M sodium acetate pH 4.8

Acetate-MOPS: 0.1 M sodium acetate
0.05 M MOPS
adjusted to pH 8

TAES: 0.04 M Tris-HCl pH 7.6
0.1 M sodium acetate pH 6.0
1.0 mM EDTA
0.1% SDS

RNase A: Ribonuclease A (Sigma) was dissolved at a concentration of 2 mg/mL in TE and boiled for 5 min.

3.8.2.1 Small Scale Plasmid Isolation for Restriction Enzyme Digestion (Birnboim, 1983)

The entire procedure was carried out in microcentrifuge tubes and centrifugations were done in an Eppendorf Centrifuge Model 5412. The

cell pellet from a 0.5 mL aliquot of an overnight culture of the plasmid bearing strain was resuspended in 0.1 mL of lysozyme dissolved in lysis buffer (1 mg/mL) and placed on ice for 5 min. Two volumes of alkaline-SDS were added, the tube was gently inverted, then returned to the ice for another 5 min. One and one half volumes of high salt A was then mixed in by inversion, and the tube was returned to the ice for 15 min. The cell debris was removed with a 15 min centrifugation and the supernatant was transferred to a tube containing 0.9 mL of ethanol. The tube contents were thoroughly mixed and then stored at -20°C for 30 min. The pellet obtained after a 5 min centrifugation of the mixture was resuspended in 0.1 mL of acetate-MOPS and reprecipitated with two volumes of cold ethanol. The DNA was pelleted with a 5 min centrifugation, vacuum dried, resuspended in 50 µL of water, and stored at -20°C.

3.8.2.2 Large Scale Plasmid Isolation

Two 500 mL cultures of the plasmid harbouring strain were grown with vigorous aeration in two-2L flasks at 37°C. When the cells reached an A_{600} of 0.6, chloramphenicol was added to a concentration of 170 µg/mL and the incubation was continued overnight (Clewell, 1972).

Plasmid DNA was isolated by the method of Birnboim and Doly (1979) as outlined below. Cells were harvested by a 10 min centrifugation at 4000 x g, resuspended in 40 mL of lysis buffer containing 1 mg of lysozyme per mL, and placed on ice for 30 min. Eighty mL of alkaline-SDS was gently mixed in and the lysed cells were returned to the ice bath for another 5 min before mixing in 60 mL of high salt B and continuing the incubation at 0°C for another hour. The cell debris was then removed by centrifugation at 27 000 x g for 10 min and the supernatant was precipitated with two volumes of ethanol for 30 min at

-60°C. The precipitate was collected with a 10 min centrifugation at 4000 x g.

The resulting pellet was dissolved in 40 mL of TAES and extracted with 40 mL of phenol-chloroform (1:1). The aqueous phase was removed after the layers were separated by a 10 min centrifugation at 4000 x g. The chloroform-phenol layer was reextracted with another 40 mL volume of TAES; the aqueous layers were combined and precipitated with two volumes of cold ethanol at -60°C. After at least 30 min, the pellet was collected by a 10 min centrifugation at 4000 x g.

The pellet was dissolved in 10 mL of H_2O and 4 mL of 1 M sodium acetate pH 8.0. The DNA was precipitated and collected as described in the previous paragraph. After repeating this step once more the pellet was dissolved in 4 mL of H_2O and 0.2 mL of RNase A was added. The mixture was incubated at 37°C for 30 min, then mixed with 75 µL of 4 M sodium acetate pH 6 and two volumes of room temperature ethanol. The DNA was allowed to precipitate at room temperature for 10 min and then collected by centrifugation at 12 000 x g for 10 min. Room temperature precipitations were repeated until the supernatant was clear. The pellet was then dried under vacuum and resuspended in 0.5-1.0 mL of TE.

Plasmid vectors for cloning were further purified on a CsCl gradient (Maniatis et al., 1982). The DNA solution was increased to a volume of 3.0 mL and mixed with 3.0 g of CsCl (Beckman scintillation grade) in a 5.0 mL nitrocellulose tube. EtBr was added to a concentration 10 mg/mL and the tube was topped with paraffin oil.

The tubes were centrifuged at 35 000 rpm in a 50.1 Ti rotor for 40 h at 20°C. The tubes were removed and the plasmid band (below the genomic DNA band) was visualized with a UV light and collected from the

top of the tube with a Pasteur pipette. The DNA was extracted four times with an equal volume of isoamyl alcohol, then dialyzed overnight at 4°C against TE and stored at -20°C.

3.9 *E. coli* Genomic DNA Isolation

Materials

Medium: 1 liter of M9 medium supplemented with
16.7 mM glucose
5.0 g tryptone
1.0 g yeast extract

Saline-EDTA: 0.15 M NaCl
0.1 M EDTA pH 8

Lysozyme Solution: 2 mg/mL lysozyme chloride (Sigma) dissolved in Saline-EDTA

Tris-SDS: 0.1 M Tris-HCL pH 7.6
0.1% SDS
0.1 M NaCl

Saline-Citrate: 0.15 M NaCl
0.015 M trisodium citrate pH 7.0
dilute 1:10 before using

RNase A: 1 mg of ribonuclease A (Sigma) dissolved in 1 mL of saline-citrate and boiled for 5 min.

One liter of *E. coli* cells was grown overnight with aeration in a 37°C incubator room. The cells were collected by a 10 min centrifugation at 4000 x g and washed once in 100 mL of saline-EDTA. The washed cells were weighed, then resuspended in 1.0 mL of saline-EDTA per gram of cells. The suspension was incubated for 30 min at 37°C, frozen in a -60°C freezer, then suspended as it thawed in 10 volumes of Tris-SDS. Following two more freeze-thaw cycles, the suspension was mixed with an equal volume of buffered phenol and stirred for 20 min at 4°C. The emulsion was separated by centrifugation at 4000 x g for 15 min and the aqueous phase was mixed with two volumes of cold ethanol. The resulting

precipitate was collected by a 10 min centrifugation at 4000 x g and then resuspended in 20 mL of saline-citrate. RNase A was added to a concentration of 50 µg/mL. The solution was incubated at 37°C for 30 min, cooled, and mixed with 0.3 mL of 4.0 M sodium acetate pH 6.0. One volume of room temperature ethanol was added, and the DNA was allowed to precipitate for 10 min at room temperature. The solution was centrifuged for 10 min at 4000 x g, and pellet was resuspended in saline-citrate, then reprecipitated as described above. When the last step had been repeated until the supernatant was no longer cloudy, the pellet was dried and resuspended in TE to a concentration of 2 mg/mL.

3.10 DNA Cloning and Restriction Mapping

3.10.1 Restriction Enzyme Digests

*Acc*I, *Ava*I, *Bam*HI, *Bcl*I, *Bgl*II, *Cla*I, *Dra*I, *Eco*RI, *Eco*RV, *Hae*III, *Hin*cII(*Hin*dII), *Hin*dIII, *Hpa*I, *Nru*I, *Pst*I, *Pvu*II, *Rsa*I, *Sal*I, *Sau*3AI, *Sma*I, *Ssp*I, *Taq*I, and *Xba*I were purchased from Boehringer Mannheim Canada (BMC), *Ba*lI, *Bgl*I, *Kpn*I, *Nde*I, and *Sph*I were purchased from Bethesda Research Labs (BRL), and *Asu*II was purchased from Promega Biotec. Restriction enzyme digests were done for 2 h in 10 or 20 µL volumes with 1 to 10 units of enzyme per µg of DNA at the temperature recommended by BMC. The incubation buffer chosen (high, medium or low salt) was the one closest to the BMC recommended buffer except in the cases of *Sma*I and *Sph*I whose buffers were prepared specifically as recommended by BMC.

Buffer	NaCl	Tris-HCl pH7.5	DTT	$MgCl_2$
low	-	10 mM	1 mM	10 mM
medium	50 mM	10 mM	1 mM	10 mM
high	100 mM	50 mM	1 mM	10 mM

(Maniatis et al., 1982)

3.10.2 Gel Electrophoresis of DNA

Materials

6 x Loading Buffer: 100 mM EDTA
0.25% bromophenol blue
15% Ficoll
(Maniatis et al., 1982)

10 x TBE Electrophoresis Buffer (Maniatis et al., 1982)
per liter:
108 g Trizma base (Sigma)
55 g boric acid
8.5 g $Na_2EDTA.2H_2O$

30% Acrylamide
per 100 mL:
29 g acrylamide (Bio-Rad Laboratories)
1 g NN'-methylene bisacrylamide (Sigma)
Make up to 100 mL with distilled water.

3.10.2.1 Preparation of DNA Samples

Samples containing 0.5 to 1.5 µg of restricted DNA were mixed 6:1 with 6 x loading buffer. *Hin*dIII, *Eco*RI, *Bam*HI, *Hpa*I, or *Hae*III digested λ DNA (BMC) and DNA fragments ranging from 1 to 12 kb in size and each differing in size by 1 kb (BRL's 1 kb ladder) were commonly used as the molecular weight markers for agarose gels while λ DNA digested with *Nde*I or pBR322 DNA digested with *Hae*III, *Hpa*I, or *Nci*I served as the molecular weight markers for acrylamide gels.

3.10.2.2 Agarose Gel Preparation and Electrophoresis

Agarose gels were used to identify or separate fragments between 0.5 and 26 kb in size. Gels were prepared by boiling agarose (Sigma Type II: medium EEO) in 1 X TBE until it was dissolved. For a 15.5 x 15.5 cm 0.8% gel, 0.64 g of agarose was dissolved in 80 mL of 1 x TBE. When the agarose had cooled to 55°C the gel was poured, the well maker was placed in the molten agarose, and the gel was left to harden. The well maker was then removed and the gel was submerged in a Bio-Rad DNA

Sub-Cell Electrophoresis system filled with 1 x TBE. The samples were pipetted into the wells and 80 V was applied for 3.5 h or until the bromophenol blue was 3/4 of the way down the gel.

3.10.2.3 Preparation and Electrophoresis of Acrylamide Gels

Polyacrylamide gels (8%) were used to identify and separate DNA fragments between 20 and 600 bp in size. A mixture of 10 mL of 30% acrylamide, and 30 mL of 1 x TBE was degassed and then mixed with 20 µL of TEMED and 30 mg of ammonium persulfate. The acrylamide mixture was then quickly poured between two 15 x 15 cm vertical gel plates and the well maker was placed 1 cm into the acrylamide. After the gel had polymerized the well maker was removed and the wells were washed before adding 1 x TBE buffer and loading the samples. The gel was placed in a Bio-Rad Protean Dual Slab Cell and run at 100V until the bromophenol blue reached the bottom (about 3 h).

3.10.2.4 Staining, Photographing, and Sizing of DNA Bands

The gels were removed from the electrophoresis apparatus and placed in 1 x TBE containing 0.05 µg/mL of EtBr for 30 min. The gel was then photographed with an indirect UV light source using a Polaroid MP4 Land Camera equipped with a Kodak 22A Wratten filter and Polaroid Type 667 film. The exposure time was 0.5 to 1.5 min at an aperture of 5.6. The sizes of the DNA fragments were determined from the photograph by measuring the distance the fragments had travelled from the origin. The distance the standard size markers had migrated from the origin was plotted as a function of the ln of their size (bp) to make a standard curve from which the size of the unknown bands could be determined by comparison.

3.10.3 Alkaline Phosphatase Treatment of DNA

Alkaline phosphatase was used to remove the 5'-terminal phosphate groups from restricted plasmid vectors, preventing re-ligation of the vector without insert DNA. Following the BMC procedure submitted by Dr. Luis Villareal, 6-8 units of calf intestine alkaline phosphatase (BMC) was added for the last 20 min of the restriction enzyme digest. The reaction was stopped by heating the sample at 65°C for 10 min in 20 mM EDTA and 0.5% SDS. The DNA was then diluted with one volume of TE + 250 mM NaCl and extracted with one volume of buffered phenol. The aqueous phase was precipitated with two volumes of ethanol and the DNA that was collected by centrifugation for 15 min was air dried and resuspended in water.

3.10.4 Ligations

Approximately 0.1 µg of the plasmid vector which had been digested and dephosphorylataed as described previously was mixed with 0.3 µg of digested insert DNA. BRL 5 x ligation buffer was added to give a final concentration of 50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5% PEG 8000, 1 mM ATP, and 1 mM DTT. The volume was made up to 10 or 20 µL with sterile distilled water. One unit of T4 DNA ligase (BRL) was added if the fragments to be ligated were blunt ended and 0.5 units were added if the ends were "sticky". The components were well mixed, sealed in a capillary tube and incubated overnight at room temperature. M13 vectors were not dephosphorylated before the ligations and only 0.07 µg of the vector was used for a ligation.

3.10.5 Transformation of *E. coli*

A 10 mL mid-log phase culture of *E. coli* grown in LB broth was

placed on ice for 15 min before collecting the cells by centrifugation for 5 min at 12 000 x g. The pellet was resuspended in 10 mL of 0.1 M $CaCl_2$ and placed on ice for 20 min. The cells were once again pelleted by centrifugation at 12 000 x g for 5 min and then gently resuspended in 0.5 mL of 0.1 M $CaCl_2$. The competent cells were stored ON at 4°C and the following day 0.2 mL aliquots were mixed with the ligation mixtures and placed on ice for 1 h. The cells were then placed in a 42°C waterbath for 90 sec, diluted with 1.0 mL of LB broth, and incubated at 37°C for 1 to 2 h to allow expression of the plasmid encoded proteins before aliquots of 0.05 to 0.2 mL were plated on LB selection plates.

3.10.6 Screening for Recombinant Plasmids

3.10.6.1 pAT153 Recombinants

The 3.6 kb plasmid vector pAT153 has genes conferring resistance to both tetracycline and ampicillin. Fragments were cloned into the unique *Hin*dIII and *Sph*I sites found in the pAT153 Tc^R gene. Transformed cells were plated on LBA plates which allowed both recombinants and nonrecombinants to grow. The colonies were screened for plasmids with inserts by picking them onto LBT plates. Cells unable to grow on the LBT plate probably had inserts in the tetracycline gene and DNA was then isolated and characterized by restriction enzyme digestion.

3.10.6.2 pKK232-8 Recombinants

The promoter cloning vector pKK232-8 carries the *bla* gene conferring ampicillin resistance as well as a promoterless chloramphenicol acetyltransferase gene (Brosius, 1984). *Bgl*II fragments were cloned into the unique *Bam*HI site, part of a multiple cloning site preceding the chloramphenicol acetyltransferase gene. Plasmid-containing colonies were first selected with LBA plates and then

streaked onto increasing concentrations of chloramphenicol to identify promoter containing recombinants. DNA from clones exhibiting the highest resistance to chloramphenicol was prepared and characterized by restriction enzyme digestion.

3.10.6.3 M13mp18/19 Recombinants

The M13mp18/19 phage are commonly used for subcloning fragments for dideoxy sequencing because they secrete ss DNA and they allow the single step identification of recombinant clones. The host strain into which these phage were transformed was Lac^- but became Lac^+ if it was complemented by the M13 vector. The Lac^+ nonrecombinant phage produced blue plaques on Xgal and IPTG while recombinant M13 phage had a fragment interrupting the β-galactosidase gene and could no longer complement the host, resulting in clear plaques.

3.10.7 Elution of DNA Fragments from Agarose Gels

3.10.7.1 Elution of DNA Fragments for Restriction Enzyme Digestion

The DNA band of interest was cut out from an EtBr stained agarose gel under indirect UV illumination taking care not to remove excess agarose with the DNA band. The DNA was then eluted as described by Selker et al (1976). The gel was suspended as thoroughly as possible in 10 mM Tris-HCl pH 8.0, 2 mM EDTA, and 1M NaCl by repeatedly forcing it through a 20-gauge needle. After incubating the suspension for 16 h at 47°C, the gel was sedimented by centrifugation for 30 min at 18 000 x g, and the supernatant was collected. Following a phenol extraction, the DNA was precipitated with three volumes of ethanol, stored at -60°C for 30 min, then collected by centrifugation at 12 000 x g for 20 min. The DNA was then dried and used for restriction enzyme digestion.

3.10.7.2 Elution of DNA Fragments for Nick Translation or Cloning

The DNA band of interest was cut out from an EtBr stained gel as described in 3.10.7.1 and the DNA was eluted using a GENECLEAN Kit, following the instructions and using the solutions supplied with the kit (Bio101 Inc.). This procedure was based on data from Vogelstein and Gillespie (1979). The agarose was dissolved in a sodium iodide solution, the DNA was bound to a silica matrix (glassmilk), and the contaminating agarose was washed away with an ethanol-salt solution. The DNA was then released from the "glassmilk" by heating it in water. DNA of excellent quality was recovered with less than a 20% loss.

3.11 DNA/DNA Hybridizations

3.11.1 Preparation of Southern Blots

Materials

Denaturing Solution: 1.5 M NaCl
0.5 M NaOH

Neutralizing Solution: 1.0 M Tris-HCl pH 8.0
1.5 M NaCl

20 X SSC: 3.0 M NaCl
3.0 M sodium citrate

Southern blots were prepared as described by Maniatis et al.(1982). DNA samples to be blotted were run on 0.8% agarose gels, stained and photographed as outlined in section 3.10.2. Excess agarose was trimmed away and the gel was agitated in several volumes of denaturing solution for 1 h at room temperature. The denaturing solution was replaced with neutralizing solution and the agitation was continued for 1 h.

If only one blot was to be prepared from the gel, a piece of Whatman 3MM paper slightly wider than the gel was placed on top of a glass plate with the ends extending into a tray of 10 X SSC allowing it

to act as a wick through which buffer could move by capillary action. The gel was placed on top of the filter paper and framed with Saran Wrap to prevent the buffer from moving around rather than through the gel. A nitrocellulose filter (Millipore HAWP) that had been floated on 2 X SSC for 2-3 min was placed on top of the gel. Any bubbles were carefully removed before 2 or 3 Whatman 3MM filter papers approximately the same size as the gel were placed on top. A three inch stack of paper towels, and finally a 500 g weight was placed on top of this. The transfer was allowed to proceed for at least 16 h, then the nitro-cellulose filter was removed and soaked in 6 X SSC for 5 min. The filter was air dried and then baked for 2 h at 80°C.

Double blots were prepared in a similar way, except that the gel was sandwiched between two nitrocellulose filters, two sets of Whatman 3MM paper, and two bundles of paper towels. The buffer in the gel provided enough moisture for the transfer of the DNA in both directions.

3.11.2 Nick Translation

Materials

5 x Nick Translation Buffer: 0.1 mM dCTP (Pharmacia)
0.1 mM dGTP "
0.1 mM dTTP "
0.05 M $MgCl_2$
0.5 mM DTT (BMC)
0.25 M Tris-HCl pH 7.6
250 µg/mL BSA (BMC)
prepared in HPLC grade water (Fisher)

DNase I: 0.1 µg/mL prepared by dilution of a 1 mg/mL stock into 1 x nick translation buffer containing 50% glycerol.

Sephadex G-50 (Sigma): Sephadex was stored at 4°C after it was allowed to expand in several volumes of TE overnight at room temperature.

The nick translation procedure was modified slightly from Maniatis et al. (1982). A 44 µL reaction mixture containing 0.5 to 1.0 µg of DNA, 10 µL of 5 x nick translation buffer and 1 µL of

DNA Polymerase I (Pharmacia 7300 u/mL) was prepared. Five µL of dATP [α^{32}P] purchased from New England Nuclear (NEN) having a specific activity of 3000 µCi/mM and a concentration of 0.0033 µM/mL was added and then 1 µL of DNase I was added. The contents were thoroughly mixed, then placed at 16°C for 1-1.5 h.

The nick translation mixture was immediately loaded on a small (5 mL) Sephadex G-50 column to separate the unincorporated nucleotides from the labelled DNA. The peaks were monitored with a Geiger counter (Ludlum Measurements Inc.) and then a 1 µL aliquot was counted in a Beckman LS-230 liquid scintillation counter. The labelled DNA was then frozen at -20°C until required.

3.11.3 Hybridization (Maniatis et al., 1982)

Materials:

10 x SSC: 1.5 M NaCl
1.5 M sodium citrate

50 x Denhardt's Solution:
per 500 mL:
5.0 g Ficoll
5.0 g polyvinylpyrrolidone
5.0 g BSA (Pentax fraction V)

Salmon sperm DNA: Stock solution of 10 mg/mL were prepared in water, sheared by passage through a 20 gauge needle, then boiled for 5 min and stored at -20°C.

Prehybridization Fluid: 6 x SSC
0.5% SDS
5 x Denhardt's Solution
100 µg/mL salmon sperm DNA

Hybridization Fluid: Prehybridization fluid plus 0.01 M EDTA

The filter to be hybridized was first wet from underneath with 6 x SSC, then immersed in 6 x SSC for 2 min. The filter was sealed in a bag with 0.2 mL /cm^2 of prehybridization fluid and incubated with agitation at 68°C. After 4 h, the prehybridization fluid was replaced

with hybridization fluid. The ^{32}P-labelled DNA probe (3-20 x 10^7 cpm) was boiled for 5 min then added to the bag. The hybridization was agitated overnight at 68°C.

After 16 h or more, the hybridization solution was removed and the filter was immersed in 2 x SSC/0.5% SDS. The filter was gently agitated for 5 min at room temperature, then transferred to 2 x SSC/0.1% SDS and agitated for another 15 min at room temperature. The filters were then placed in 0.1 x SSC/0.5% SDS and agitated for 2 h at 68°C. The wash was replaced and the incubation was continued another 30 min. The filters were then dried and exposed to Kodak X-Omat-AR X-ray film in a Picker Source One cassette containing one Cronex Quantalll intensifying screen (Du Pont) and kept at -70°C for 8-72 h before developing the film. X-ray film was developed by placing it in developer for 3 min, then rinsing it well with water, and placing it in fixer for 3 min. The film was then well rinsed with water and dried.

3.12 Isolation of Plasmid Deletion Mutants

3.12.1 Deletion of Specific Fragments from Plasmids

When one or more restriction fragments surrounded by the same restriction enzyme site were to be deleted, 0.5 µg of the DNA was digested with the enzyme surrounding the fragment, the DNA was diluted to 50 µL, and ligated overnight at 16°C.

3.12.2 BAL31 Deletion Mutagenesis of pBT22

Nuclease BAL31's ability to degrade double stranded DNA from the ends inward makes it a useful tool for the definition of phenotypic boundaries. This principle was applied to the definition of one end of *katG* using a unique *Sal*I site located in a nonessential part of the

vector, near one end of the insert.

The BAL31 deletion mutagenesis was done using the method described by Frey et al. (1984). Ten µg of pBT22 DNA was digested with *Sal*I in a 40 µL volume and a 2 µL sample was run on a gel to verify that the restriction digest was complete. Additions of 0.1 volumes of 3M sodium acetate pH 4.8 and 0.7 volumes of isopropanol were made and the sample was placed at -20°C for at least 30 min. The precipitated DNA was collected with a 10 min centrifugation, the pellet was washed twice with 80% ethanol, and left to air dry.

The pellet was dissolved in 110 µL of BAL31 incubation buffer (20 mM Tris-HCl pH 8.0, 600 mM NaCl, 12 mM $MgCl_2$, 12 mM $CaCl_2$, 1 mM EDTA) and the sample was prewarmed at 30°C. At 0, 2, 5, 10, and 15 min, a 25 µL sample was taken, 5 µL of ice cold 0.2 M EDTA pH 8.0 and 80 µL of ice cold water were added, then the sample was placed on ice. When the time course was completed, each sample was precipitated in the same way as in the first part of the procedure. Each DNA sample was dissolved in 10 µL of TE buffer and 5 µL aliquots of the 0 and 15 min time samples were run on a gel to make sure that BAL31 was active. The remaining samples were ligated as described in section 3.10.4, except that they were kept at 14°C for 2-3 h and then diluted with 90 µL of ligation buffer before continuing the incubation overnight at 14°C. The dilution was to prevent intramolecular ligation.

The ligations were transformed into UM228 and plasmid containing cells were selected on LBA. Catalase was checked by the H_2O_2 drop test and colonies were screened for deletion plasmids by rapid plasmid screening prepartions. Plasmids of interest were further characterized by restriction enzyme digestion.

3.13 Maxicell Analysis of Plasmid Encoded Polypeptides

The maxicell procedure was carried out as described by Sancar et al. (1979), but with minor modifications. Plasmid bearing cells were grown to early mid-log phase (3×10^8 cells/mL) in K Medium containing the appropriate plasmid selecting antibiotic. The cells were then irradiated in a sterile Petri dish with a Mineralight 115V UV lamp at a distance of 36 cm. The irradiation time (earliest time at which 100% killing occurred) was assayed and found to be 12 sec for CSR603 and 30 sec for UM262. The irradiated cells were transferred to a sterile 125 mL Erlenmeyer flask and shaken at 37°C for 1 h. Cycloserine (100 mg/mL prepared in 0.1 M phosphate buffer pH 8.0) was added to a final concentration of 100 µg/mL and then the 37°C incubation was continued overnight. The following morning the cells were collected by centrifugation at 12 000 x g and washed twice in Hershey's Salts. The washed pellet was then resuspended in 5.0 mL of Hershey's Medium and 20 µCi of ^{14}C-labelled mixed amino acids (0.05 mCi/0.5 mL) purchased from NEN were added. The cells were then incubated at 37°C for 1.5 h with aeration to allow ^{14}C-labelled plasmid encoded polypeptides to be synthesized. The cells were harvested in a microcentrifuge tube by repeated 3 min spins in an Eppendorf Model 5412 centrifuge. After one wash in 0.5 mL of SM, the cells were evenly resuspended in 40 µL of sample buffer (2% SDS, 0.2 M MSH, 8% glycerol, 63 mM Tris pH 6.8) and frozen at -20°C until they were to be loaded on a gel.

Just before loading the SDS-PAGE gel, 15 µL of the sample was mixed with 2 µL of tracking dye (0.25% bromophenol blue) and boiled for 3 min. The samples were loaded, electrophoresed, stained, and

destained as described in Section 3.14.3.

Following destaining, the gel was prepared for fluorography using the procedure of Boulnois and Timmis (1984). The fixer solution was replaced with DMSO and agitated gently for 30 min. The DMSO was replaced with fresh DMSO and agitation was continued another 30 min. The gel was placed in 22% (w/v) PPO in DMSO, and after 1 h, the solution was removed and replaced with water. After replacing the water several times over the course of an hour, the gel was dried on filter paper.

The dried gel was exposed to x-ray film as described in section 3.11.3, except the cassette contained two intensifying screens, and the exposure was done for 5-16 days at -70°C.

3.14 HPI Protein Purification and Analysis

3.14.1 Partial Purification of Catalase HPI

Two-1 L LB cultures of the desired *E. coli* strain were grown overnight at 37°C on a rotary shaker. The cells were collected by centrifugation, weighed, and frozen at -60°C.

The frozen cells were thawed at 37°C and resuspended at 4°C in 400 mL of KPi (50 mM potassium phosphate pH 7.0) per 75 g of cells until a smooth slurry was obtained. The cells were then lysed by passing them through an Aminco French Pressure Cell Press at 20 000 psi. The lysate was centrifuged at 7000 x g for 15 min and the supernatant was collected. The preceding step was repeated and then 2.5% (w/v) streptomycin sulphate (BMC) was added and the slurry was allowed to stir for 50 min at 4°C. The precipitate was removed by centrifugation for 15 min at 7000 x g and the supernatant was fractionated by precipitation with 65% ammonium sulfate (Schwarz Mann). After stirring 45 min at 4°C,

the precipitate was collected by centrifugation for 15 min at 12 000 x g. The pellet was resuspended in a minimal volume of KPi and dialyzed overnight against 4 L of KPi. Five hundred mg of protein was loaded onto a 2.8 x 60 cm DEAE-Sephadex column and washed with KPi until the A_{260} fell below 0.01. Fifty mM NaCl was then run through the column and 75 drop fractions were collected. The protein peak was pooled and concentrated by an 80% ammonium sulfate precipitation. The pellet was resuspended in the smallest volume possible, dialyzed overnight against KPi, centrifuged to clarify, and stored at -20°C.

If catalase was active, it was assayed after each step.

3.14.2 Immunodiffusion Analysis

Immunodiffusion plates containing 0.9% (w/v) NaCl and 1.0% (w/v) agar were prepared. HPI-antisera prepared in a rabbit (Loewen et al., 1985b) was placed in the centre well with the test samples surrounding it and diffusion was allowed to proceed overnight at 37°C.

3.14.3 SDS-Polyacrylamide Gel Electrophoresis (PAGE)

Gels for SDS-PAGE were prepared and electrophoresed with the Bio-Rad Protean Dual Slab Cell apparatus described in section 3.10.2.3 for polyacrylamide DNA electrophoresis. An 8% polyacrylamide, 0.1% SDS running gel was prepared in Tris-HCl pH 8.8 and after polymerization it was overlayered with a 4% polyacrylamide, 0.1% SDS stacking gel prepared in Tris-HCl pH 6.8.

Samples to be run on the gel were mixed with an equal volume of sample buffer (see section 3.13 for contents), boiled for 3 min, mixed 1:10 with 0.25% bromophenol blue, and loaded on the gel. The gels were run in 0.1% SDS, 25 mM Tris, 0.2 M glycine at 25 mA for about 3 h. The gel was then removed from the apparatus and stained overnight in Coomassie

Brilliant Blue (0.25% in 10% acetic acid and 40% methanol). The stain was replaced the following day with destain, containing 7% methanol and 15% acetic acid. The destain was repeatedly replaced until the protein bands appeared as dark blue bands against a light blue background. The gel was then transferred to filter paper and dried under vacuum for 2-3 h.

Molecular weight standards were purchased from Sigma. The ln of the known protein molecular weights was plotted as a function of the distance the proteins had travelled from the origin. Unknown protein molecular weights were then calculated by a comparison of the distances migrated with the standard curve.

3.14.5 Cyanogen Bromide Cleavage of HPI Catalase

HPI catalase was purified by J. Switala using the method previously described by Loewen and Switala (1986). Cyanogen bromide peptides were generated from catalase following the method of Steers et al. (1965). Forty mg of protein was lyophilized, then resuspended at a concentration of 15 mg/mL in 70% formic acid. Cyanogen bromide was added to a concentration of 18 mg/mL and the sample was allowed to react overnight at room temperature in the dark. Five volumes of H_2O were added and the sample was lyophilized. The peptides were citraconallated as described by Atassi and Habeeb (1972). The sample was resuspended in 9.0 mL of 6 M guanidine-HCl (BRL), 0.2 M N-ethylmorpholine acetate pH 8.5 and 300 µL of citraconic anhydride was added in 50 µL aliquots while maintaining the pH at 8-9. After stirring for 2 h, the sample was dialyzed against 0.1 M NH_4HCO_3, lyophilized, and resuspended in 0.5 mL of H_2O. The sample was then loaded on a Sephadex G-75-40 column and the resulting peaks were further separated by HPLC. Protein from the peaks was lyophilized and stored at -20° until it was prepared

for amino acid sequencing.

3.14.6 Amino Acid Sequence Determination

The amino acid sequence determinations were done by Dr. H. Duckworth and W. Taylor in the Department of Chemistry, University of Manitoba. The cyanogen bromide generated peptides were sequenced in an Beckman Model 890C Sequencer. Samples were submitted to 8 to 21 cycles of Edman degradation and the amino acid derivatives in the fractions collected were identified by HPLC. Amino acid assignments were made by comparing the cycle of interest to the preceding cycle and looking for the appearance of new peaks. The elution time of the peak was then compared to the amino acid standards to identify the peak. In cases where more than one amino acid had almost the same elution time, the proper identification was made by running the amino acid standards and the unknown simultaneously.

3.15 Dideoxy Chain Termination Sequencing

3.15.1 Preparation of M13 Vectors

JM101 was prepared for M13 infection as described by Dr. L. Donald (personal communication). A flask containing 100 mL of 2xYT was inoculated with 1.0 mL of an overnight culture of JM101 and aerated on a rotary shaker bed at 37°C for 1.5 h. A 5.0 mL aliquot was then transferred to 500 mL of 2xYT and after incubation of this flask under the same conditions for 1.5 h, a 1.0 mL inoculum of M13 phage was added. (The M13 phage was prepared by inoculating a single plaque into 1.5 mL of log phase JM101 grown in 2xYT and incubating this overnight at 37°C with aeration.) The 500 mL culture was allowed to continue shaking at 37°C for 5-8 h before the cells were collected, washed in SM, and the RF

(replicative form) DNA was prepared as described for large scale plasmid preparation in section 3.8.2.2.

3.15.2 Cloning into M13 Vectors

Restriction enzyme digests and ligations were carried out as described in Section 3.10. JM101 was prepared for the transformation as described in section 3.10 with two exceptions: competent JM101 cells were always used immediately, and the heat shock was done at 37°C for 15 min. After the heat shock, the cells were mixed with 100 mM IPTG, 2% Xgal(in dimethyl formamide), and 0.2 mL of log phase JM101 grown in 2xYT. This mixture was then mixed with 2.5 mL of LB top agar and immediately plated on an LB plate. These plates were then inverted and incubated at 37°C for 8-20 h.

3.15.3 Preparation of SS Template

Materials

PEG/NaCl: 20% (w/v) polyethylene glycol 6000 (Fisher)
2.5 M NaCl
Store at 4°C

The recombinant M13 phage plaques, prepared as described in the preceding section were clear (sometimes referred to as white) but mixed with numerous wild type M13 blue plaques. A clear plaque was inoculated with a toothpick into a 50 mL tube containing 1.5 mL of log phase JM101 growing in 2xYT. This was incubated with vigorous shaking in a Dubnoff Metabolic Shaking Waterbath at 37°C for 7-8 h. The culture was transferred to a microcentrifuge tube and the cells were pelleted with a 5 min centrifugation. The supernatant was transferred to a clean tube and the centrifugation was repeated a second time. The cell pellet was retained for preparation of RF DNA, and the ss DNA was prepared as described by Schreier and Cortese (1979). The supernatant was transferred to a

tube containing 0.2 mL of cold PEG/NaCl, mixed, and kept at room temperature for 15 min. The phage pellet was collected by a 5 min centrifugation and the supernatant was aspirated off. A second 3 min centrifugation of the pellet was done and the remaining traces of PEG/NaCl were removed by aspiration. The phage pellet was resuspended in 100 µL of TEN_{250}, then mixed with 50 µL of phenol/chloroform (1:1). The phases were once again separated by centrifugation and the aqueous phase was transferred to a tube containing 50 µL of chloroform. After this extraction, 0.25 mL of ethanol was added to the aqueous phase, and the tube was kept overnight at -20°C. The DNA was collected by a 20 min centrifugation, dried at room temperature, and resuspended in 15-25 µL of TE or HPLC water.

3.15.4 Preparation of RF DNA

After removing the supernatant in section 3.15.3, RF DNA was prepared from the remaining pellet as described in section 3.8.2.1.

3.15.5 Preparation and Electrophoresis of Sequencing Gels

Materials

TBE Running Buffer (10 x Stock)
per liter:

108 g	Trizma base (Sigma)
55 g	boric acid (Fisher)
9.3 g	$Na_2EDTA.2H_2O$ (Sigma)

(gives a stock of pH 8.3)

40% Acrylamide Stock
per 100 mL:

38 g	acrylamide (Bio-Rad Laboratories)
2 g	NN'-methylene bisacrylamide (Sigma)

The solution was filtered and stored in the dark at room temperature.

5 x TBE Stock: Prepared by diluting the 10 x Stock

Sialyzing Solution: 2% dichloromethylsilane (Aldrich) (v/v) in carbon tetrachloride (MCB)

Ammonium Persulfate Solution: 100 mg of ammonium persulfate (Fisher) dissolved in 1.0 mL of

distilled water was freshly prepared each time.

Two 38.5 x 20 cm gel plates were thoroughly scrubbed with soap and water, then rinsed with distilled water. One plate was sialyzed on one side and thoroughly rinsed with distilled water. After both plates were clean and dry, two 0.05 cm spacers were placed between them and the plates were taped in place. For each gel

21 g urea (BRL-ultra pure)
10 mL 5 x TBE stock
7.5 mL 40% acrylamide stock
and distilled water to 50 mL

was dissolved at 37°C and then allowed to cool to room temperature before adding 50 µL TEMED (BRL-ultrapure) and 0.25 mL of ammonium persulfate solution. The acrylamide solution was slowly injected between the glass plates with a 50 mL syringe.

Samples were loaded into wells with BRL "sharkstooth" combs and the gels were electrophoresed at 1200V (constant voltage) using 1 x TBE buffer prepared by diluting the 10 x TBE stock. The first part of the sequence was obtained by ending the electrophoresis after the bromophenol blue from the tracking dye reached the bottom (about 1.5 h) and sequence beyond that which could be read on this gel was obtained by ending electrophoresis after the xylene cyanol reached the bottom (about 3 h). For sequence between 350-550 bp the samples were run from 6-11 h. Gels to be run for 6-11 h were prepared and run in 1 x TBE diluted from a 10 x TBE stock containing 162 g Tris, 27.5 g boric acid, and 9.5 g $Na_2EDTA.H_2O$ per liter.

3.15.6 Sequencing and T-Tracking SS Templates

Materials

Klenow Buffer (10 x): 100 mM Tris-HCL pH 8.0
50 mM $MgCl_2$

(prepared in deionized distilled H_2O)

Klenow Dilution Buffer: 10 mM Tris-HCl pH 8.0
(prepared in deionized distilled H_2O)

Dideoxynucleotide Solutions (Pharmacia sequencing grade): 10 mM stocks were prepared in HPLC grade H_2O and frozen at -60°C.

Working solutions were prepared by diluting the stocks in HPLC grade H_2O) to the following concentrations:

0.1 mM ddATP
0.1 mM ddCTP
0.3 mM ddGTP
0.5 mM ddTTP

Deoxynucleotide Solutions: 10 mM stocks were prepared in 5 mM Tris-HCL pH 7.4 and 0.1 mM EDTA. dNTP's were purchased from Pharmacia.

0.5 mM working solutions of dCTP, dGTP, and dTTP were prepared from the stocks in HPLC H_2O. The nucleotide mixes were prepared with the nucleotide to be sequenced limiting as follows.

Nucleotide to be sequenced	0.5 mM dCTP	0.5 mM dGTP	0.5 mM dTTP	50 mM Tris pH8
A^o	20 µL	20 µL	20 µL	5 µL
C^o	1 µL	20 µL	20 µL	5 µL
G^o	20 µL	1 µL	20 µL	5 µL
T^o	20 µL	20 µL	1 µL	5 µL

Chase: 0.5 mM dATP
0.5 mM dCTP
0.5 mM dGTP
0.5 mM dTTP
prepared in HPLC grade H_2O

[α^{32}P]dATP for Sequencing: High specific activity ^{32}P-labelled dATP was purchased from NEN. Before use, the label was diluted with 0.125 mM cold dATP (1 µL for every 4 µL of label) to give a specific activity of 300 Ci/nM.

M13 Primer: The 17 base long universal sequencing primer was purchased from Regional DNA Synthesis Laboratory, University of Calgary. The primer was dissolved in HPLC grade H_2O at a concentration of 2 ng/µL.

Formamide Dye: 10 mL of formamide was stirred gently with 0.5 g Amberlite™ MB1 for 30 min. The Amberlite was removed by filtration and the formamide was mixed with 0.03 g xylene cyanol FF, 0.03 g bromophenol blue, and 0.75 g $Na_2EDTA \cdot 2H_2O$.

3.15.6.1 Reactions for T-Tracking 10 Clones

To anneal the primer to the template, 2 µL of ss DNA was mixed with a 2 µL aliquot of primer mix (4 µL M13 primer, 6 µL 10 x Klenow buffer,

12 µL HPLC grade H_2O) and placed at 65°C for 5 min. The tubes were then allowed to cool for 20 min at room temperature. A T-reaction mixture was prepared from 8 µL ddTTP, 8 µL dTTP, 4 µL [α^{32}P]dATP, and 1 µLof Klenow (DNA Polymerase I Klenow Fragment, Pharmacia). Two µL was added to each tube of primed template. After a 15 min incubation at 42°C, 1.5 µL of chase was added to each tube and the incubation was continued for another 15 min. The tubes were removed from the waterbath, 2 µL of formamide dye was added, and the tubes were centrifuged a few seconds in a Beckman Microfuge II. The tubes were then boiled 3 min and 2.5 to 3.5 µL was loaded on the gel.

3.15.6.2 Reactions for Sequencing SS Templates

For samples to be sequenced 7 µL of DNA was mixed with 1.5 µL of 10 x Klenow buffer, and 2 µL of primer. The samples were placed at 65°C for 5 min and then allowed to cool at room temperature for 20 min. For each clone to be sequenced A, C, G, and T reaction tubes containing 1 µL Of the appropriate ddNTP, 1 µL of the corresponding dNTP solution and 0.5 mL of [α^{32}P]dATP were prepared (ie: 1 µL ddATP and 1 µL A^o solution= A reaction). Two µL of the annealed primer-template was added to each of the four reaction tubes. Two µL of diluted Klenow. (1 µL in 24 µL of Klenow dilution buffer) was then added to each tube. After a 15 min incubation at 42°C, 2 µL of chase was added and the incubation was continued for 30 min. Four µL of formamide dye was added, the tubes were centrifuged, and then boiled 3 min before loading the gel.

3.15.7 Preparation for and Autoradiography of Sequencing Gels

After the samples had been electrophoresed as far as required, the taped plates were removed from the apparatus. The tape was removed and

the gel plates were carefully pried apart with a spatula, leaving the gel on the non-sialyzed plate surface. The gel was then transferred on the plate to a 5% methanol/5% acetic acid bath and soaked for at least 20 min. The gel was then blotted with Kim Wipes to remove excess liquid and transferred to Whatman 3MM paper. The partially dried gel readily stuck to the paper and peeled away from the glass plate. Using a Bio-Rad Model 483 Slab Dryer, the gel was dried under vacuum for 1 h at 80°C.

The dried gel was then placed on Kodak X-Omat AR X-ray film and exposure was allowed to proceed for 18 h at room temperature before developing the film as described in section 3.11.3.

3.15.8 Interpretation of the DNA Sequence

Gels were read on a light box keeping the following rules in mind. In sequences of more than one C residue, the first C is much weaker in intensity than the second C. Runs of G's give a similar pattern to C's when preceded by a T residue. In runs of consecutive A residues, the first A will be the most intense with the rest of equal or gradually decreasing intensity. Runs of T residues following a C residue show a similar pattern. These and other hints on the interpretation of DNA sequence are described by Hindley and Staden (1983).

Analysis of the DNA sequence was done both manually and by computer. The computer software package, Microgenie (Beckman) was used to predict the amino acid sequence, calculate the protein molecular weight, determine the codon usage, determine the amino acid content, compare sequences to each other and with the data bank, and predict the hydrophilicity of the amino acid residues.

RESULTS

4.0 RESULTS

4.1 Genetic Mapping of Transposon Tn*10* Catalase Mutants

4.1.1 Isolation, Characterization, and Mapping of *katE12*::Tn*10* and *katF13*::Tn*10*

Transposon Tn*10* insertions in MP180 were isolated by spreading cells infected by λ561 on LBT plates and incubating them overnight at 42°C. Approximately 10 000 colonies were screened for catalase activity by the colony H_2O_2 drop test. Two colonies with lower catalase levels were picked and stored as UM120 and UM122. Both mutants had low stationary phase catalase levels when assayed on agar plates, but the stationary phase oxygraph values of UM120 and UM122 were similar to the parental strain (See Table 3 for oxygraph values). High stationary phase oxygraph values had been previously observed for the HPII mutant, UM181 by Loewen and Triggs (1984). The high stationary phase oxygraph values in the absence of HPII may result because the H_2O_2 concentrations become elevated in stationary phase, inducing HPI. On agar plates, the absence of aeration prevents the H_2O_2 from accumulating to a concentration high enough to induce HPI. UM120 and UM122 were ascorbate inducible like their parent, MP180 (see Table 4). These characteristics were typical of HPII mutants that had been previously isolated.

Cell free extracts from stationary phase cultures of UM120 and UM122 were electrophoresed on polyacrylamide gels, then stained for catalase and peroxidase activity. The results are illustrated in Figure 2. In both the UM120 and UM122 extracts, the slow running HPII band was absent, confirming that the mutations in these strains were

Table 3. Catalase levels in Tn*10* generated catalase mutants and their parents

Strain	Catalase Activity (U/mg [dry cell wt]) Log Phase 1.5 mM[a]	Log Phase 60 mM[b]	Stationary Phase 1.5 mM	Stationary Phase 60 mM
MP180	4.2	16.8	9.6	39.9
UM120 (*katE12*::Tn*10*)	3.7	15.5	7.8	24.6
UM122 (*katF13*::Tn*10*)	6.5	21.2	11.2	49.3
UM202 (*katG17*::Tn*10*)	ND[c]	3.2	2.7	22.8
UM178 (*katE1*)	1.5	NA[d]	NA	NA
UM196 (*katE1 katG17*::Tn*10*)	ND	NA	NA	NA
CSH57a	3.9	16.3	8.3	34.2
UM197 (*katG17*::Tn*10*)	ND	ND	ND	13.7

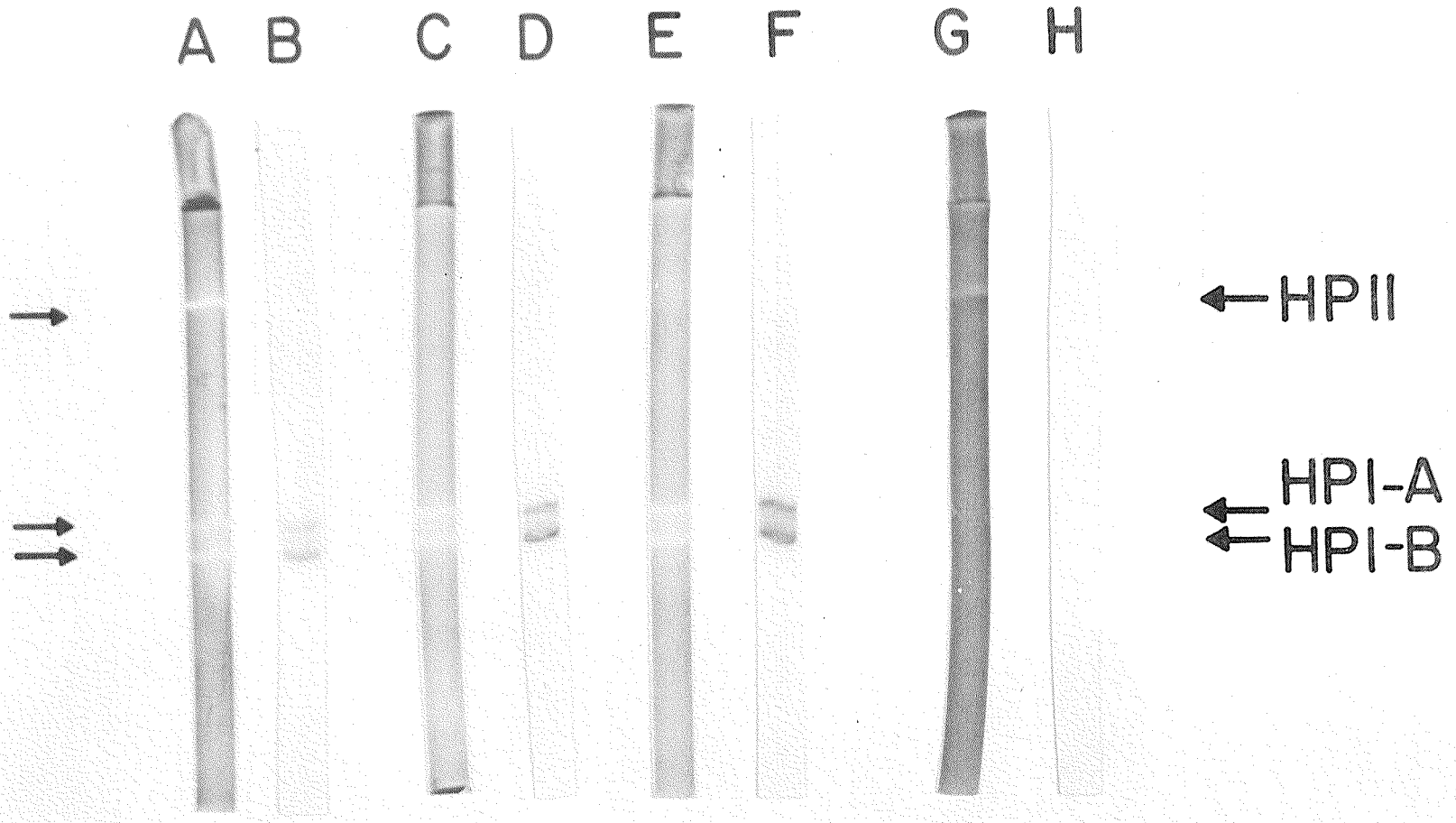
[a] concentration of H_2O_2 in assay was 1.5 mM
[b] concentration of H_2O_2 in assay was 60 mM
[c] not detectable
[d] not assayed

Table 4. Effect of ascorbate on catalase levels of various strains

Strain	Catalase (U/mg [dry cell wt])[a] at time (min)[b]					
	-10	-5	+10	+15	+20	+30
MP180	4.2	3.7	15.1		21.0	
UM120(*katE12*::Tn*10*)	3.5	3.7	12.3			27.3
UM122(*katF13*::Tn*10*)	4.9	4.9	15.2			17.1
UM178(*katE1*)	1.5	2.3		7.7		7.9
UM196(*katE1 katG17*::Tn*10*)	ND[c]	ND		ND		ND
CSH57a	2.5	2.4		6.1		8.6
UM197(*katG17*::Tn*10*)	0.5	ND		ND		ND

[a] H_2O_2 concentration in assay was 1.5 mM
[b] Ascorbate was added at 0 time
[c] not detectable

Figure 2. Visualization of catalase and peroxidase in crude extracts of MP180 and the transposon Tn*10* derivatives UM120, UM122, and UM202. Lanes A, C, E, and G were stained for catalase activity and lanes B, D, F, and H were stained for peroxidase activity. The following samples were run: A and B from MP180, 540 µg protein; C and D from UM120, 535 µg protein; E and F from UM122, 545 µg protein; G and H from UM202, 560 µg protein (Loewen et al., 1985a).


A
B
C
D
E
F
G
H
HPII
HPI-A
HPI-B

affecting the HPII catalase.

Prior to mapping the transposon insertions, the stability of the mutations was tested. No catalase positive revertants of either strain were found in 1000 Tc^R colonies. Attempts to select revertants out of the population in low levels of H_2O_2 or on Maloy Tc^S selection plates were unsuccessful. The mutations were stable and therefore suitable for mapping by interrupted mating and P1 transduction. Transposons have been shown to be suitable for the mapping and cloning of genes in several organisms (DeBruijn and Lupski, 1984).

The general locations of the Tn*10* insertions were found by interrupted mating, then specifically located by P1 transduction. Interrupted matings between UM120 or UM122 and the recipient, CSH57a (*rpsL*) were done and samples were spread on LBST plates to select *Tn*10 containing (Tc^R) recipient cells. The time of entry data were plotted (Figure 3) and the Tn*10* in UM120 was found to enter at approximately 39 min while the Tn*10* of UM122 entered at about 63 min. Since the Tn*10* of UM122 entered relatively late, the precision of the time of entry data dropped significantly (Miller, 1972). The loss of precision is evident from the slower rate of increase of Tc^R recombinants during the mating. These loci appeared to be the same loci that had been identified by NG mutagenesis and mapped as *katE* at 37.8 min and *katF* at 59 min (Loewen, 1984; Loewen and Triggs, 1984). P1 transductions were then used to look for the linkage of the Tn*10* (Tc^R) mutations with loci previously found to cotransduce with *katE* or *katF*. The *kat12*::Tn*10* (UM120) locus cotransduced with the temperature sensitive locus, *pheS*, 62% of the time (see Table 5), which converted to a 0.3 min separation, agreeing well

Figure 3. Determination of time of entry for *katE12*::Tn*10* and *katF13*::Tn*10*. Tc^R recombinants from the UM120/CSH57a (□), and the UM122/CSH57a (o) matings were enumerated by taking samples at various time intervals, vortexing them, and plating aliquots on LBST plates. The best straight line through each set of data points above the background level was determined by linear regression. The line was extrapolated to the X axis to give the approximate time of entry for the two genes, *katE* and *katF*.

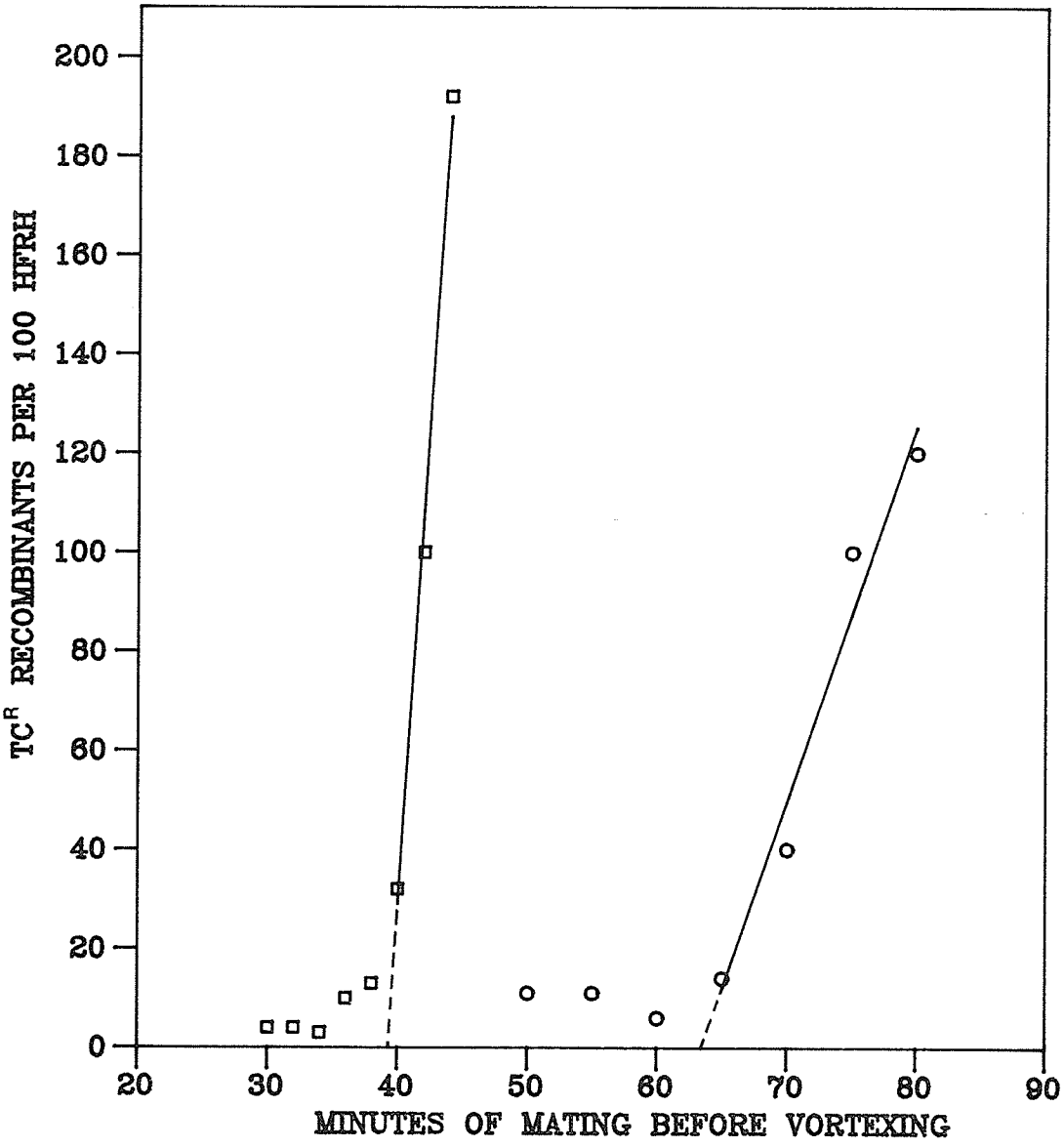

200
180
160
140
120
100
80
60
40
20
0
TC^R RECOMBINANTS PER 100 HFRH
20
30
40
50
60
70
80
90
MINUTES OF MATING BEFORE VORTEXING

Table 5. Mapping of *katE12*::Tn*10* and *katF13*::Tn*10* by P1 transduction

Donor	Recipient	Selected Marker	Unselected Marker Class	No.	(%)	min[a]
UM120 (*katE12*::Tn*10*)	4913 (*pheS*)	Tc^R	*pheS*$^+$	148	(62)	0.3
			pheS	92	(38)	
UM122 (*katF13*::Tn*10*)	JF627 (*pyrG*)	Tc^R	*pyrG*$^+$	21	(11)	1.0
			pyrG	163	(89)	
UM122 (*katF13*::Tn*10*)	JF627 (*pyrG*)	*pyrG*$^+$	Tc^R	6	(3)	1.4
			Tc^S	234	(97)	
UM122 (*katF13*::Tn*10*)	JM96 (*cysH*)	Tc^R	*cysH*$^+$	185	(51)	0.4
			cysH	175	(49)	
UM122 (*katF13*::Tn*10*)	JM96 (*cysH*)	*cysH*$^+$	Tc^R	95	(26)	0.7
			Tc^S	265	(74)	

[a] separation of markers in minutes as calculated from the cotransduction frequencies by the method of Wu (1966) (Loewen and Triggs, 1984).

with other mapping data for *katE* (Loewen, 1984). The *kat13*::Tn*10* (UM122) mutation cotransduced with *pyrG* and *cysH* at average frequencies of 7% (1.2 min) and 38% (0.55 min) respectively (see Table 5), mapping *katF13*::Tn*10* to approximately 58.9 min. The data for *katF13*::Tn*10* fit well with other data mapping *katF* (Loewen and Triggs, 1984). The general locations of *katE* and *katF* on the *E. coli* chromosome are illustrated in Figure 4.

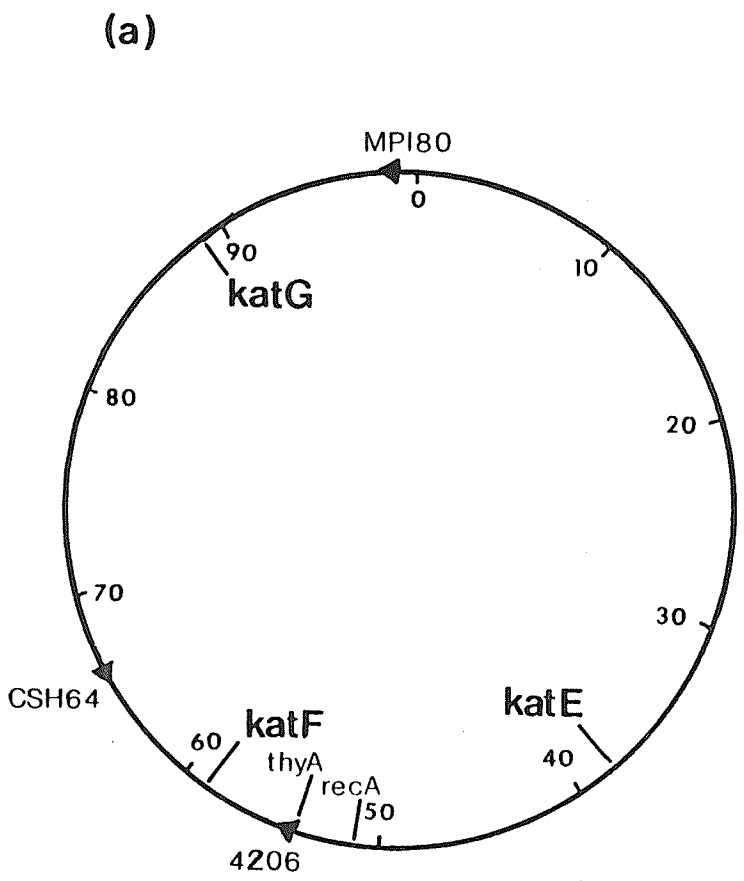
4.1.2 Isolation, Characterization, and Mapping of *katG17*::Tn*10*

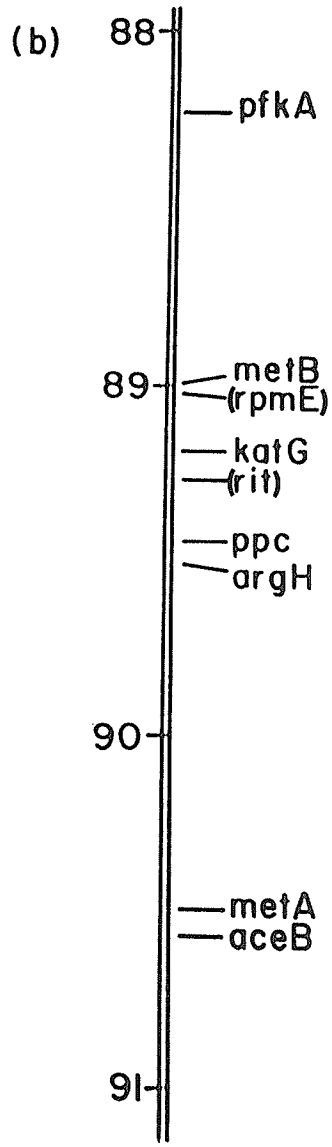
katG mutants could not be isolated directly in MP180 because the catalase levels of *katG* mutants on agar plates are almost the same as wild type levels. Consequently, a *katE katG*$^+$ strain, UM178, was mutagenized with Tn*10*. After screening more than 12 000 colonies, a Tn*10* mutant lower in catalase (*kat17*::Tn*10*) was isolated and stored as UM196. To allow certain experiments to be performed in *katE*$^+$ strains the mutation was transferred into MP180 and CSH57a by P1 transduction and these strains were stored as UM202 and UM197 respectively.

UM196 (*katE kat17*::Tn*10*) had almost no catalase activity on agar plates and no assayable catalase in the oxygraph. UM202, which was *katE*$^+$*kat17*::Tn*10*, had normal catalase levels on the agar plate but little assayable catalase during log phase. During stationary phase catalase was present (see Table 3). Strains carrying the *kat17*::Tn*10* mutation were not inducible with ascorbate (see Table 4). The properties of these mutants led to the understanding of the independent induction of catalases HPI and HPII (Loewen et al., 1985a).

Cell free extracts of stationary phase UM202 were electrophoresed and stained for catalase and peroxidase activities. As shown in Figure 2 this mutant lacked the faster running HPI A and B doublet in gels

Figure 4. a) Genetic map of *E. coli* K12 based on 100 min. The locations of *katE*, *katF*, and *katG* as well as some other relevant markers are shown. The arrows indicate the origin and direction of transfer of the Hfr strains used in this study (Bachmann, 1983).
b) Genetic map orienting *katG* relative to other nearby markers. *metB* acted as a reference point for the placement of *katG* (Loewen et al., 1985b).


(a)
MPI80
0
10
20
30
40
50
60
70
80
90
katG
katE
katF
thyA
recA
CSH64
4206


(b)
88
89
90
91
pfkA
metB
(rpmE)
katG
(rit)
ppc
argH
metA
aceB

stained for both catalase and peroxidase.

The characteristics of the *kat17*::Tn*10* mutation (little assayable catalase during log phase growth and no catalase induction by ascorbate) suggested that it was a mutant of the *katG* gene which mapped around 90 min. P1 transductions using UM196 as the *kat17*::Tn*10* donor showed that this locus was *katG*, and these data, in combination with mapping data from NG mutants (Loewen et al., 1985b), mapped *katG* to 89.2 min on the *E. coli* chromosome (see cotransduction frequencies and equivalent map units in Table 6). The linkage between *argH* and the Tn*10* was found to vary depending on the selected marker, Tc^R or *argH*$^+$. A second set of transductions was done with another *argH* strain, PA309 but the frequencies were still fairly inconsistent, suggesting that this was somehow a consequence of this particular selection system. For instance, one sequence may recombine more frequently than the other. The general location of the gene on the chromosome is shown in Figure 4 and the more precise location of *katG* is illustrated beside it. The cotransduction frequencies placed the genes in the order *metB-katG-ppc-argH*. The other markers, *rpmE*, coding for a ribosomal protein, and *rit*, coding for a protein that affects the thermolability of the 50S ribosomal subunit, were placed by comparing our data to those for *rpmE* (Dabbs, 1981) and *rit* (Ono, 1978) (Loewen et al., 1985b).

4.2 Isolation of Transformable Catalase Mutants

4.2.1 Isolation of Catalase Mutants by NG Mutagenesis

Catalase mutants had been previously isolated by NG mutagenesis of CSH7 and CSH57a (Loewen, 1984; Loewen and Triggs, 1985; Loewen et al., 1985b), but these strains had very low transformation efficiencies.

Table 6. Mapping of *katG17*::Tn*10* by P1 transduction

Donor	Recipient	Selected Marker	Unselected Marker Class	No.	(%)	min[a]
UM196 (*katG17*::Tn*10*)	5128 (*metB*)	Tc^R	*metB*$^+$	188	(78)	0.16
			metB	52	(22)	
UM196 (*katG17*::Tn*10*)	5128 (*metB*)	*metB*$^+$	Tc^R	196	(82)	0.13
			Tc^S	44	(18)	
UM196 (*katG17*::Tn*10*)	6374 (*argH*)	Tc^R	*argH*$^+$	67	(28)	0.7
			argH	170	(72)	
UM196 (*katG17*::Tn*10*)	6374 (*argH*)	*argH*$^+$	Tc^R	107	(45)	0.23
			Tc^S	133	(55)	
UM196 (*katG17*::Tn*10*)	5236 (*ppc*)	Tc^R	*ppc*$^+$	155	(65)	0.27
			ppc	85	(35)	
UM196 (*katG17*::Tn*10*)	5236 (*ppc*)	*ppc*$^+$	Tc^R	173	(72)	0.21
			Tc^S	67	(28)	
UM196 (*katG17*::Tn*10*)	PA309 (*argH*)	Tc^R	*argH*$^+$	24	(30)	0.66
			argH	55	(70)	
UM196 (*katG17*::Tn*10*)	PA309 (*argH*)	*argH*$^+$	Tc^R	117	(49)	0.42
			Tc^S	123	(51)	

[a] separation of markers in minutes was calculated from the cotransduction frequencies by the method of Wu (1966) (Loewen et al., 1985b).

HB101, a readily transformable *E. coli* K12 strain was mutagenized with NG and two mutants that were lower in catalase, UM68 and UM69, were isolated. Unfortunately, the background catalase levels were still so high that after transformation with the catalase carrying plasmid pLC36-19, pLC36-19 containing colonies could not be differentiated from non-plasmid bearing colonies. UM69 was further mutagenized with NG producing UM123, a strain with lower catalase than UM69 but still not completely negative. UM123 had variable catalase levels depending on the media content and often the background levels were too high to allow the differentiation of pLC36-19 containing colonies and non-plasmid harbouring colonies.

The NG mutagenesis of HB101 was repeated and three mutants lower in catalase, UM228, UM229, and UM230 were isolated. UM228 had the lowest catalase levels on agar plates although it had equal or higher catalase levels in the oxygraph (see Table 7). Because of its low catalase activity on agar plates, UM228 was chosen for the transformation of and screening for plasmid encoded catalase. UM228 was found to be readily transformable and colonies containing catalase bearing plasmids (pA30 and pE6) were obvious on the agar plate and in the oxygraph (see Table 7). A genetic characterization of the catalase mutations was not possible because the strain was *recA*, preventing the efficient recombination required for mapping. Extracts of UM228 contained only HPI which was still ascorbate inducible, (see Table 8), suggesting that the regulation of HPI had not been affected.

4.2.2 Construction of UM262

As noted above, all of the transformable catalase mutants available were *recA* which prevented mapping by interrupted mating or P1 trans-

Table 7. **Catalase activities of nitrosoguanidine mutants**

Strain	Agar Plate[a]	Catalase Activity (U/mg [dry cell wt]) 1.5 mM H_2O_2[b]	60 mM H_2O_2[c]
HB101	high	6.9	26.0
UM228	very low	7.6	20.1
UM229	low	6.2	17.7
UM230	low	6.6	20.0
UM228/pA30	high	15.8	84.6
UM228/pE6	high	15.0	84.6

[a] levels were determined by the colony H_2O_2 drop test
[b] the concentration of H_2O_2 in the assay was 1.5 mM
[c] the concentration of H_2O_2 in the assay was 60 mM

Table 8. **Effect of ascorbate on catalase levels of HB101 and UM228**

Strain	Catalase(U/mg [dry cell wt])[a] at time min[b] -10	-5	+15	+30
HB101	2.4	2.5	7.4	9.0
UM228	2.7	2.8	5.0	7.2

[a] H_2O_2 concentration in assay was 1.5 mM
[b] ascorbate was added at 0 min

duction. Furthermore, one NG mutant, UM123, was found to contain a suppressor of catalase activity, reinforcing the need for catalase mutants of known genotype to be used in the cloning of catalase genes.

The *recA*$^+$ parent of HB101, RR1, was used to construct transformable catalase mutants of known genotype. The *katG17*::Tn*10* locus from UM196 was transduced into RR1 by selecting for Tc^R colonies. The *katG17*::Tn*10* derivative of RR1 was plated on 10 µg/mL of trimethoprim to select out a *thyA* mutant which was stored as UM163. Cultures of the *thyA* mutant were always supplemented with 100 µg/mL thymine (Miller, 1972) to allow growth. UM163 was mated with 4206 (a *recA* Hfr with its origin of transfer at 55 min, [Miller, 1972]) and after 10 min of mating, samples were plated on M9 agar plates lacking thymine to select for *thyA*$^+$ cells. Since *recA* (51 min) was close to *thyA* (54.8 min) (see Figure 4), a large proportion of the *thyA*$^+$ recombinants were also *recA*. Cells were then streaked in duplicate onto LB plates and one was exposed to a UV light (Mineralight 115V) for 15 sec at a 15 cm distance to screen for *recA* cells. After an overnight incubation in the dark, cells that grew poorly were picked and rechecked in the same way but with a *recA*$^+$ control. A strain that was *recA* (grew poorly or not at all after the UV irradiation) and *katG17*::Tn*10* was stored as UM262. Other mutant markers of RR1 were checked after each step to make sure the strain retained the original genotype.

4.3 Restriction Mapping of pLC36-19

As described in section 2.11, the Clarke and Carbon plasmid, pLC36-19 carries the structural gene for HPI. This 25.5 kb plasmid contained a 19.2 kb insert of *E. coli* genomic DNA in a 6.3 kb ColE1

vector (Loewen et al., 1983). The plasmid was mapped with single and double restriction enzyme digests. The number of restriction sites in a single enzyme digest equalled the number of fragments generated, and in a double digest, the total number of restriction enzyme sites or fragments was equal to the sum of the number of restriction enzyme sites or fragments from the two single enzyme digests.

The sizes of the restriction enzyme fragments generated from pLC36-19 are given in Table 9. These fragments were used to map pLC36-19 by comparing the fragments generated in double digests with the fragments generated from a single enzyme digest. Fragments from a single enzyme digest which disappeared in the double digest must be cut by the second enzyme. The resulting fragments could be identified by adding a combination of the new fragments up to the same size as the original fragment. In cases where two fragments of the same size were generated from a digest, the double band was identified because it appeared brighter than adjacent bands. Fragments larger than 600 bp were usually sized on agarose gels and fragments smaller than 600 bp were normally sized on acrylamide gels. Fragments larger than the largest standard band from *Hin*dIII restricted λ DNA (23.7 kb) could not be sized accurately and are represented by >23 700 bp. Subsequent restriction maps in this thesis were generated in a similar manner. Fragments believed to originate from the vector were confirmed by restricting ColE1 DNA (see Table 10).

This approach was used to map sites for 14 restriction enzymes in pCL36-19 for a total of 57 sites. The final map is shown in Figure 5. The locations of some of the small *Pvu*II fragments were found by digesting DNA isolated from the two largest *Ava*I bands which were

Table 9. Restriction enzyme fragments obtained by digestion of pLC36-19

Restriction Enzyme	Fragment Size (bp)	Sum (bp)
*Cla*I	9870, 8580, 7080	25 530
*Ava*I	12 600, 10 040, 2730	25 370
*Bam*HI	>23 700	-
*Ba*II	16 010, 9220, 170, 160	25 560
*Eco*RI	>23 700, 380	-
*Hin*cII	7900, 7050, 5200, 2100, 1800, 1350, 340, 100, 90	25 930
*Hin*dIII	21 920, 3640	25 560
*Hpa*I	14 500, 8950, 2100, (100)[a]	25 650
*Pst*I	17 260, 7010, 1080, 410	25 760
*Pvu*II	7860, 6150, 3600, 3460, 1170, 1150, 650, 480, 380, 370	25 270
*Sma*I	12 700, 12 700	25 400
*Sa*II	>23 700	-
*Eco*RV	8100, 6060, 4440, 3360, 1370, 920, 730, 710	25 690
*Sph*I	10 400, 7550, 6190, 1380, 350, 200, 100	26 170
*Eco*R1-*Pst*I	12 630, 7010, 4530, 1080, 410, 380	26 040
*Eco*R1-*Hin*dIII	21 920, 3450, 260	25 630
*Eco*R1-*Bam*HI	>23 700, 760, 380	-
*Eco*R1-*Sma*I	12 700, 6350, 6040, 380	25 470
*Hin*dIII-*Pst*I	12 210, 7010, 3640, 1080, 1150, 410	25 500
*Hin*dIII-*Sa*II	21 000, 3640, 1010	25 650
*Hin*dIII-*Bam*HI	21 920, 2550, 1050	25 520
*Hpa*I-*Pst*I	12 800, 5400, 2100, 1900, 1550, 1080, 410, (100)	25 340
*Hpa*I-*Sma*I	7580, 7080, 5370, 3180, 2100, (100)	25 410
*Hpa*I-*Hin*dIII	13 950, 8080, 2100, 950, 750, (100)	25 930
*Hin*cII-*Bam*HI	7900, 7050, 5200, 1800, 1650, 1350, 340, 300, (100, 90)	25 780
*Hin*cII-*Eco*RI	7900, 7050, 5200, 2100, 1800, 1060, 330, 300, (100)	25 840
*Hin*cII-*Pst*I	6210, 5500, 5200, 2100, 1800, 1550, 1350, 1080, 410, 340, (100, 90)	25 730
*Hin*cII-*Ba*II	7900, 5200, 4700, 2350, 2100, 1800, 1270, 160, (100, 90)	25 670
*Ba*II-*Sma*I	10 100, 6560, 5600, 2910, 170, 160	25 500
*Ba*II-*Pst*I	12 750, 4620, 3650, 3150, 1080, 410, 170, 160	25 990
*Ba*II-*Hin*dIII	15 900, 6010, 3300, 170, 160	25 540
*Bam*HI-*Pst*I	13 550, 7010, 3750, 1080, 410	25 800
*Ava*I-*Pst*I	11 020, 7010, 3120, 2730, 1080, 410	25 370
*Ava*I-*Ba*II	7320, 6820, 5870, 3120, 2730, 170, 160	26 190
*Pvu*II-*Pst*I	7860, 6150, 3050, 1640, 1610, 1150, 1100, 1080, 650, 480, 440, 380, 370	25 960
*Pvu*II-*Sma*I	7860, 3460, 3380, 2660, 1820, 1760, 1170, 1150, 650, 480, 380, 370	25 140
*Pvu*II-*Bam*HI	6250, 4650, 3600, 3460, 3050, 1170, 1150, 650, 480, 380, 370	25 210
*Cla*I-*Hin*dIII	9870, 7080, 5040, 3130, 400	25 520

*Cla*I-*Hin*cII	5200, 4860, 4710, 3060, 2460, 2100, 1800, 1350, 240 (100, 90)	25 970
*Cla*I-*Sma*I	7080, 6910, 5740, 2680, 2640	25 050
*Cla*I-*Pst*I	9320, 4540, 3830, 2880, 2850, 1080, 410	24 910
*Eco*RV-*Ava*I	6060, 5370, 4440, 3360, 2730, 1370, 920, 730, 710	25 690
*Eco*RV-*Bam*HI	8100, 6060, 4440, 3360, 920, 830, 730, 710, 490	25 640
*Eco*RV-*Cla*I	8100, 4440, 3360, 3230, 2440, 1000, 920, 730, 710, 330	25 260
*Sph*I-*Ba*II	7550, 6190, 5830, 2440, 1380, 1150, (350, 200, 150, 140, 100)	25 480
*Sph*I-*Cla*I	7550, 6190, 3900, 2700, 2100, 1380, 1350, (350, 200, 100)	25 820
*Sph*I-*Eco*R1	10 400, 6190, 5540, 1440, 1380, (390, 350, 200, 100)	25 990
*Sph*I-*Hin*cII	7900, 4820, 4020, 1810, 1410, 1350, 990, 860, 820, 350, 340, 200, 100, 90, 90	25 150
*Sph*I-*Sa*II	7550, 7550, 6190, 2570, 1380, (350, 200, 100)	25 890

[a] brackets surround fragments too small to see on an agarose gel, but presumed to be present

(Loewen et al., 1983)

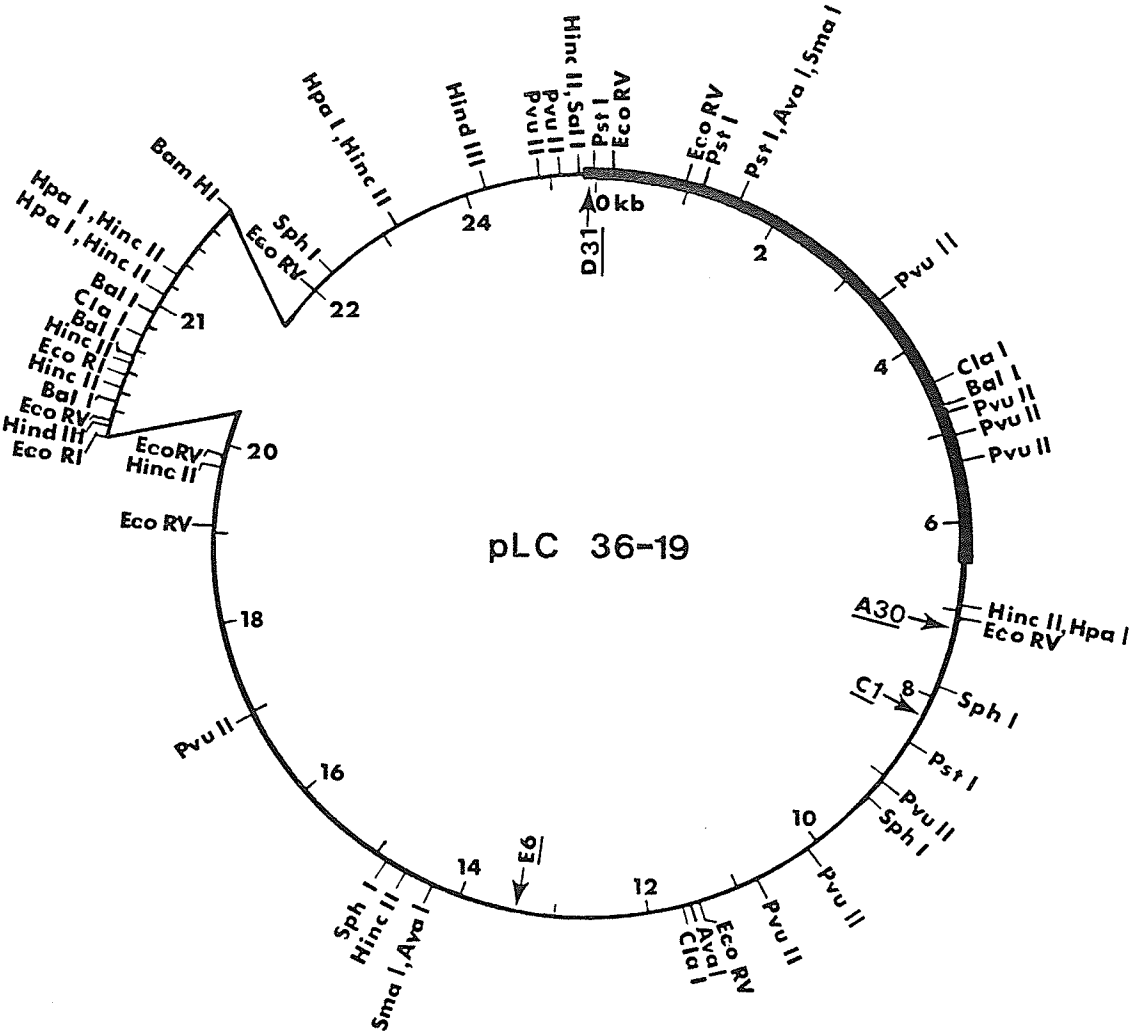
Table 10. **Restriction enzyme fragments obtained by digestion of ColE1 DNA**

Restriction Enzyme	Fragment Sizes (bp)	Sum(bp)
*Eco*RV	5650, 920	6570
*Pst*I	5010, 1080, 410	6500
*Pvu*II	4570, 1160, 380, 370	6480
*Eco*RV-*Cla*I	3230, 2440, 920	6590
*Eco*RV-*Ava*I	5650, 920	6570
*Eco*RV-*Pst*I	5220, 920, (410)[a]	6550
*Pvu*II-*Bal*I	4570, 1140, 380, 370	6460
*Pvu*II-*Cla*I	4570, 870, 380, 370, 260	6450
*Pvu*II-*Ava*I	2670, 1720, 1160, 380, 370	6320
*Ava*I-*Bal*I	3260, 3070	6330
*Ava*I-*Cla*I	3390, 2880	6270
*Pst*I-*Cla*I	2830, 2230, 1080, 410	6550

[a] size was estimated because the standards were inappropriate to size fragments this small

(Loewen et al., 1983)

Figure 5. Restriction map of pLC36-19. The thick black line represents the ColE1 vector and the thin black line represents the *E. coli* genomic insert DNA. The map was generated using the restriction enzyme fragments given in Tables 9 and 10. Three small *Sph*I fragments (350 bp, 200bp, and 100bp in size) have not been placed on the map because their positioning was questionable. The arrows on the inside of the circle point to the positions of the Tn*5* insertions in pLC36-19 generating the plasmids pA30, pC1, pE6, and pD31. The positions of the Tn*5* insertions were mapped using the restriction enzyme fragments given in Table 11 (Loewen et al., 1983; Triggs-Raine and Loewen, 1987).


pLC 36-19
0 kb
2
4
6
8
10
12
14
16
18
20
21
22
24
D31
A30
C1
E6
Pst I
Eco RV
Hinc II, Sal I
Pvu II
Hind III
Hpa I, Hinc II
Bam HI
Sph I
Eco RV
Pst I, Ava I, Sma I
Pvu II
Cla I
Bal I
Pvu II
Pvu II
Pvu II
Hinc II, Hpa I
Eco RV
Sph I
Pst I
Pvu II
Sph I
Pvu II
Pvu II
Eco RV
Ava I
Cla I
Sph I
Hinc II
Sma I, Ava I
Pvu II
Eco RV
Hinc II
Eco RV
Hpa I, Hinc II
Hpa I, Hinc II
Bal I
Cla I
Bal I
Hinc II
Eco RI
Hinc II
Bal I
Eco RV
Hind III
Eco RI

isolated from an agarose gel. Of the nucleases studied, *Bam*HI, *Sal*I, *Pst*I, and *Sma*I cut only once, and *Bgl*I, *Xba*I, *Kpn*I, and *Bcl*I did not cut at all. Three small *Sph*I restriction fragments could not be unequivocally placed because no restriction sites were found within them. These fragments have not been included in the pLC36-19 map, although the best fit of the *Sph*I fragments with the other restriction sites placed them in the 9 to 10 kb region of pLC36-19.

4.4 Isolation and Mapping of Transposon Tn*5* Insertions in pLC36-19

The *E.coli* genomic DNA insert in pLC36-19 was 19.2 kb in size but only 2.5 kb was required to code for an 84 kd protein, the size of the HPI subunit. To try to localize the gene encoding HPI on pLC36-19, transposon Tn*5* insertions were made in the plasmid.

Small scale alkaline-SDS plasmid preparations for restriction enzyme digestion were done on approximately 220 potential pLC36-19::Tn*5* containing colonies. The quality of the DNA from rapid plasmid preparations on these large plasmids was poor and it could only be used for comparing the size of the plasmids. Large scale plasmid preparations were done on plasmids which appeared larger than pLC36-19 so that they could be restricted to look for the presence of new bands indicative of the presence of Tn*5*. Seventeen pLC36-19 plasmids containing Tn*5* insertions were identified, but of these, 14 were insertions in almost the same position. The Tn*5* insertions in four different insertion plasmids, pA30, pE6, pC1, and pD31 were mapped by restriction enzyme digestion with *Hin*dIII, *Ava*I, and *Pst*I. *Pst*I digestions of all the other insertion plasmids were almost identical to pA30 suggesting that this area was a "hotspot" for Tn*5* insertion. This was unexpected as Tn*5* has been shown

at the phenotypic level to have low insertional specificity. However, Tn*5* has been shown to insert with higher frequencies into DNA sequences which have some homology to the Tn*5* inverted repeat (De Bruijn and Lupski, 1984). No Tn*5* insertions were found within the *katG* gene itself which has since been shown to have no obvious homologies with Tn*5*. The restriction fragment sizes for the 4 pLC36-19::Tn*5* plasmids are in Table 11. None of the Tn*5* insertions in pLC36-19 inactivated the catalase gene as is evident from the oxygraph values in Table 16 of the UM53 plasmid transformants. All of the plasmids complemented the catalase deficiency of UM53. The approximate locations of the mapped insertions are shown in Figure 5. A more detailed map of the plasmid pC1 is shown in Figure 6.

4.5 Localization of the HPI Coding Region by Deletion of pC1

Although the transposon Tn*5* insertions in pLC36-19 did not inactivate the catalase gene, the new sites provided by the transposon were exploited for deletion mutagenesis to localize the catalase gene. One of the Tn*5* containing plasmids, pC1, was restricted with *Hin*dIII, religated, transformed into UM123, and selected on LBK plates. Because they had been selected on kanamycin, all of the resulting plasmids had to carry at least the *Hin*dIII fragment from the transposon which contained the kanamycin resistance gene. All of the colonies appeared to have the same catalase levels on the agar plate.

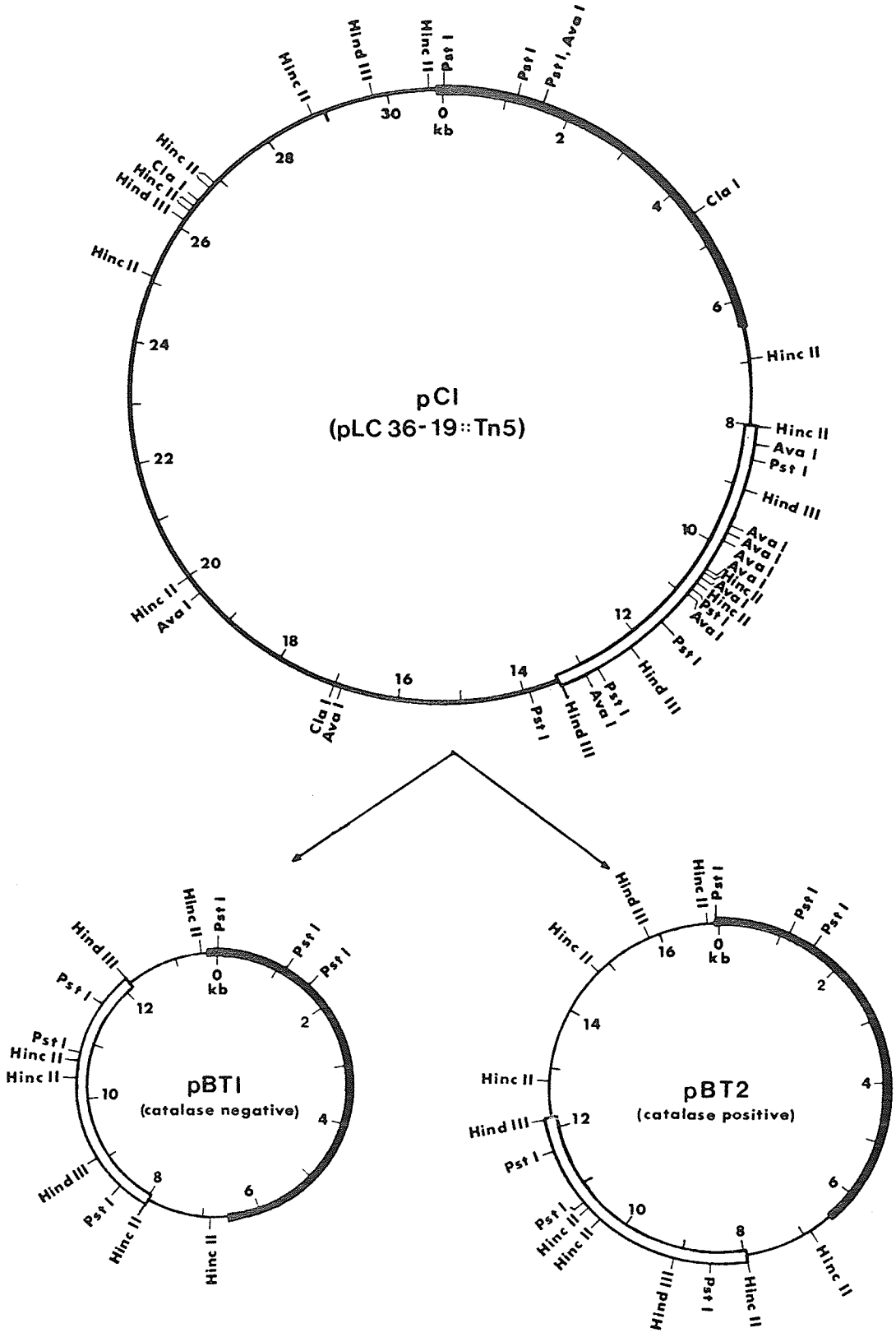
Eighteen colonies were grown up and plasmids were prepared by the alkaline-SDS small scale procedure. Fifteen of the eighteen clones were smaller than pLC36-19, and two of these (labelled pBT1 and pBT2) were chosen for further characterization. Aliquots of large scale plasmid preparations of pBT1 and pBT2 were restricted by *Hin*dIII and the

Table 11. Restriction enzyme fragments used to map Tn*5* insertions in pLC36-19

Plasmid	Restriction Enzyme	Fragment sizes (bp)	Sum (bp)
pLC36-19	*Aval*	12 600, 10 040, 2730	25 370
pLC36-19	*Hin*dIII	21 920, 3800	25 720
pLC36-19	*Pst*I	17 260, 7010, 1080, 410	25 760
pLC36-19	*Eco*RV	8100, 6060, 4440, 3960, 1370, 920, 730, 710	25 690
pC1	*Aval*	12 600, 6470, 4230, 2730, 1900, 1180, 980, 250, (250)[a]	30 340
pCI	*Hin*dIII	14 050, 9900, 3830, 3300	31 080
pCI	*Pst*I	17 260, 7010, 2480, 1140, 1080, 920, 790, 410	31 090
pA30	*Aval*	12 600, 5880, 5320, 2730, 2000, 1180, 670, 370	30 750
pA30	*Hin*dIII	14 000, 10 000, 3830, 3300	31 130
pA30	*Pst*I	17 260, 5900, 2230, 1900, 1080, 900, 800, 410	30 480
pA30	*Eco*RV	10 740, 8100, 4440, 3360, 1370, 920, 730, 710	30 370
pE6	*Aval*	12 600, 10 040, 2170, 2000, 1700, 1180 670, 370	30 730
pE6	*Hin*dIII	16 000, 8200, 3830, 3300	31 330
pE6	*Pst*I	13 500, 7010, 4920, 2470, 1080, 900, 850, 410	31 140
pD31	*Aval*	12 000, 10 040, 2770, 2000, 1270, 1150, 750, 400	30 380
pD31	*Hin*dIII	(20 700), 3830, 3300, 2770	30 600
pD31	*Pst*I	17 260, 7010, 2470, 1200, 1080, 900, 850, (410)	31 180

[a] the brackets indicate bands which either could not be seen on the gel but were known to be present or bands which could not be sized accurately on the gel

Figure 6. Restriction maps of pCI and the *Hin*dIII deletion derivatives pBT1 and pBT2. The thick black lines represent the ColEI vector, the thin black lines represent the insert DNA, and the open box represents the transposon Tn*5*. The *Ava*I and *Cla*I sites of pCI were not included in the pBTI or pBT2 maps although they are still presumed to be present. The 4 *Hin*cII sites between 26 and 27 kb on the pCI map are represented by a single *Hin*cII site at 12.8 kb on the pBT2 map.


pCI
(pLC 36-19::Tn5)
pBTI
(catalase negative)
pBT2
(catalase positive)
kb

restriction patterns differed in only one fragment: pBT2 contained a 3.8 kb *Hin*dIII fragment not present in pBT1 (see Figure 7). When these same plasmids were transformed into a completely catalase deficient mutant, UM53, their catalase levels also differed. pBT2 which contained the extra 3.8 kb *Hin*dIII fragment complemented the catalase deficiency of UM53, giving it catalase levels in the oxygraph of 3.7 u per mg dry cell wt with 1.5 mM H_2O_2 and 29.8 u per mg dry cell wt with 60 mM H_2O_2, while pBT1 had no detectable catalase in the oxygraph. The oxygraph values for these strains are given in Table 16, and the construction of these deletion plasmids is shown in Figure 6. These results suggested that all or part of the HPI gene was encoded in this 3.8 kb fragment.

The orientation of the 3.8 kb *Hin*dIII fragment with respect to the vector was shown to be the same as in pCL36-19. Both pBT2 and pC1 were restricted with *Hin*cII and the 1800 bp and 2100 bp fragments which cover the region between the beginning of the 3.8 kb fragment and the ColE1 vector were identified in both plasmids, confirming the orientation to be the same. If the fragment had been reversed in orientation, different restriction fragments would have resulted. The fragment sizes from the digests are given in Table 12.

4.6 Cloning and Restriction Mapping of the Catalase Gene

The 3.8 kb *Hin*dIII fragment identified by deletion analysis to carry all or part of the catalase gene was cloned into the *Hin*dIII site of the small (3570 bp) high copy number vector, pAT153. The catalase activities of the clones were screened in the catalase deficient NG mutant, UM228 and plasmid DNA was prepared from 4 catalase positive colonies. All of the catalase positive clones contained the 3.8 kb

Figure 7. *Hin*dIII restriction endonuclease digestion showing the presence of a 3.8 kb fragment in pBT2 not found in pBT1. DNA samples were digested with *Hin*dIII, electrophoretically separated on agarose gels, stained, and photographed. Lanes: a, λ DNA; b, pBT1; c, pLC36-19; d and e, pCI; and f, pBT2. Samples in lanes a, b, c, and d were run on a separate gel from the samples in lanes e and f and consequently, the λ- *Hin*dIII standard in lane a (fragment sizes shown in photograph) cannot be used to size the bands in lanes e and f. The pCI-*Hin*dIII digest in lane d however has also been included in lane e of the second gel as a size reference.

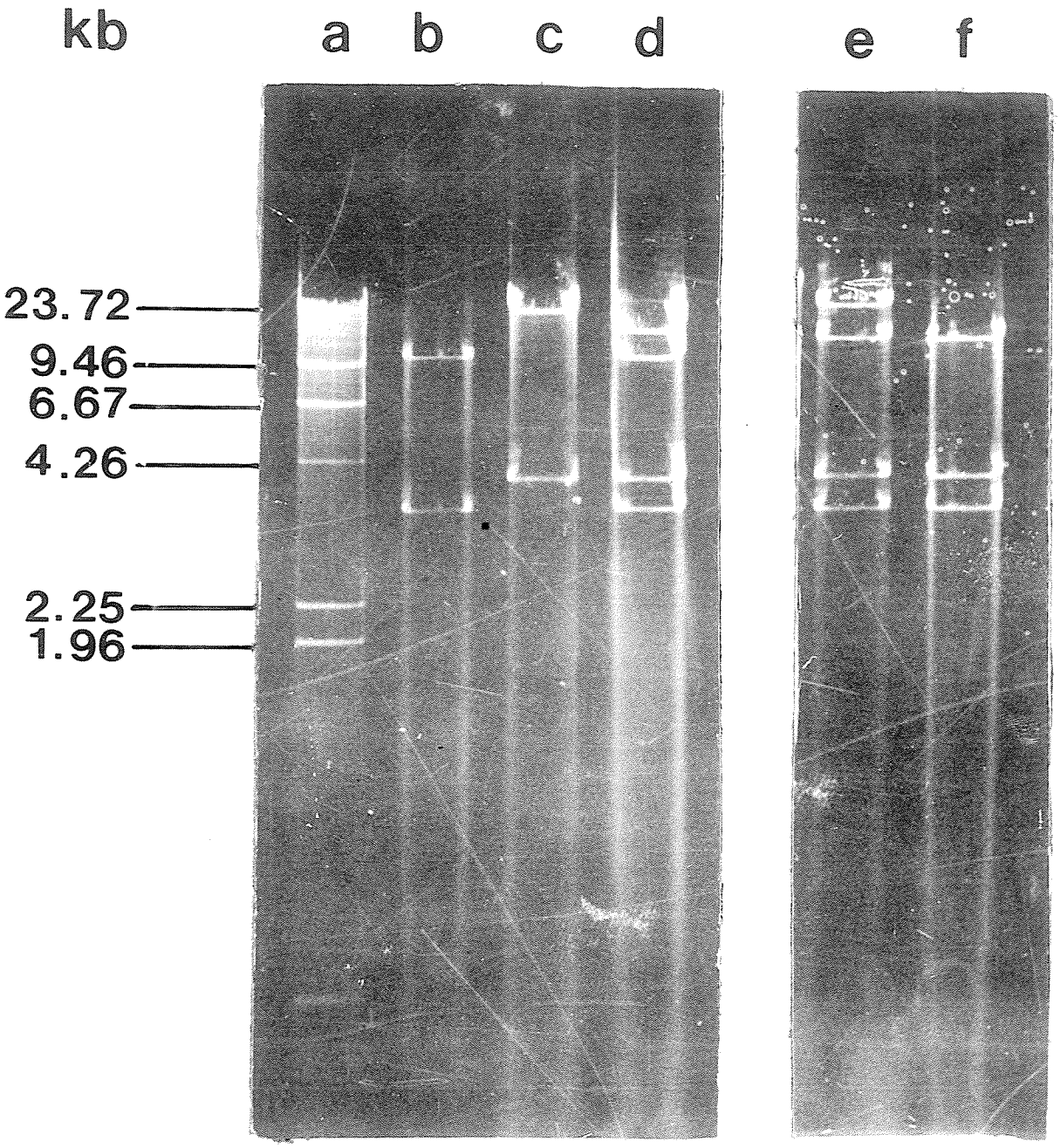

kb
a
b
c
d
e
f
23.72
9.46
6.67
4.26
2.25
1.96

Table 12. **Restriction fragments generated from pBT2, pLC36-19, and pC1 by *Hin*cII**

Plasmid	Restriction fragment sizes (bp)	Sum (bp)
pLC36-19	7900, 7050, 5200, 2100, 1800, 1350, 340	25 740
pC1	7050, 7050, 5200, 2690, 2430, 2100, 1800, 1350, 1100, 340	29 260
pBT2	7050, 2690, 2100, 1800, 1720, 1120, 340	16 820

*Hin*dIII fragment, but two types of restriction patterns were found when they were digested with *Eco*RV. One clone of each type of plasmid was prepared in large scale and these were labelled as pBT22 and pBT54 (sizes of the *Eco*RV restriction fragments generated from these two plasmids are given in Tables 13 and 15). These two plasmids contained the same *Hin*dIII insert, but in opposite orientations. The fact that the catalase gene was active in both orientations confirmed that the entire gene was present in this fragment, including the promoter, and that expression of the catalase gene was not affected by the close proximity of the insert to the tetracycline promoter in pBT22.

Crude extracts from the cells were electrophoresed on polyacrylamide gels and stained for catalase and peroxidase (see Figure 8). Although UM228 did produce residual HPI it was obvious that the plasmid containing strains were producing much larger quantities of the HPI A and B doublet.

The plasmids were transformed into UM53 and extracts of the plasmid harbouring bacteria were assayed for catalase using the oxygraph. The plasmids conferred very high levels of catalase activity to UM53 both on the agar plate and in the oxygraph. The levels were much higher than when the same fragment was cloned in the lower copy number vector ColE1 (pBT2). The oxygraph values are given in Table 16.

pBT22 was restriction mapped in detail. The restriction sites were first placed using the restriction digests given in Table 13, and then ordered in detail with the help of restriction fragments electrophoresed and sized on acrylamide gels (see Table 14). In total, sites for 14 restriction enzymes were mapped in the plasmid for a total of 37 restriction sites, the majority of which were clustered at the 0 kb end

Figure 8. Visualization of catalase and peroxidase activities in UM228 and the plasmid containing derivatives UM228/pBT22 and UM228/pBT54. Lanes b, d, and f were stained for catalase activity and lanes a, c, and e were stained for peroxidase activity. The following samples were run: lanes a and b from UM228/pBT54, 676 µg protein; lanes c and d from UM228/pBT22, 490 µg protein, and lanes e and f from UM228, 654 µg protein.

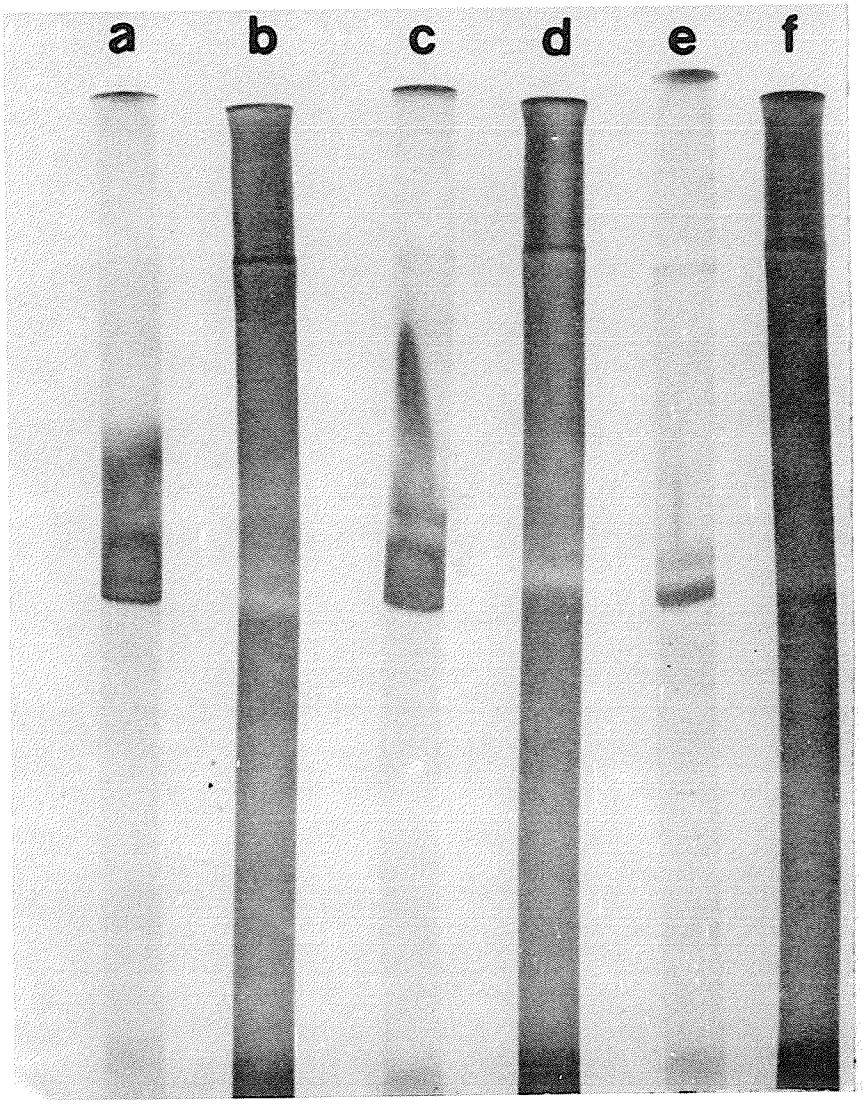

a
b
c
d
e
f

Table 13. Restriction enzyme fragments generated from pBT22

Enzyme	Fragment sizes (bp)	Sum (bp)
*Acc*I	3940, 3230	7170
*Bgl*II	6370, 810, 320, 140	7640
*Cla*I	6900, 480	7380
*Eco*RV	3460, 2280, 1380	7120
*Hin*dII	2520, 2120, 1620, 735, 340, 110, 90	7535
*Hin*dIII	3830, 3570	7400
*Nru*I	2770, 1900, 1250, 1250	7170
*Acc*I-*Cla*I	3230, 2860, 650, (480)[a]	7220
*Acc*I-*Eco*RV	3040, 2280, 1040, 500, 450	7310
*Acc*I-*Hin*dII	2440, 1700, 1530, 735, 360, 340, 110, 90	7305
*Eco*RV-*Cla*I	3500, 2300, 1400, (480)	7680
*Hin*dII-*Eco*RV	2520, 1250, 1150, 950, 680, 440, 340, 110, 100, 90	7600
*Hin*dII-*Sph*I	2520, 1620, 1150, 1000, 735, 340, (100, 90)	7555
*Hin*dII-*Bam*HI	2520, 1720, 1330, 735, 370, 340, 310, 110, 90	7525
*Hin*dIII-*Hin*dII	2520, 2120, 850, 620, 545, 340, 270, 110, 90	7465
*Hin*dIII-*Sph*I	3030, 1980, 1980, 530	7520
*Hin*dIII-*Bam*HI	3230, 2640, 1220, 350	7440
*Hin*dIII-*Cla*I	3570, 3360, 480	7410
*Nru*I-*Hin*dII	2080, 1610, 1250, 700, 640, 340, 280, 170, 110, 90	7270
*Nru*I-*Cla*I	2540, 1900, 1250, 1040, 240, 240	7210
*Nru*I-*Eco*RV	2770, 1250, 1200, 1100, 700, 260, 100	7380

[a] Brackets surround fragments which could not be accurately sized on the gel or fragments which could not be seen but were known to be present

Table 14. Restriction fragments of pBT22 sized on acrylamide gels and used to order restriction enzyme sites

Enzyme	Fragment sizes (bp)
Ball	170, 160
*Eco*R1	340
*Ball-Eco*RV	170, 160, 90
*Ball-Hind*III	740, 170, 140, 110, 90
*Ball-Hind*III	260, 170, 160
*Ball-Cla*I	260, 160, 100, 90
*Ball-Eco*R1	270, 170, 80, 70
*Ball-Bgl*II	630, 210, 170, 140, 130
*Cla*I-*Hind*III	520, 270, 250, 110, 90, 90
*Acc*I-*Bgl*II	650, 320, 140, 110
*Eco*R1-*Bgl*II	700, 290, 140, 95
*Eco*R1-*Hind*III	490, 340, 300, 90
*Eco*R1-*Nru*I	230, 100
*Bgl*II-*Hind*III	590, 470, 340, 220, 140, 100
*Bgl*II-*Eco*RV	700, 280, 220, 140, 100
*Bgl*II-*Nru*I	200, 140, 140, 80
*Hind*III-*Eco*RV	160, 155

of the pBT22 insert. The restriction map of the plasmid is shown in Figure 9.

pBT54, which contained the same insert but in the opposite orientation was restricted with *Hin*cII to confirm the opposite orientation of the insert. A map containing the restriction fragments from Table 15 is shown in Figure 10.

4.7 Definition of the Boundaries of the HPI Gene

4.7.1 Definition of the 0 kb End by a Specific *Eco*RI Deletion

The plasmid pBT22 was restricted with *Eco*RI, religated, and transformed into UM228. Colonies selected on LBA plates were screened for catalase activity by the H_2O_2 drop test and a large scale plasmid preparation was done on one catalase negative colony. The plasmid was called pBT24. An *Eco*RI digest of pBT24 produced one band on an agarose gel and no small bands on an acrylamide gel, confirming that the 340 bp *Eco*R1 fragment had been deleted. When pBT24 was transformed into UM53, no catalase activity was produced as shown by the oxygraph values in Table 16. The placement of this deletion with reference to the 0 kb end of the plasmid is shown in Figure 12. These results showed that part of the catalase coding region or controlling region must start before the *Eco*RI site, placing it within 300 bp of the 0 kb end.

4.7.2 Cloning of a 7550 bp *Sph*I fragment from pLC36-19

A *Sph*I restriction digest of the Tn*5* containing derivative of pLC36-19, pA30, was ligated with *Sph*I restricted pAT153 and then transformed into UM123. The plasmid containing cells were selected on LBA plates and screened for recombinant plasmids (Tc^S) on LBT plates. Plasmid screening preparations were done on twelve Tc^S clones, and DNA

Figure 9. Restriction map of pBT22. The thick line represents the vector pAT153, and the thin line represents the insert DNA. The 0 kb end marks the start of the 3.8 kb *Hin*dIII insert. The map was made using the restriction fragments in Tables 13 and 14 (Triggs-Raine and Loewen, 1987).

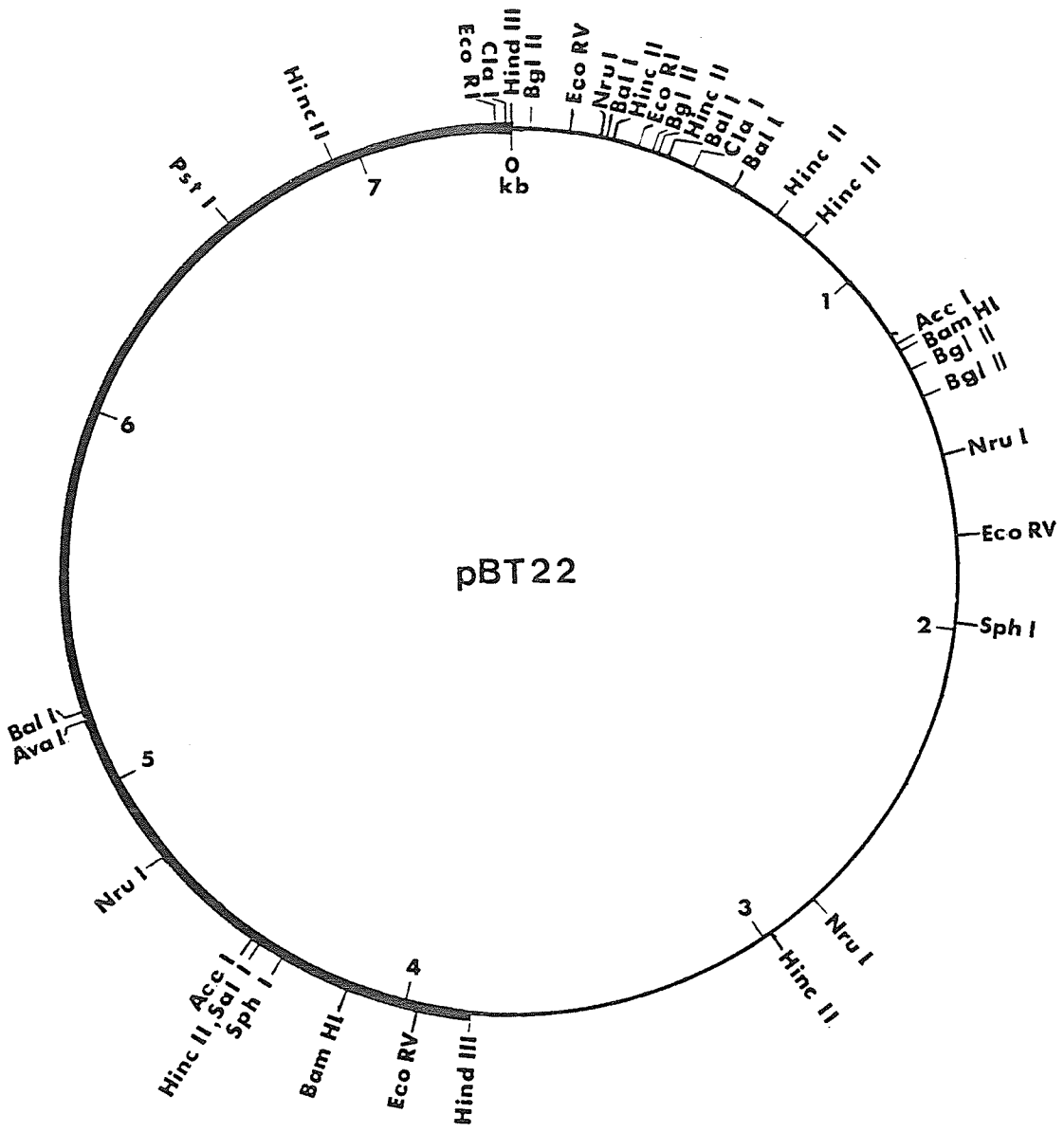

pBT22
0 kb
1
2
3
4
5
6
7
Eco RI
Cla I
Hind III
Bgl II
Eco RV
Nru I
Bal I
Hinc II
Eco RI
Bgl II
Hinc II
Bal I
Cla I
Bal I
Hinc II
Hinc II
Acc I
Bam HI
Bgl II
Bgl II
Nru I
Eco RV
Sph I
Nru I
Hinc II
Hinc II
Eco RV
Bam HI
Acc I
Sal I
Sph I
Nru I
Bal I
Ava I
Pst I
Hinc II

Table 15. **Restriction enzyme fragments generated from pBT54**

Enzyme	Fragment sizes (bp)	Sum (bp)
*Hin*dIII	3830, 3570	7400
*Hin*cII	2520, 2100, 1450, 930, 340, (100, 90)[a]	7530
*Eco*RV	5500, 1380, 400	7280

[a] Brackets surround bands not visible on an agarose gel

Figure 10. Restriction map of pBT54. The thick line represents the vector, pAT153, and the thin black line represents the insert DNA. Although not substantiated by restriction enzyme digests, the same restriction enzyme sites present in the pBT22 insert should be present in this insert since it is simply reversed in orientation.

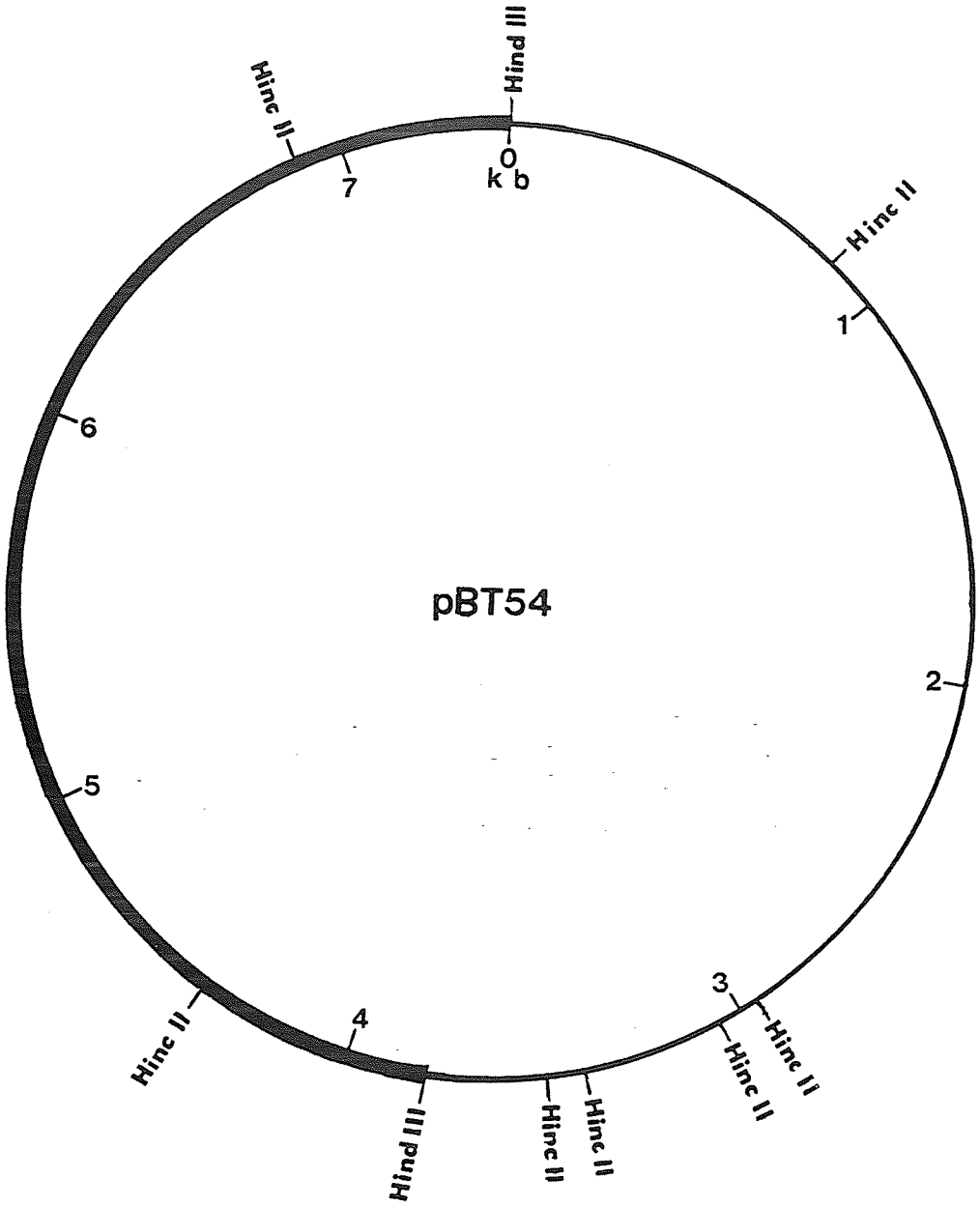

Hind III
0
kb
Hinc II
1
Hinc II
2
3
Hinc II
Hinc II
Hinc II
Hinc II
Hind III
4
Hinc II
5
6
Hinc II
7
pBT54

Table 16. Catalase activity in log phase UM53 containing various plasmids

Strain	Catalase Activity (u/mg dry cell wt) 1.5 mM H_2O_2	60 mM H_2O_2
UM53	ND[a]	0.2
UM53/pLC36-19	4.9	31.9
UM53/pC1	3.5	21.0
UM53/pE6	4.0	26.4
UM53/pA30	3.6	20.6
UM53/pD31	0.3	3.3
UM53/pBT5	ND	ND
UM53/pBT22	7.1	91.6
UM53/pBT54	6.2	80.0
UM53/pBT30	ND	1.5[b]
UM53/pBT29	ND	ND
UM53/pBT28	8.4	109.3
UM53/pBT24	ND	ND
UM53/pBT2	3.7	29.8
UM53/pBT1	ND	ND

[a] not detectable
[b] the value decreased to 0 after incubation at 0°C

from the largest plasmid, pBT5, was prepared for restriction enzyme digestion. pBT5 was found to contain the 7550 bp *Sph*I fragment from pLC36-19 that included a large portion of the 3.8 kb *Hin*dIII fragment cloned into pBT22 and extended through the 0 kb end of the *Hin*dIII fragment into pLC36-19. The orientation of the *Sph*I fragment in pAT153 was determined from the restriction enzyme digests given in Table 17 and the restriction map is shown in Figure 11.

pBT5 was transformed into the catalase deficient mutant UM53, but it was unable to complement the catalase mutation (see Table 16 for the oxygraph assay values). This showed that the entire HPI gene was not contained in pBT5 and that the HPI gene must extend from within 300 bp of the 0 kb end to beyond the *Sph*I site at approximately 2 kb on the pBT22 map (See Figure 12).

4.7.3 Definition of the 3.8 kb End of the Gene by BAL31 Deletions

BAL31 deletion plasmids were constructed as described in section 3.12.2. Several plates of Ap^R colonies were picked onto another LB plate and scored for catalase by the H_2O_2 drop test. Rapid plasmid preparations from five catalase negative and two catalase positive colonies were run on an agarose gel. The largest of the catalase negative plasmids was named pBT25. Plasmid DNA from a large scale preparation of pBT25 was restricted by *Hin*dIII, *Hin*cII, and *Nru*I (data not shown). The deletion of pBT25 was found to extend from between the *Nru*I site and the *Sal*I site of pAT153 to between the *Hin*cII site closest to the 3.8 kb end of the insert and the *Eco*RV site at 1.75 kb (refer to the pBT22 map in Figure 9). All *Sph*I sites were absent. pBT25 was then used as a size control to look for larger catalase negative plasmids.

Table 17. Restriction enzyme fragments generated from pBT5

Enzyme	Fragment size (bp)	Sum (bp)
*Sph*I	7550, 3570	11 120
*Bam*HI	9410, 860	10 270
*Eco*RI	8430, 2000, 400	10 830
*Hin*dIII	8630, 2350	10 980
*Cla*I	9000, 1830	10 830

Figure 11. Restriction map of pBT5. The thick line represents the vector, pAT153 and the thin line represents the insert DNA.

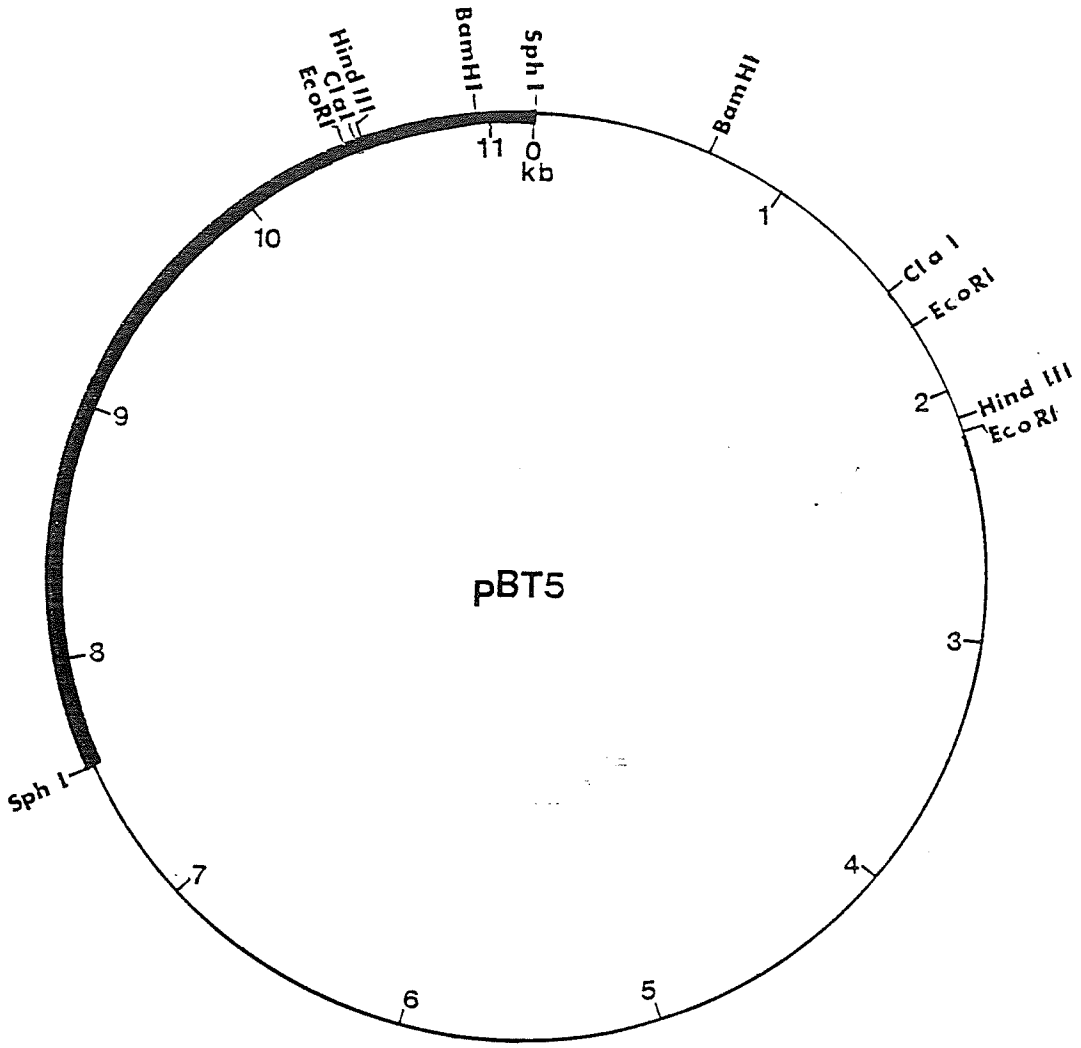
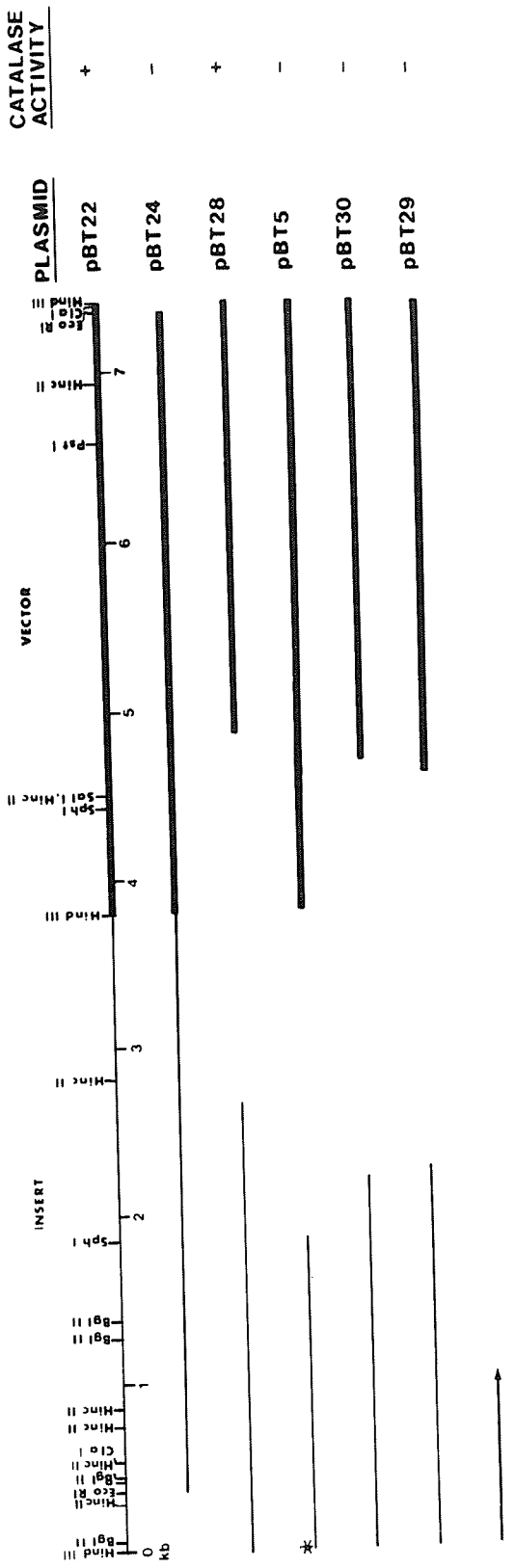

pBT5
Sph I
BamHI
Hind III
ClaI
EcoRI
0
kb
1
2
3
4
5
6
7
8
9
10
11
BamHI
Cla I
EcoRI
Hind III
EcoRI
Sph I

Figure 12. Localization of *katG* by deletion mapping of pBT5, pBT24, pBT28, pBT29, and pBT30. The thick line represents the vector DNA while the thin line represents the insert DNA. The asterisk indicates that pBT5 extends beyond the *Hin*dIII site at the 0 kb end to the next *Sph*I in the *E. coli* genomic DNA insert of pLC36-19. The arrow indicates the direction of transcription of the *katG* gene (Triggs-Raine and Loewen, 1987).


CATALASE ACTIVITY
PLASMID
pBT22
pBT24
pBT28
pBT5
pBT30
pBT29
+
−
+
−
−
−
VECTOR
INSERT
Hind III
Hinc II
Pst I
Sal I
Sph I
Hinc II
Hind III
Hinc II
Bgl II
Bgl II
Sph I
Hinc II
Hinc II
Cla I
Hinc II
Bgl II
Eco RI
Hinc II
Hind III
Bgl II
0
1
2
3
4
5
6
7
kb

Plasmids from approximately 80 colonies were screened by electrophoresis on an agarose gel with pBT22 and pBT25 as upper and lower limit size controls. The catalase activity of these colonies on agar plates was also scored. From these 80 plasmids, one catalase positive plasmid, pBT28, and two catalase negative plasmids (pBT29 and pBT30), all larger than pBT25 were chosen for further characterization. The deletions in the plasmids were roughly mapped using the restriction enzyme digests given in Table 18. These restriction digests were also useful because they allowed *Rsa*I sites to be mapped in pBT28 and this enzyme was later used to clone fragments for sequencing. Approximately 2100 bp had been deleted from pBT28 leaving a 5300 bp plasmid which complemented the catalase deficiency of UM53 (See table 16). The deletion in pBT28 was found to extend from the *Ba*lI site in the vector past the *Hin*cII site at the 3.8 kb end of the insert but ending before the *Nru*I site following this. The extent of the deletion is illustrated in Figure 12. pBT29 and pBT30 had deletions extending from approximately the same place in the vector through the *Hin*cII site at the 3.8 kb end of the insert and beyond the *Nru*I site, but ending before the *Sph*I site (see Figure 12). The deletions of pBT29 and pBT30 were almost exactly the same, pBT30 having only a slightly larger deletion. The extra DNA which has been deleted could be from the vector or the insert since there were no restriction enzymes mapped in this area to allow more precise definition. In Figure 12, the extra deleted DNA in pBT30 has simply been divided between the vector and the insert, although the exact boundaries are unknown. The *Eco*RV-*Sph*I digests of pBT29 and pBT30 shown in Table 18 confirmed that the *Sph*I sites were still present in both instances, but the resulting small fragment could not be seen on the agarose gel making

Table 18. Restriction enzyme fragments used to map pBT28, pBT29, and pBT30

Plasmid	Enzyme(s)	Fragment size (bp)	Sum (bp)
pBT28	*Bgl*II	3980, 810, 330, 140	5260
pBT28	*Eco*RV	3830, 1380	5210
pBT28	*Hin*cII	4030, 735, 340, 110, 90	5305
pBT28	*Nru*I	3910, 1250	5160
pBT28	*Rsa*I	2840, 1200, 640, 480, 200	5360
pBT28	*Bgl*II-*Rsa*I	2780, 640, 480, 360, 330, 300, 200, 150, 140	5380
pBT28	*Eco*RV-*Rsa*I	2840, 700, 520, 500, 480, (200)[a]	5040 (5240)[b]
pBT28	*Eco*RV-*Ava*I	2320, 1500, 1380	5200
pBT28	*Sph*I-*Ava*I	4110, 1110	5220
pBT28	*Rsa*I-*Bam*HI	2840, 1200, 530, 480, (200)	5050 (5250)
pBT28	*Rsa*I-*Cla*I	2840, 640, 520, 480, 480, (200), (200)	4960 (5360)
pBT28	*Rsa*I-*Sph*I	2840, 1200, 640, 480	5160
pBT29	*Eco*RV	3530, 1380	4910
pBT29	*Nru*I	2720, 1250, 810	4780
pBT29	*Rsa*I	2640, 1200, 640, 460, 200	5140
pBT29	*Eco*RV-*Sph*I	3290, 1380	4670
pBT30	*Eco*RV	3360, 1380	4740
pBT30	*Nru*I	2750, 1250, 810	4810
pBT30	*Rsa*I	2570, 1200, 640, 460, 200	5070
pBT30	*Eco*RV-*Sph*I	3260, 1380	4640

[a] brackets surround fragments whose presence was inferred but no acrylamide gel was run to confirm this

[b] brackets surround sums that include the values of the fragments inferred to be present

the total size of the plasmids appear smaller for these digests.

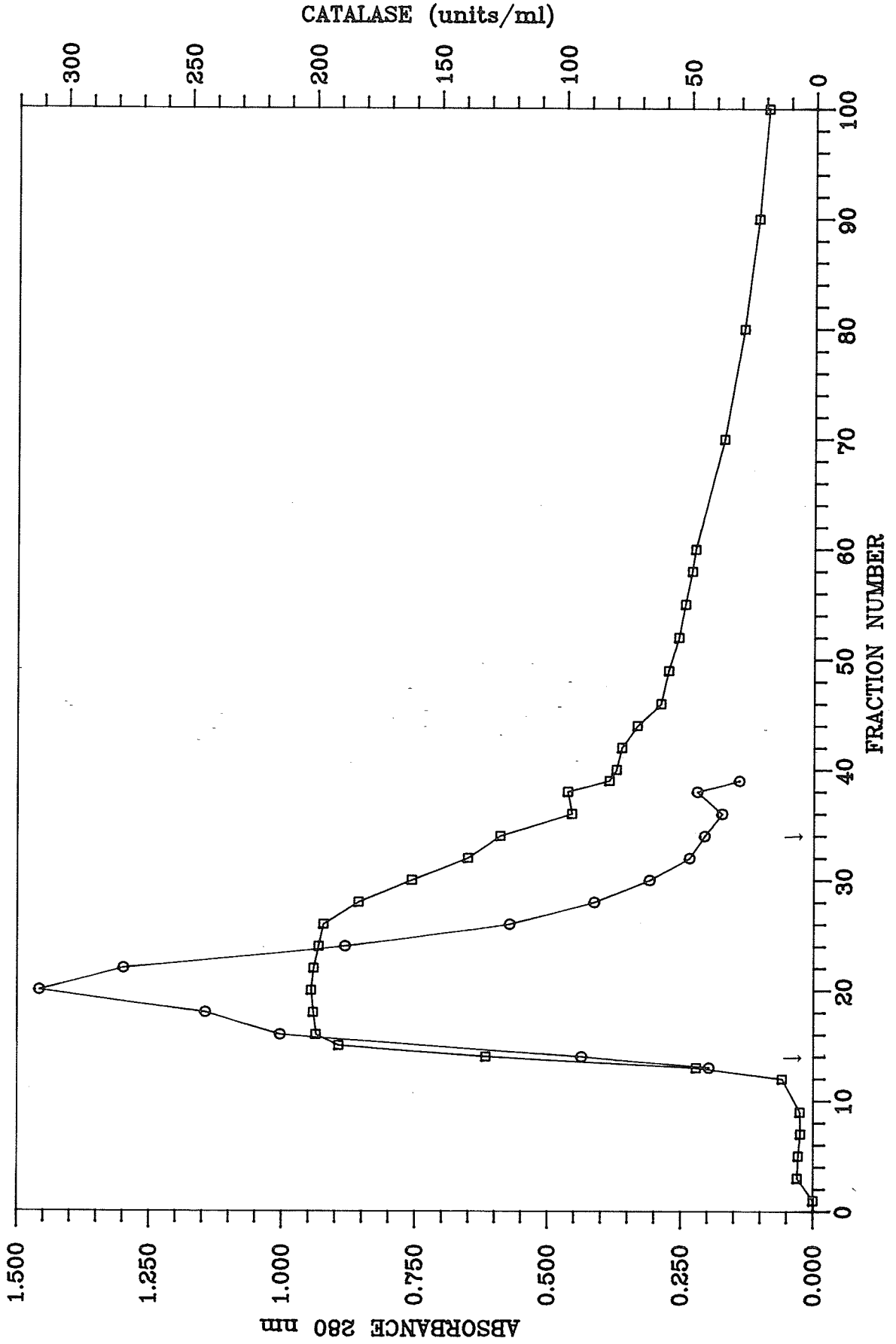
4.8 Identification of the Promoter

4.8.1 Immunodiffusion Analysis of Partially Purified HPI

Crude catalase preparations were prepared from UM262, UM262/pBT22, UM262/pBT24, and UM262/pBT5 following the partial purification procedure for HPI catalase outlined in section 3.14.1. UM262 was chosen as the host strain for these preparations because it produced no HPI catalase and therefore should produce no anti-HPI precipitable protein. The UM262/pBT22 preparation provided a model for the other preparations because it was the only strain carrying a catalase positive plasmid and thus allowed catalase activity to be assayed throughout the preparation. As illustrated in Figure 13, when protein from UM262/pBT22 was run through a DEAE-Sephadex column, the fractions containing the large protein peak, tubes 14 to 34, also contained the HPI catalase peak. It was hoped that by pooling the fractions covering the protein peak, during the subsequent preparation of UM262/pBT24 and UM262/pBT5, any inactive HPI protein with similar properties would be included. The inactive HPI protein would then be detected with anti-HPI.

The deletions in pBT5 and pBT24 were known to be at opposite ends of the 3.8 kb *Hin*dIII fragment carrying the HPI gene, yet both plasmids were catalase negative. If a cross reacting protein from the partial HPI purification was detected with anti-HPI in one strain but not in the other, this might suggest at which end of the insert the promoter was situated. The presence of an immunoprecipitable protein without catalase activity would suggest that the deletion in the plasmid in this strain was in the end of the gene encoding the carboxyl terminus of the HPI protein. A deletion in the promoter end would likely produce no anti-HPI

Figure 13. Determination of protein concentrations and catalase activity in fractions collected from a DEAE-Sephadex column. Approximately 500 mg of crude protein from UM262/pBT22 was loaded on a DEAE-Sephadex column and after washing the column with phosphate buffer, 0.5 M NaCl was run through the column and 100-75 drop fractions were collected. The protein peak was followed by taking the A_{280} of certain fractions (□), and the catalase peak was followed by assaying the fractions with 1.5 mM H_2O_2 in the oxygraph. (○) The arrows indicate the fractions which were pooled to provide a crude catalase preparation for immunodiffusion analysis.


CATALASE (units/ml)
300
250
200
150
100
50
0
100
90
80
70
60
50
40
30
20
10
0
FRACTION NUMBER
ABSORBANCE 280 nm
1.500
1.250
1.000
0.750
0.500
0.250
0.000

detectable protein.

The crude catalase preparations from the three plasmid bearing strains and UM262 itself were used for immunodiffusion analysis against anti-HPI. Fifty µL of each sample was placed in a well in an immunodiffusion plate surrounding a central well containing 50 µL of anti-HPI. After allowing the samples to react, only the UM262/pBT22 preparation was found to contain immunoprecipitable protein (see Figure 14). The absence of immunoprecipitable protein in the strains bearing the deletion plasmids prevented the identification of any one end as promoter containing. Since one deletion must be in the carboxyl terminus of the gene the absence of immunoprecipitable protein suggested that the loss of the carboxyl end of the protein may prevent the polypeptide from folding into a structure recognizable by any of the polyclonal population of antibodies present in the HPI antisera or that the protein was so unstable that it was quickly degraded, preventing it from being detected with the antisera. The absence of any cross reacting protein in the UM262/pBT5 deletion plasmid will be further alluded to in the next section.

4.8.2 Maxicell Analysis of Plasmid Encoded Polypeptides

The plasmid encoded polypeptides from pBT22, pBT54, pBT5, and the related deletion plasmids pBT28, pBT29, pBT30, and pBT24 were analyzed by maxicell expression. The autoradiogram of the plasmid encoded polypeptides is shown in Figure 15. Cells containing no plasmids produced no protein bands, only a fast running band of cell debris and associated label. A single protein band with a relative molecular weight (M_r) of 32 000 was produced by cells containing the plasmid vector pAT153. The Ap^R gene of pAT153 codes for two proteins weighing 28 000 daltons and

Figure 14. Immunodiffusion analysis of crude catalase preparations against anti-HPI. Wells: a, UM262; b, UM262/pBT5; c, UM262/pBT22; d, UM262/pBT24; e, UM262/pBT22.

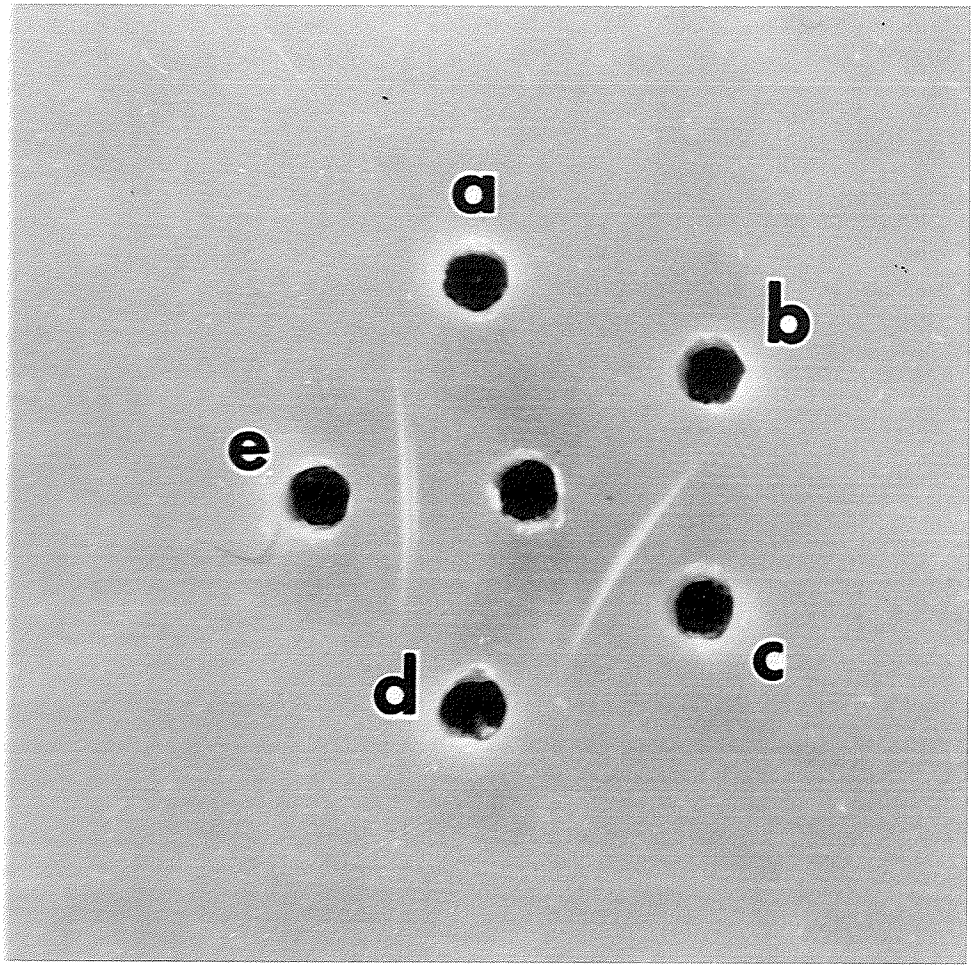
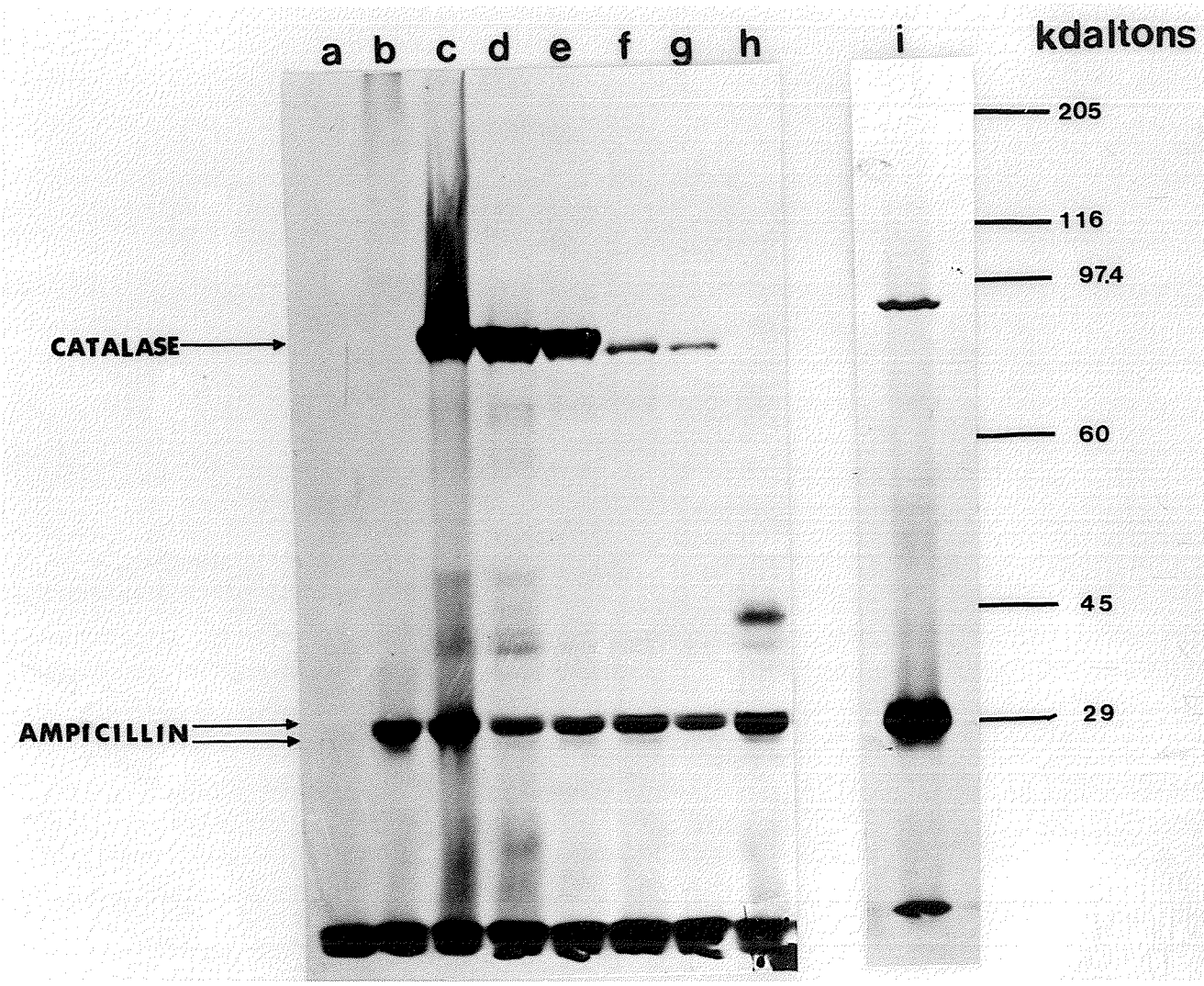

a
b
e
c
d

Figure 15. Maxicell analysis of plasmid encoded polypeptides. ^{14}C-labelled polypeptides were analyzed by electrophoresis on an 8.0% SDS-polyacrylamide gel. Lanes: a, CSR603; b, CSR603/pAT153; c, CSR603/pBT22 (*katG*$^{+}$); d, CSR603/pBT54 (*katG*$^{+}$); e, CSR603/pBT28 (*katG*$^{+}$); f, CSR603/pBT29 (*katG*); g, CSR603/pBT30 (*katG*); h, CSR603/pBT24 (*katG*); i, UM262/pBT5 (*katG*) (Triggs-Raine and Loewen, 1987).


a b c d e f g h
i
kdaltons
205
116
97.4
60
45
29
CATALASE
AMPICILLIN

31 000 daltons (Sancar et al., 1979) which appeared here as the strong 32 000 dalton band under which a weaker band was sometimes visible. The Tc^R gene produces a weak band at 37 000 daltons (Sancar et al., 1979) which could not be seen on this autoradiogram. The catalase positive plasmids, pBT22, pBT54, and pBT28 contained the Ap^R gene products but also a new protein band at 84 000 daltons, the same size as the HPI subunit. The intensity of the HPI protein band relative to the ampicillin resistance protein indicated that all three plasmids produced a stable HPI protein at comparable efficiencies.

The deletion plasmids pBT29, pBT30, and pBT24 were useful in defining the location of the promoter. The two catalase negative BAL31 deletion plasmids, pBT29 and pBT30 produced a band of the same size or slightly smaller than the 84 000 dalton band produced from the catalase positive plasmids, as well as the Ap^R gene products. The weak intensity of the 84 000 dalton band in comparison to the Ap^R gene products suggested that the protein was unstable. The production of an unstable, but almost full size protein from these plasmids suggested the deletions at the 3.8 kb end of the insert were in the carboxyl terminal coding portion of the gene. If this were the promoter containing end of the gene the deletion would probably prevent the production of any HPI protein. pBT24, which had a deletion at the 0 kb end of the insert would then be expected to produce no protein if it were in the promoter region. However, besides the Ap^R proteins a second protein band appeared at 45 000 daltons. Although the presence of this band was not understood it was known that if the carboxyl terminus were at the 0 kb end, such a small deletion would not result in a protein this small, so it still supported the hypothesis that the 0 kb end contained the promoter.

The *Sph*I clone, pBT5, produced an unusual protein product as well. Because pBT5 contains an even larger deletion at the 3.8 kb end than either pBT29 or pBT30, it was expected to produce an even smaller HPI derivative. Instead it produced a protein band slightly larger than HPI, having a M_r of 92 000. This polypeptide is unlikely to have resulted from the in-frame fusion of the existing portion of the HPI gene and the Tc^R gene because this fusion protein would be about 75 amino acids smaller than the HPI protein. The larger deletion in the HPI coding sequence of pBT5 may have made the protein so unstable that it was no longer present in large enough quantities to be visible in the autoradiogram. The new protein band at 92 000 daltons could then be a polypeptide coded for by the large piece of DNA extending beyond the 0 kb end of the insert to the next *Sph*I site of pLC36-19. It was known from the immunodiffusion analysis in the previous section that the protein product of pBT5 was either lost during purification because of its altered properties or that it was not similar enough to cross-react with HPI antisera.

4.8.3 Cloning of the Promoter

As final confirmation that the HPI gene promoter was at the 0 kb end of the 3.8 kb *Hin*dIII fragment, a 320 bp *Bgl*II promoter containing fragment was selected in the promoter cloning vector pKK232-8 (Pharmacia).

The pBR322 derivative, pKK232-8 contains a promoterless chloramphenicol acetyltransferase (CAT) gene which can be activated by inserting a promoter in front of it. A multiple cloning site which precedes the gene makes the insertion of fragments very simple. Other features engineered into the plasmid include transcription terminators

flanking the gene to prevent pBR322 promoters from interfering with CAT expression and translational stop signals in all three reading frames between the multiple cloning region and the CAT gene to prevent the production of fusion proteins. A Shine-Dalgarno sequence is appropriately placed in front of the promoterless CAT gene so that only the transcriptional promoter sequences are required to express the CAT gene. The plasmid also contains an Ap^R gene for the selection of the plasmid (Brosius, 1984).

pKK232-8 was restricted with *Bam*HI which cuts in the multiple cloning region and *Bgl*II fragments from pBT22 were ligated in. The ligation mixture was transformed into UM228 and plasmid containing colonies were selected on LBA plates. Approximately 240 colonies were picked onto LB plates containing 10 µg/mL and 20 µg/mL of chloramphenicol. Plasmids were isolated from two colonies which grew better than the others. On an agarose gel, both plasmids appeared the same size or slightly larger than pKK232-8. One of the plasmids was named pGprm1 and prepared on a large scale.

The *Bgl*II insert in pGprm1 could no longer be cut out with the same enzyme because *Bam*HI/*Bgl*II hybrid sites are not recognized by either enzyme. The fragment was cut out by restriction with *Sma*I and *Sal*I which cut on either side of the *Bam*HI site in the multiple cloning region. As illustrated in Figure 16, the resulting fragment was 330 bp in size, almost exactly the same size as the 320 bp *Bgl*II fragment produced from pBT22. The fragment was slightly larger because of the additional base pairs from the multiple cloning region. Restrictions of pGprm1 with the asymmetrically placed *Eco*RI and *Eco*RV sites (shown in Figure 16 and sized in Table 19) were used to show that the orientation

Figure 16. Restriction enzyme digests identifying the 320 bp *Bgl*II fragment cloned into the promoter cloning plasmid pKK232-8 and confirming the orientation of the promoter to be the same as it is in *katG*. The DNA samples were electrophoresed on polyacrylamide gels and the sizes of the standard fragments used to size the other fragments are given beside the gel. Lane a: pBR322-*Hae*III; Lane b: pBT22-*Bgl*II; Lane c: pGprm1-*Sma*I/*Sal*I; Lane d:pGprm1-*Eco*RI; Lane e: pGprm1-*Hin*dIII/*Eco*RV; Lane f: pBR322-*Nci*I.

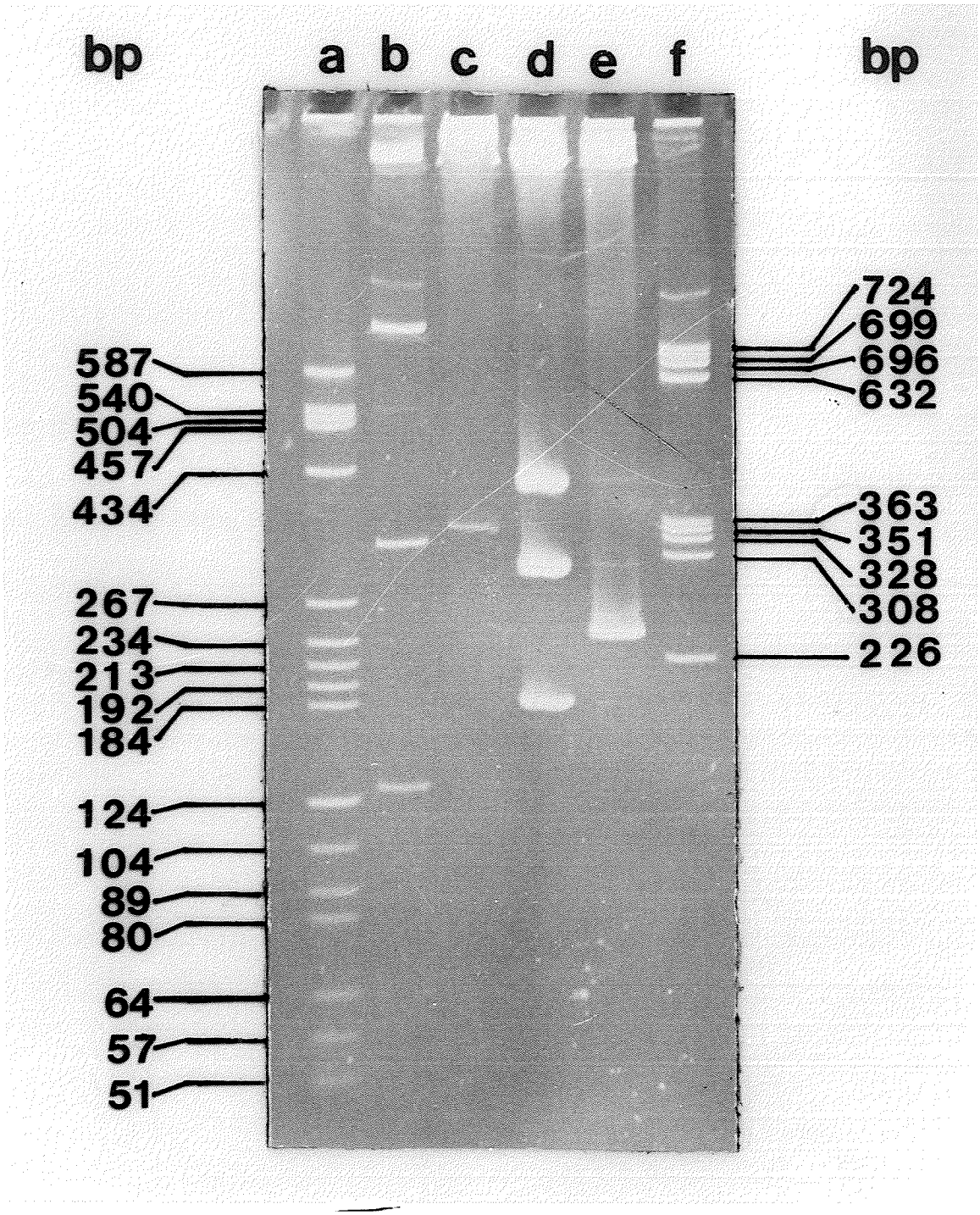

bp
a
b
c
d
e
f
bp
724
699
696
632
587
540
504
457
434
363
351
328
308
267
234
213
192
184
226
124
104
89
80
64
57
51

Table 19. Restriction fragments generated from pGprm1

Enzyme(s)	Fragment size (bp)
*Sma*I-*Sal*I	330
*Eco*RI	430; 290; 180
*Hin*dIII-*Eco*RV	230

of the *Bgl*II fragment was the same in both pBT22 and the promoter cloning plasmid pGprm1. The restriction map of pGprm1 is shown in Figure 17.

pGprm1 was retransformed into UM228 and 60 colonies were picked onto LBC plates to show that all of the colonies were chloramphenicol resistant. Colonies were also picked onto increasing concentrations of chloramphenicol and pGprm1 was found to confer resistance to more than 60 µg/mL of choramphenicol.

4.9 Identification of the Structural Gene for HPI as *katG*

The results presented in the previous sections showed that pBT22 contained the structural gene for the *E. coli* hydroperoxidase, HPI. A transposon Tn*10* insertion in the *katG* locus at 89.2 min was found to prevent the synthesis of the HPI catalase, but the role of *katG* in this phenomenon was still unclear; *katG* could be either the structural gene encoding HPI or a gene encoding a regulatory protein affecting HPI synthesis.

The following experiment was designed to answer the question,"was *katG* the structural gene encoding HPI?" Genomic DNA was prepared from both the wild type *E. coli* strain MP180, and the *katG17*::Tn*10* derivative, UM202. Approximately 2 µg of each DNA sample was restricted with *Hin*cII and the digests were run on an agarose gel with pBT22 digested by *Hin*cII as a standard at one end and the BRL 1 kb ladder standard at the opposite end. A Southern blot of the gel was prepared and then hybridized with nick translated pBT22 DNA under stringent conditions (6 x SSC, 65^{o}C). A photograph of the resulting autoradiogram is shown in Figure 18.

The first lane in the autoradiogram contained *Hin*cII restricted

Figure 17. Restriction map of pGprm1. The thick line represents the vector, pKK232-8 and the thin line represents the insert DNA. Only the restriction sites of interest in mapping the orientation of the *Bgl*II insert have been included (Triggs-Raine and Loewen, 1987).

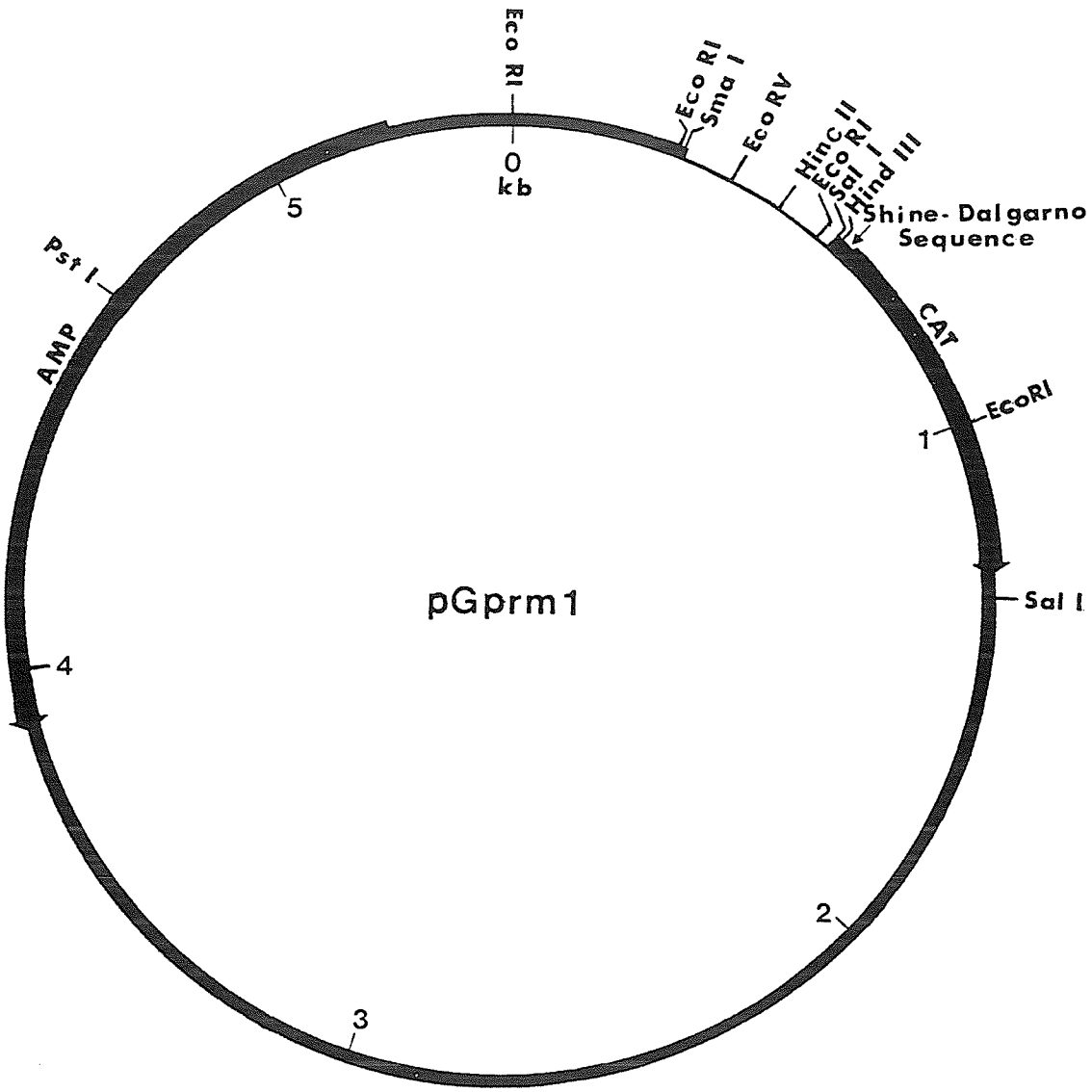
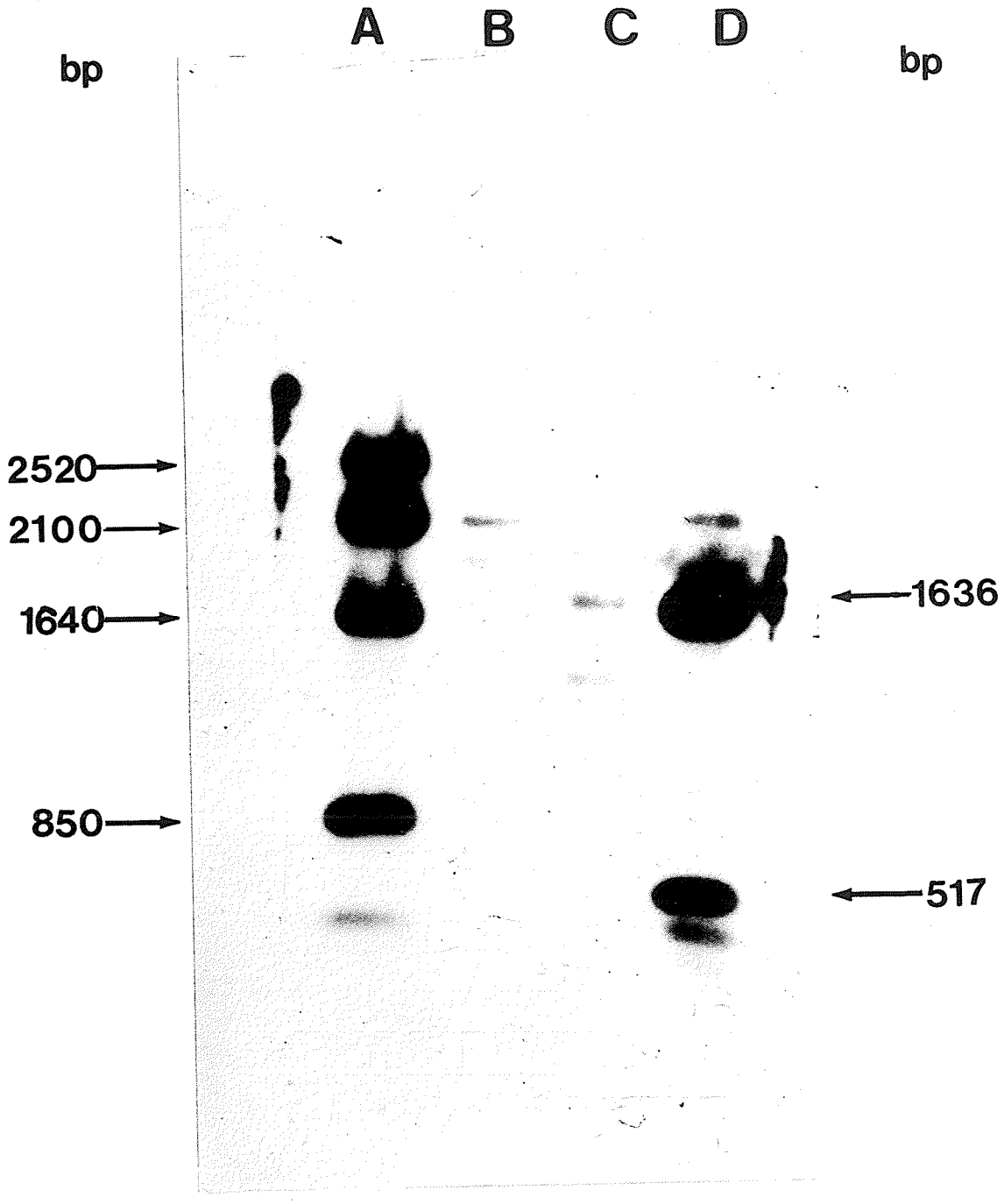

Eco RI
0
kb
Eco RI
Sma I
Eco RV
HinC II
Eco RI
Sal I
Hind III
Shine-Dalgarno
Sequence
CAT
EcoRI
1
Sal I
2
3
4
5
Pst I
AMP
pGprm1

Figure 18. Autoradiogram of the change in hybridization pattern when genomic DNA with or without a transposon Tn*10* insertion in *katG* was digested with *Hin*cII and probed with ^{32}P-labelled pBT22. Lanes A, pBT22 *Hin*cII; B, MP180 genomic DNA digested with *Hin*cII; C, UM202 (*katG*::Tn*10*) genomic DNA digested with *Hin*cII; D, BRL 1 kb ladder molecular weight standard containing some pBR322 fragments which hybridized with the vector portion of pBT22. The numbers along the side indicate the sizes in bp of the known fragments (Triggs-Raine and Loewen, 1987).


A
B
C
D
bp
bp
2520
2100
1640
850
1636
517

pBT22 DNA probed by itself. Of the five hybridizing bands in this lane, only the second (2100 bp) and the last band (340 bp) are solely composed of pBT22 insert DNA. The first band (2520 bp) was made up of the vector DNA only while the third and fourth bands (1640 bp; 850 bp) were hybrid fragments composed of both vector and insert DNA. Consequently, the bands expected to hybridize in a *Hin*cII digest of genomic DNA included the 2100 bp and 340 bp bands as well as two other bands extending from the *Hin*cII site of the pBT22 insert to the next *Hin*cII site in the *E. coli* genomic DNA. The sizes of these *Hin*cII bands were known to be about 1800 bp and 1350 bp from the restriction mapping of pLC36-19 which contained the genomic DNA surrounding the *Hin*dIII fragment cloned into pBT22.

The second lane in the autoradiogram contained *Hin*cII digested MP180 genomic DNA. The expected bands at 2100 bp, 1800 bp, and 1350 bp were present but the 340 bp fragment could not be seen. This was not surprising since the large fragments were only weakly visible and the small fragment would be expected to show weaker hybridization as was apparent for this band in the first lane. The third band in the hybridization (about 1650 bp) was not expected, and it is likely that this band resulted from a partial digest between the 1350 bp fragment the neighbouring 100 bp and 340 bp *Hin*cII bands (See Figure 9). This may have resulted because it is often more difficult for restriction enzymes to cut sites which are very close to each other. The fact that this band was not visible in the next lane of UM202 genomic DNA digested by *Hin*cII also suggested that it was the result of a partial digest.

The third lane containing the *katG17*::Tn*10* derivative of MP180 showed an altered pattern of hybridization. The 2100 bp band composing the majority of the pBT22 insert was absent and was replaced by two new

bands of about 1700 bp and 1100 bp in size. These fragments add up to more than 2100 bp because they each contain a portion of the 2100 bp fragment and 360 bp of the Tn*10* attached to the end. The first *Hin*cII sites in Tn*10* are about 360 bp from each end (Way et al., 1984) (see Figure 19).

The change in the hybridization pattern confirmed that the gene containing the transposon insertion, *katG* was the same gene that was encoding the HPI protein on pBT22.

4.10 The Sequencing of *katG* and Analysis of This Sequence

4.10.1 Sequencing Strategy

Restriction fragments from pBT22, pBT24, and pBT28 were subcloned into M13mp18 and M13mp19 for dideoxy chain termination sequencing. The M13 vectors (mp18 and mp19) contain a multiple cloning region (MCR) preceded by a universal primer binding site (see Figure 20). The opposite orientations of the MCR's in these two vectors allow the same inserts to be cloned, but the end of the insert closest to the MCR may vary depending on the vector choice. For instance, if a clone is constructed with the enzymes *Eco*RI and *Hin*dIII in M13mp18, the *Hin*dIII end of the fragment will be closest to the MCR, while in M13mp19, the *Eco*RI end will be closest to the MCR. Therefore by selecting the appropriate vector, the end of the fragment where sequencing begins can be controlled. In cases where the vector was cut with only one enzyme, the choice of vector became unimportant because the orientation of fragment insertion could not be controlled.

The sequencing strategy for *katG* is illustrated in Figure 21. The sequenced fragments are labelled to allow cross referencing to Table 20 for a description of their construction. The first fragments cloned for

Figure 19. Restriction map of transposon Tn*10*. The thicker lines represent the left and right inverted repeat segments of the 9300 bp transposon. The *Hin*cII (H) restriction sites are indicated (Way et al., 1984).

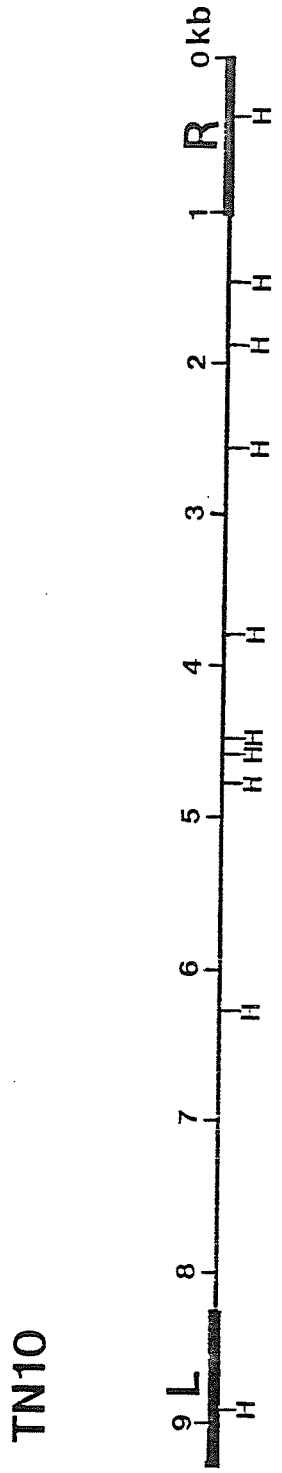

TN10
L
R
0kb
1
2
3
4
5
6
7
8
9
H

Figure 20. The multiple cloning regions of the M13mp18 and M13mp19 vectors. The order of the restriction sites with reference to the universal primer binding site is shown.

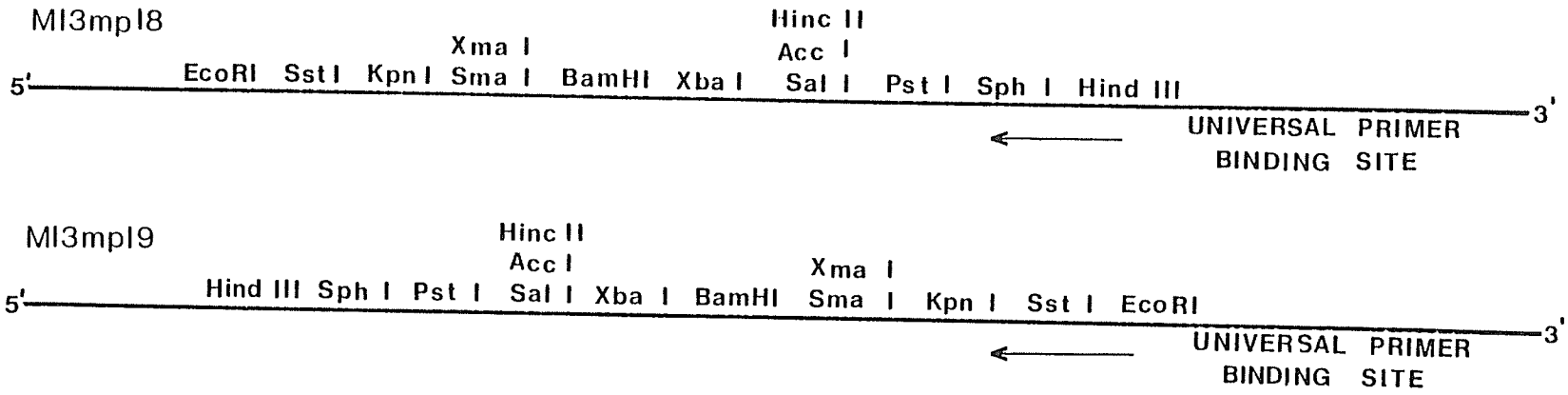

M13mp18
Hinc II
Acc I
Xma I
EcoRI Sst I Kpn I Sma I BamHI Xba I Sal I Pst I Sph I Hind III
5'
3'
UNIVERSAL PRIMER
BINDING SITE
M13mp19
Hinc II
Acc I
Xma I
Hind III Sph I Pst I Sal I Xba I BamHI Sma I Kpn I Sst I EcoRI
5'
3'
UNIVERSAL PRIMER
BINDING SITE

Figure 21. The sequencing strategy of *katG*. The origin of the arrows indicate the point at which sequencing began and the arrows point in the direction in which sequencing was done. The letters above the arrows can be cross referenced to Table 20 to see how the clones were constructed. The asterisks indicate the beginning and end of the *katG* coding region.

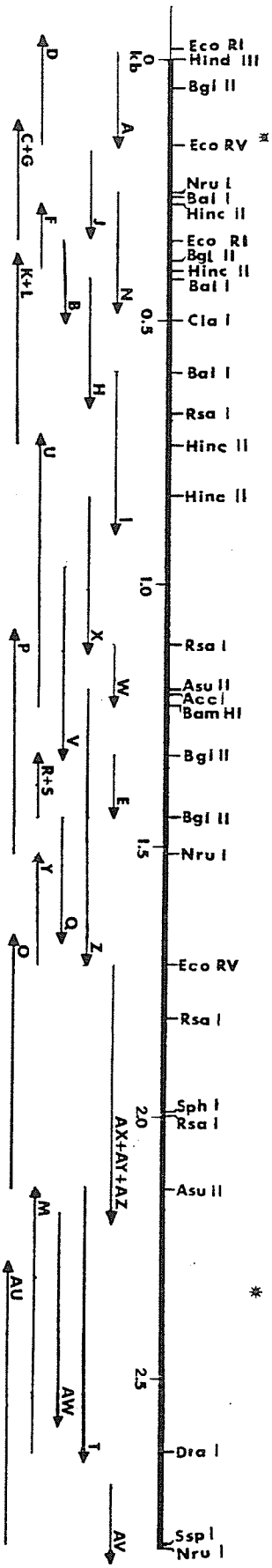

Eco RI
Hind III
Bgl II
Eco RV
Nru I
Bal I
Hinc II
Eco RI
Bgl II
Hinc II
Bal I
Cla I
Bal I
Rsa I
Hinc II
Hinc II
Rsa I
Asu II
Acc I
Bam HI
Bgl II
Bgl II
Nru I
Eco RV
Rsa I
Sph I
Rsa I
Asu II
Dra I
Ssp I
Nru I
kb
0
0.5
1.0
1.5
2.0
2.5
A
B
C+G
D
E
F
H
I
J
K+L
M
N
O
P
Q
R+S
T
U
V
W
X
Y
Z
AU
AV
AW
AX+AY+AZ

Table 20. Fragments cloned into M13mp18/19 for sequencing *katG*

Fragment	Vector	Enzyme(s)	Insert Origin	Enzyme(s)
A	mp19	*Eco*RI-*Sma*I	pBT22	*Eco*RI-*Eco*RV
B	mp19	*Eco*RI-*Acc*I	pBT22	*Eco*RI-*Cla*I
C	mp19	*Eco*RI-*Sma*I	pBT22	*Eco*RI-*Eco*RV
D	mp19	*Hin*dIII-*Sma*I	pBT22	*Hin*dIII-*Eco*RV
E	mp19	*Bam*HI	pBT22	*Bgl*II
F	mp19	*Hin*cII	pBT22	*Hin*cII
G	mp19	*Eco*RI-*Hin*dIII	pBT22	*Eco*RI-*Hin*dIII
H	mp18	*Bam*HI-*Hin*cII	pBT24	*Bam*HI-*Ba*II
I	mp18	*Bam*HI-*Hin*cII	pBT24	*Bam*HI-*Ba*II
J	mp18	*Bam*HI-*Hin*cII	pBT22	*Bam*HI-*Eco*RV
K	mp18	*Bam*HI-*Hin*cII	pBT22	*Bam*HI-*Hin*dIII
L	mp19	*Hin*cII	pBT22	*Hin*cII
M	mp19	*Acc*I-*Sma*I	pBT22	*Dra*I-*Asu*II
N	mp19	*Sma*I	pBT22	*Nru*I
O	mp19	*Hin*dIII-*Acc*I	pBT22	*Asu*II-*Hin*dIII
P	mp19	*Sma*I	pBT22	*Nru*I
Q	mp19	*Bam*HI-*Hin*cII	pBT22	*Bgl*II-*Eco*RV
R	mp18	*Bam*HI-*Hin*cII	pBT22	*Bgl*II-*Eco*RV
S	mp19	*Bam*HI	pBT22	*Bgl*II
T	mp19	*Acc*I-*Hin*dIII	pBT22	*Asu*II-*Hin*dIII
U	mp19	*Bam*HI-*Hin*cII	pBT24	*Bam*HI-*Ba*II
V	mp18	*Bam*HI-*Hin*cII	pBT24	*Bgl*II-*Eco*RV
W	mp18	*Bam*HI-*Hin*cII	pBT22	*Bam*HI-*Rsa*I
X	mp18	*Bam*HI-*Hin*cII	pBT28	*Bam*HI-*Hin*cII
Y	mp18	*Bam*HI-*Hin*cII	pBT22	*Bgl*II-*Eco*RV
Z	mp19	*Acc*I	pBT22	*Asu*II
AU	mp19	*Acc*I-*Sma*I	pBT22	*Ssp*I-*Asu*II
AV	mp19	*Bam*HI	pBT22	*Sau*3A
AW	mp19	*Sma*I	pBT22	*Hae*III
AX	mp19	*Sma*I	pBT22	*Eco*RV
AY	mp19	*Sma*I-*Hin*dIII	pBT22	*Hin*dIII-*Eco*RV
AZ	mp18	*Bam*HI-*Sma*I	pBT22	*Bgl*II-*Ssp*I

sequencing were *Bgl*II and *Hin*cII fragments. To determine if the clones contained insert DNA from pBT22, the RF DNA was prepared from all of the clones, run on an agarose gel, Southern blotted, and hybridized with nick translated pBT22. All of the cloned *Bgl*II fragments hybridized, indicating that they did contain inserts from pBT22, but there were several *Hin*cII clones that did not hybridize. It was found that *Hin*cII was a poor choice for restricting the vector because it appeared to contain a contaminating exonuclease which resulted in many white plaques containing no insert DNA. In the subsequent cloning procedures, the presence of pBT22 DNA in the recombinant M13 was not confirmed by hybridization. Two of the cloned fragments, M and AY, had been cut from an agarose gel, purified, and cloned. The *Sau*3A (AV) and *Hae*III (AW) clones were selected by cloning all possible fragments from pBT22, then hybridizing the RF DNA with the large 2100 bp *Hin*cII fragment of pBT22, which had been cut from an agarose gel, purified, and nick translated.

Single stranded DNA was prepared from a number of white plaques from each cloning experiment. All of the clones were T-tracked unless specific clones of interest had been identified by hybridization and in these instances only selected clones were T-tracked. Selection of clones for sequencing from the T-tracks was straightforward if part of the sequence of the same or the opposite strand was already known. If the sequence was unknown, and more than one type of clone was expected in the experiment, the following guidelines were helpful. The smallest DNA fragment greater than 100 bp in size normally cloned with the highest frequency and should therefore be the most common T-track. Small white plaques were often associated with very large inserts and they also had decreased yields of ssDNA which appeared as a light T-track.

Finally, vector deletions containing no insert DNA were common and the T-tracks were always examined for vector sequence.

The entire region between the *Hin*dIII site at the start of the pBT22 insert and the *Dra*I site, 2624 bp into the insert was sequenced in both directions (Figure 23). All of the restriction sites were overlapped to be sure that no small pieces of sequence were missed. Sequence beyond the *Dra*I site was not determined in both directions but it is included in a subsequent section in the interest of future research in the areas surrounding this gene.

4.10.2 The *katG* DNA Sequence

The sequence of the 2624 bp portion of pBT22, determined using the strategy in Figure 21, corresponded well with the restriction enzyme map previously generated for pBT22. Only one restriction site, *Cla*I at 499 bp was not identified by restriction enzyme digestion. This site was probably not found because the *Cla*I restriction site, 5'ATCGAT3', may not be recognized if the underlined A residue is methylated (Mayer et al., 1981).

The 2624 bp sequence contained only one open reading frame large enough to encode the HPI protein. The open reading frame started at base 146 with an ATG triplet specifying methionine and extended for 726 amino acids until it was ended by a TAA codon at 2326 bp. The DNA sequence between nucleotides 655 and 723 is shown in Figure 22 as an example of the DNA sequencing and interpretation. The entire DNA sequence is presented in Figure 23 with its predicted amino acid sequence underneath.

The *katG* DNA sequence was compared to the GenBank (1986) of DNA sequences as well as to yeast catalase T using the computer program,

Figure 22. Photograph of a portion of an autoradiogram of a sequencing gel. This sequence was obtained from clone I and allows the nucleotide sequence between bases 655 to 723 of the *katG* sequence to be read by reading from the bottom to the top of the photograph. The A, C, G, and T tracks are labelled above the lanes. The sequence reading from bottom to top is 5'GCTAGAAAACTCC GGCTTCCGTACCTTCGGTTTTGGTGCCGGTCGTGAAGACGTCTGGGAACCGGATCT 3'.

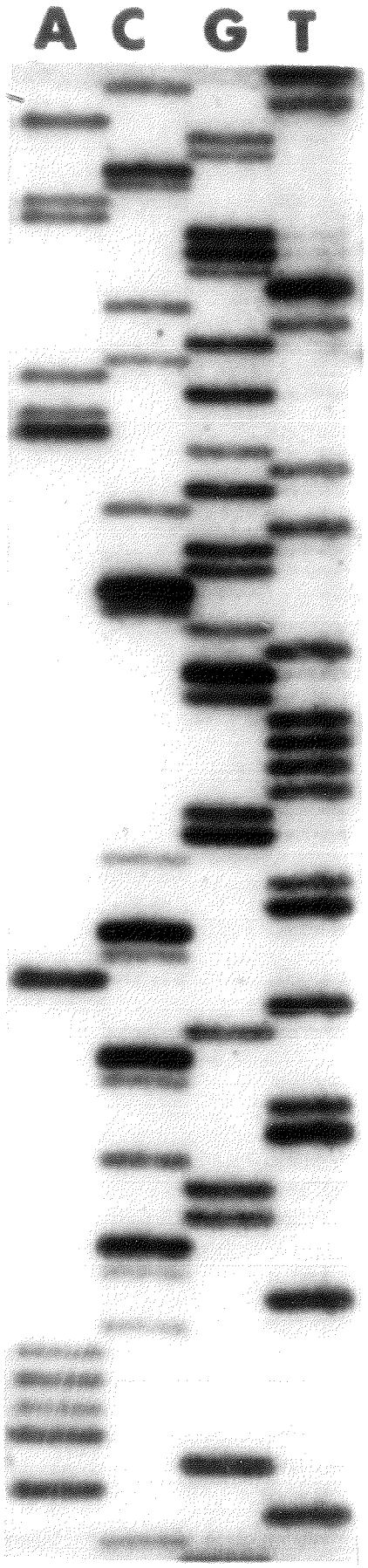

A C G T

Figure 23. The *katG* DNA sequence and the HPI predicted amino acid sequence. Numbers on the right hand side of the sequence specify the nucleotide number and those on the left hand side specify the amino acid number. Restriction sites relevant to the material presented in this thesis are italicized and labelled. Amino acids confirmed by direct sequencing of the purified HPI protein are underlined and in boldface print. A potential promoter (-35 and -10) sequence, Shine-Dalgarno sequence, and terminator sequence are emphasized in darker type.

HindIII 28
*AAGCTT*AATTAAGATCAATTTGATCTAC

BglII 100
ATCTCTTTAACCAACAATATGTA*AGATCT*CAACTATCGCATCCGTGGATTAATTCAA**TTATAA**CTTCTCTCT

EcoRV 172
AACGCTGTG**TATCGT**AACGGTAACACTGTA**GAGG**GGAGCACATTGATGAGCACGTCAGAC*GATATC*CATAAC
M S T S D D I H N

244
ACCACAGCCACTGGCAAATGCCCGTTCCATCAGGGCGGTCACGACCAGAGTGCGGGGGCGGGCACAACCAC*T*
T T A T G K C P F H Q G G H D Q S A G A G T T T
10
NruI BalI 316
*CGCGA*CTGG*TGGCCA*AATCAACTTCGTGTTGACCTGTTAAACCAACATTCTAATCGTTCTAACCCACTGGGT
R D W W P N Q L R V D L L N Q H S N R S N P L G
34
EcoRI BglII 388
GAGGACTTTGACTACCGCAAA*GAATTC*AGCAAATTAGATTACTACGGCCTGAAAAAA*AGATCT*GAAAGCCCT*G*
E D F D Y R K E F S K L D Y Y G L K K D L K A L
58
HincII BalI 460
*TTGAC*AGAATCTCAACCGTGG*TGGCCA*GCCGACTGGGGCAGTTACGCCGGTCTGTTTATTCGTATGGCCTGG
L T E S Q P W W P A D W G S Y A G L F I R M A W
82
ClaI 532
CACGGCGCGGGGACTTACCGTTCA*ATCGAT*GGACGCGGTGGCGCGGGTCGTGGTCAGCAACGTTTTGCACCG
H G A G T Y R S I D G R G G A G R G Q Q R F A P
106
BalI 604
CTGAACTCCTGGCCGGATAACGTAAGCCTCGATAAAGCGCGTCGCCTGTTG*TGGCCA*ATCAAACAGAAATAT
L N S W P D N V S L D K A R R L L W P I K Q K Y
130
676
GGTCAGAAAATCTCCTGGGCCGACCTGTTTATCCTCGCGGGTAACGTGGCGCTAGAAAACTCCGGCTTCC*GT*
G Q K I S W A D L F I L A G N V A L E N S G F R
154
RsaI HincII 748
*AC*CTTCGGTTTTGGTGCCGGTCGTGAAGACGTCTGGGAACCGGATCTGGAT*GTTAAC*TGGGGTGATGAAAAA
T F G F G A G R E D V W E P D L D V N W G D E K
178
820
GCCTGGCTGACTCACCGTCATCCGGAAGCGCTGGCGAAAGCACCGCTGGGTGCAACCGAGATGGGTCTGATT
A W L T H R H P E A L A K A P L G A T E M G L I
202
HincII 892
TAC*GTTAAC*CCGGAAGGCCCGGATCACAGCGGCGAACCGCTTTCTGCGGCAGCAGCTATCCGCGCGACCTTC
Y V N P E G P D H S G E P L S A A A A I R A T F
226
964
GGCAACATGGGCATGAACGACGAAGAAACCGTGGCGCTGATTGCGGGTGGTCATACGCTGGGTAAAACCCAC
G N M G M N D E E T V A L I A G G H T L G K T H
250

1036
GGTGCCGGTCCGACATCAAATGTAGGTCCTGATCCAGAAGCTGCACCGATTGAAGAACAAGGTTTAGGTTGG
G A G P T S N V G P D P E A A P I E E Q G L G W
274

1108
GCGAGCACTTACGGCAGCGGCGTTGGCGCAGATGCCATTACCTCTGGTCTGGAAGTAGTCTGGACCCAGACG
A S T Y G S G V G A D A I T S G L E V V W T Q T
298

RsaI 1180
CCGACCCAGTGGAGCAACTATTTCTTCGAGAACCTGTTCAAGTATGAGTGG*GTAC*AGACCCGCAGCCCGGCT
P T Q W S N Y F F E N L F K Y E W V Q T R S P A
322

AsuII AccI BamHI 1252
GGCGCAATCCAG*TTCGAAGCG**GTAGAC*GCACCGGAAATTATCCC*GGATCC*GTTTGATCCGTCGAAGAAACGT
G A I Q F E A V D A P E I I P D P F D P S K K R
346

BglII 1324
AAACCGACAATGCTGGTGACCGACCTGACGCTGCGTTTTGATCCTGAGTTCGAGA*AGATCT*CTCGTCGTTTC
K P T M L V T D L T L R F D P E F E K I S R R F
370

1396
CTCAACGATCCGCAGGCGTTCAACGAAGCCTTTGCCCGTGCCTGGTTCAAACTGACGCACAGGGATATGGGG
L N D P Q A F N E A F A R A W F K L T H R D M **G**
394

BglII 1468
CCGAAATCTCGCTACATCGGGCCGGAAGTGCCGAAAGA*AGATCT*GATCTGGCAAGATCCGCTGCCGCAGCCG
P K S R **Y I G P E V P K E** D L I W Q D P L P Q P
418

ClaI NruI 1540
ATCTACAACCCGACCGAGCAGGACATT*ATCGAT*CTGAAAT*TCGCGA*TTGCGGATTCTGGTCTGTCTGTTAGT
I Y N P T E Q D I I D L K F A I A D S G L S V S
442

1612
GAGCTGGTATCGGTGGCCTGGGCATCTGCTTCTACCTTCCGTGGTGGCGACAAACGCGGTGGTGCCAACGGT
E L V S V A W A S A S T F R G G D K R G G A N G
466

1684
GCGCGTCTGGCATTAATGCCGCAGCGCGACTGGGATGTGAACGCCGCAGCCGTTCGTGCTCTGCTTGTTCTG
A R L A L M **P Q** R **D W D V N A A A V** R **A** L L V L
490

EcoRV 1756
GAGAAAATCCAGAAAGAGTCTGGTAAAGCCTCGCTGGCG*GATATC*ATAGTGCTGGCTGGTGTGGTTGGTGTT
E K I Q K E S G K A S L A D I I V L A G V V G V
514

RsaI 1828
GAGAAAGCCGCAAGCGCCGCAGGTTTGAGCATTCAT*GTAC*CGTTTGCGCCGGGTCGCGTTGATGCGCGTCAG
E K A A S A A G L S I H V P F A P G R V D A R Q
538

1900
GATCAGACTGACATTGAGATGTTTGAGCTGCTGGAGCCAATTGCTGACGGTTTCCGTAACTATCGCGCTCGT
D Q T D I E M F E L L E P I A D G F R N Y R A R
562

1972
CTGGACGTTTCCACCACCGAGTCACTGCTGATCGACAAAGCACAGCAACTGACGCTGACCGCGCCGGAAATG
L D V S T T E S L L I D K A Q Q L T L T A P E M
586

SphI RsaI
2044
ACTGCGCTGGTGGGCG*GCATGC*GT*GTAC*TGGGTGGCAACTTCGATGGCAGCAAAAACGGCGTCTTCACTGAC
T A L V G G M R V L **G A** N F **D G S K N** G V F T D
610

2116
CGCGTTGGCGTATTGAGCAATGACTTCTTCGTGAACTTGCTGGATATGCGTTACGAGTGGAAAGCGACCGAC
R V G V L S N D F F V N L L D M **R Y E W K** A **T D**
634

AsuII HaeIII
2188
GAATCGAAAGAGCTG*TTCGAA*GGCCGTGACCGTGAAACCGGCGAAGTGAAATTTAC*GGCC*AGCCGTGCGGAT
E S K E L F E G R D R E T G E V K F T A S R A D
658

2260
CTGGTGTTTGGTTCTAACTCCGTCCTGCGTGCGGTGGCGGAAGTTTACGCCAGTAGCGATGCCCACGAGAAG
L V F G S N S V L R A V A E V Y A S S D A H E K
682

2332
TTTGTTAAAGACTTCGTGGCGGCATGGGTGAAAGTGATGAACCTCGACCGTTTCGACCTGCTGTAATCTGAC
F V K D F V A A W V K V M N L D R F D L L END
706

2404
CC**CGTTCAGCGGCTGCTTGCTGGCAGTCGCTGAACG**TTCTTTACCAGCGTATAGTGGGCGAACGAAAACTAC

2476
ACACTGGATCTCTCATGTCTGCCGCAGGAAAGAGCAACCCACTGGCAATCAGTGGCCTGGTTGTGCTCACAC

2548
TTATCTGGAGTTATAGCTGGATTTCATGAAGCAAGTCACCAGTTACATCGGTGCCTTCGACTTTACCGCCTT

2620
ACGCTGCATTTTCGGCGCTCTCGTTTTATTCATCGTCCTTTTATTACGTGGTCGCGGAATGCGCCCGACACC

DraI
G*TTTAAA*

Microgenie. No homologous sequences were found, confirming this sequence to be novel. The DNA sequence of both *katG* and the surrounding region was compared to the sequence of *oxyR* (Gisela Storz, personal communication) but no homology with this sequence was identified.

4.10.3 Control Sequences of the *katG* Gene

Sequences typical of *E. coli* transcriptional and translational control elements were found in close proximity to the *katG* coding sequence. The potential control sequences presented in this section are illustrated in Figure 24 with the appropriate *E. coli* consensus sequences for comparison.

As discussed in section 2.11, sequences upstream from the point of transcription initiation are required for RNA polymerase binding. These include the Pribnow box or -10 sequence (TATAAT) and the -35 sequence (TTGACA) which are normally separated by 16-18 bp. Initiation of transcription often occurs at an A that may be the central base of a CAT triplet 5-9 bp downstream from the -10 sequence (Harley and Reynolds, 1987). Although the transcriptional start site of the *katG* mRNA had not been experimentally determined, the promoter was known to be within the 322 bp *Bgl*II fragment previously shown to contain the promoter. Using the *Bgl*II site at 55 bp and the protein start codon at 146 bp as boundaries, the sequence showing the best homology to the promoter consensus sequence was found to start 75 bp into the sequence. A potential -35 sequence of TTATAA was located 18 bp upstream from a potential -10 sequence of TATCGT which in turn was located 9 bp upstream of an A residue centred in a CAC triplet. These sequences fit well with the consensus sequences, having all of the highly conserved nucleotides, (bases conserved in more than 75% of *E. coli* promoter sequences, [Harley

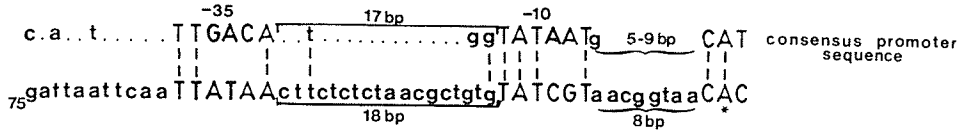
Figure 24. The control sequences of the *katG* gene.

a) Promoter sequence. The potential *katG* promoter sequence is shown below the *E. coli* consensus promoter sequence with dashed lines connecting the conserved bases. The portion of the *E. coli* consensus sequence which is weakly conserved is represented by small letters, while that which is strongly conserved is represented by capital letters. The starting nucleotide of the *katG* sequence shown in this figure is numbered on the left hand side. The asterisk denotes the base at which transcription is expected to start.

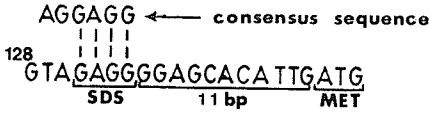
b) Shine-Dalgarno Sequence (SDS). The potential *katG* Shine-Dalgarno sequence is shown below the *E. coli* consensus sequence and the conserved bases are connected with slashed lines. The first nucleotide of the *katG* sequence is numbered on the left hand side.

c) The Termination Sequence. A potential rho independent termination sequence for the *katG* gene is shown with the secondary structure the RNA would be expected to form underneath. The palindromic portion of the sequence in underlined and the first nucleotide shown of the *katG* sequence is numbered on the left hand side.

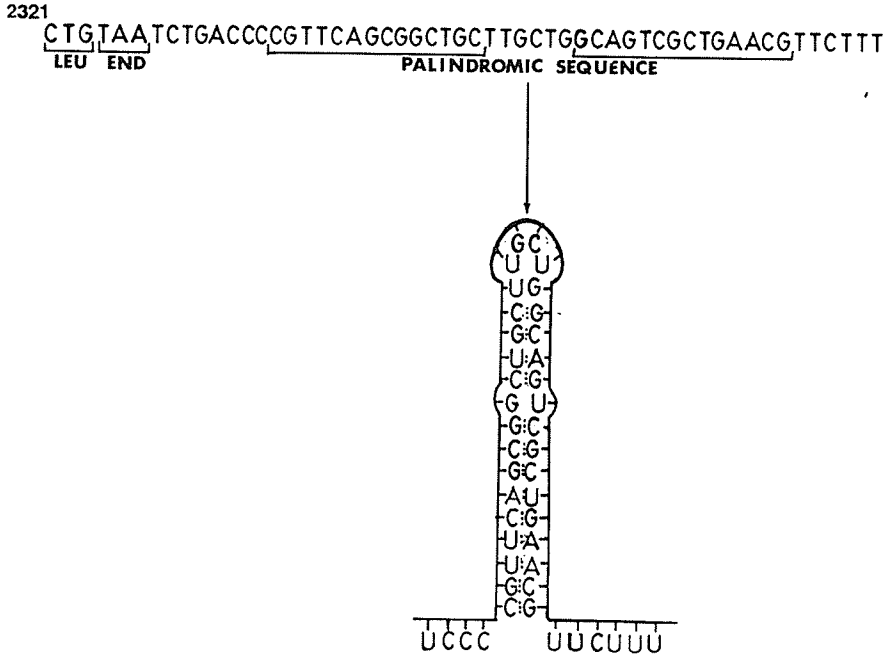
(a) Promoter Sequence



(b) Shine-Dalgarno Sequence



(c) Termination Sequence



and Reynolds, 1987]) conserved in this sequence. This included the two T's in the -35 sequence and the TA...T portion of the -10 sequence. An A was identified as a potential transcription start point 8 bp downstream from the -10 sequence, centred in a CAC triplet which showed good homology to the CAT triplet which has been weakly associated with transcriptional start points (Harley and Reynolds, 1987). Weakly conserved bases (occuring in 37% or less of *E. coli* promoters, [Harley and Reynolds, 1987]) are also shown in Figure 24. Two of these seven bases are conserved in the potential *katG* promoter.

A consensus sequence complementary to the 3' end of the 16S rRNA, called the Shine-Dalgarno sequence, is required for translational initiation. This consensus sequence of AGGAGG is found 3-12 bp before the first amino acid of the protein (Glover, 1984). A 9 bp polypurine region containing a sequence homologous to the Shine-Dalgarno consensus sequence was found 11 bp upstream of the initiating methionine of the HPI protein (see Figure 24). The sequence surrounding the potential Shine-Dalgarno sequence in Figure 23 forms a hyphenated dyad of CACTGT-ACATTG with a 1 bp mismatch. The purpose of this potential pairing sequence is unknown since Ganoza et al.(1987) have found that most sequences 5' to the start codon have little pairing potential as this would likely interfere with translation.

E.coli transcriptional termination sequences show no real consensus sequence at a primary level but they do have consensus at the level of secondary structure. All *E. coli* termination sequences contain a palindromic sequence, often G-C rich, which could allow stable base pairing of the resulting mRNA. Termination may be rho dependent or independent, but typically the hyphenated dyads of rho independent

terminators are followed by a polyT tail (Rosenberg and Court, 1979). The *katG* sequence has a potential rho independent terminator consisting of a 14 bp hyphenated dyad 8 bp beyond the end of the coding sequence followed by a 6 bp stretch containing 5 U's. The palindrome contains only a single base mismatch and a predominance of GC bases which would allow stable stem and loop formation as shown by the model in Figure 24. Based on the calculations of Borer et al. (1974), a stability of ΔG=-21.4 kcal can be calculated confirming the extreme stability of the palindrome. Termination of the RNA transcript always occurs 16-24 bp downstream from the center of the hyphenated dyad (Rosenberg and Court, 1979). The *katG* mRNA would be expected to terminate between base 2367 and 2375 in Figure 23.

4.10.5 Confirmation of the Predicted HPI Amino Acid Sequence

4.10.5.1 Sequencing of Cyanogen Bromide Peptides from HPI

To confirm the 726 amino acid sequence predicted for the HPI protein a portion of the HPI amino acid sequence had to be determined. Two attempts to sequence the N-terminus of the protein were unsuccessful, suggesting that the N-terminus was blocked as a result of some type of modification. The identity of the blocking group was unknown, but N-terminal blocking groups have been reported in bovine liver catalase. The group blocking the N-terminus of bovine liver catalase has not been identified even with the help of x-ray crystallography and positive and negative ion mass spectrometry (Schroeder et al., 1982a).

Peptides with amino termini accessible for amino acid sequencing were generated by cyanogen bromide cleavage of the HPI protein. Six peptides labelled Ia, Ic, 3, III, 4 and IV, were sequenced, and the amino acids were identified in as many reaction cycles as possible (see Table

21). From the six sequenced HPLC peaks, seven polypeptide sequences were identified in the HPI predicted sequence. The extra identifiable polypeptide was the result of a mixture of polypeptides in peak IV which were separable into two peptide sequences based on knowledge of the DNA sequence.

The amino acids in Table 21 which confirm portions of the predicted HPI sequence are underlined as are the confirmed portions of the HPI predicted sequence in Figure 23. Unfortunately, the small quantity of the peptides and the presence of more than one polypeptide in some of the peaks made the amino acid sequence difficult to interpret. Specifically, amino acids which were of low yield during the final hydrolysis in the reaction, serine and threonine, were particularly difficult to identify. Amino acids which did not confirm the predicted amino acid sequence were rechecked and in many cases the inconsistencies could be accounted for. In no instance where there were inconsistencies unaccounted for was there any evidence that the predicted sequence was incorrect.

The identification of all the peptides in the HPI sequence confirmed that the reading frame was correct. Although there were some differences between the peptide sequences and the predicted amino acid sequence, the fact that amino acids could be confirmed in cycles following the inconsistencies showed that the reading frame was correct.

4.10.5.2 Molecular Weight Determination and Amino Acid Composition

The molecular weight of the predicted HPI sequence was calculated to be 80 049, consistent with the relative molecular weights of 84 000 (Loewen et al., 1985b) and 78 000 (Claiborne and Fridovich, 1979) estimated by SDS-PAGE.

Table 21. Amino acid sequence of cyanogen bromide generated peptides of HPI

HPLC Peak	Amino Acids Identified from Cycles																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Ia	G A*	R	F H	G	A	G	P	Y	G	P	I K								
Ic[a]	G A Y	Y L P	F K I	G	V P	S	M V P												
3-M[b]	R S	Y	E N	G	F K	G	T	D	G	S	K I	E							
3-m[c]	G	N	T	W	V A	P E V	P N E	Y	K I F E	P	P	-							
III-M	S	Y	N	G	F	G	T	D	G	S[d]	K I	E							
III-m	G	N	T	W	V A	P E V	P N E	Y	K I F E	P	P	-							
4	N	D	E	E	T	V	A	L	I	A	G	G	H	T	L	G	K I	T	H
IV-M	G	E N V P	K	G A W	A	Y	F I	G	P	E N	V	P	I K L	E N					
IV-m	V P	Q	L	D	W	D	V P	E N	A	A	K L I A	V	G	A					

[a] extra tyrosine present throughout the analysis
[b] major amino acids
[c] minor amino acids
[d] extra alanine present
All amino acids confirming the HPI predicted sequence are underlined
* Cycles where a) more than one amino acid was identified in cycles if more than one amino acid was abundant in comparison to the previous cycle and/or where b) the amino acid peak in the cycle could be one of two or three amino acids having similar residence times.

The amino acid composition of the predicted HPI sequence is shown in Table 22 with the amino acid compositions previously determined for HPI catalase (Claiborne and Fridovich, 1979) and bovine liver catalase (Schroeder et al., 1982). Fridovich's determination of the HPI amino acid composition was relatively consistent with that of the predicted HPI protein. There were obvious differences in the serine, threonine, aspartate/asparagine and tryptophan values. The method of determination of the serine and threonine values could explain their difference. Because serine and threonine are destroyed during the hydrolysis of the protein in preparation for the determination of the amino acid composition, serine and threonine content must be determined at various times of hydrolysis and then the original content determined by extrapolation back to zero time. Tryptophan content must also be determined separately, which could have led to the high value. The reason for the low aspartate/asparagine value is unknown.

The amino acid composition of the HPI protein was also compared to the bovine liver catalase amino acid composition. The compositions were similar, but there were some notable differences. The histidine and tyrosine content was much higher in the bovine liver catalase, while the alanine and leucine content was much higher in the HPI protein.

4.10.6 Codon Usage in katG

Codon usage is organism specific, reflecting the isoaccepting tRNA

Table 22. Amino acid composition of catalase proteins.

Amino Acid	Predicted HPI Sequence #	Predicted HPI Sequence %	Fridovich[a] HPI #	Fridovich[a] HPI %	Bovine Liver[b] Catalase %
Lysine	35	(4.8)	38.4	(5.1)	5.3
Histidine	13	(1.8)	8.9	(1.3)	4.0
Arginine	41	(5.6)	44.5	(5.9)	6.1
Aspartic Acid[c]	83	(11.5)	57.7	(7.7)	13.6
Threonine	39	(5.4)	27.6	(3.7)	4.3
Serine	44	(6.1)	100.1	(13.3)	4.9
Glutamic Acid[e]	68	(9.4)	77.2	(10.3)	9.1
Proline	39	(5.4)	38.8	(5.2)	7.5
Glycine	65	(8.9)	60.7	(8.1)	6.7
Alanine	77	(10.6)	67.8	(9.0)	6.9
Cysteine	1	(0.1)	0	(0.0)	0.8
Valine	43	(5.9)	46.2	(6.1)	6.7
Methionine	13	(1.8)	11.6	(1.5)	2.0
Isoleucine	28	(3.9)	23.6	(3.1)	3.6
Leucine	64	(8.8)	60.3	(8.0)	6.9
Tyrosine	15	(2.1)	15.1	(2.0)	4.0
Phenylalanine	36	(5.0)	31.3	(4.2)	6.1
Tryptophan	22	(3.0)	42.0	(5.6)	1.2
Total	726		751.8		

[a] Claiborne and Fridovich, 1979
[b] Schroeder et al., 1982
[c] aspartic acid and asparagine are both included in this value
[d] glutamic acid and glutamine are both included in this value

population in the organism (Ikemura and Ozeki, 1982). In yeast and *E. coli*, it has also been shown that the level of expression of a gene is reflected in the codon usage (Bennetzen and Hall, 1982; Grosjean and Fiers, 1982).

Grosjean and Fiers (1982) have found that the codon usage in highly expressed *E. coli* genes is very biased. Codons corresponding to low abundance isoaccepting tRNA's such as AUA for isoleucine, CGG/AGA/AGG/CGA for arginine, CUA for leucine, and GGA (and possibly GGG) for glycine are clearly avoided. In contrast, weakly expressed genes exhibit a random use of codons. Highly expressed *E. coli* genes also optimize the codon-anticodon interaction energies through bias in codon usage. The codons marked by an asterisk in Table 23 are those whose preferential usage is likely to optimize the codon-anticodon interaction energies.

The codon usage of *katG* is given in Table 23 and it is clear that the usage of codons whose isoaccepting tRNA's are in low abundance is avoided. The preferential usage of most of the other codons was typical of highly expressed genes. Two exceptions to the general trend of high expression codon usage were the preferential use of GCC over GCT for alanine and GAT over GAC for aspartate. Overall the codon bias of *katG* was typical of a highly expressed gene. It is thought that codon usage may be a way of fine tuning the rate of translation; the preferential use of GCC over GCT for alanine in HPI (HPI having a high alanine content) may be involved in the fine tuning of HPI translation.

4.10.7 Comparison of HPI to Other Catalase and Peroxidase Proteins

The predicted amino acid sequence of HPI was compared to the amino acid sequences of BLC (bovine liver catalase)(Schroeder et al., 1982),

Table 23. Codon usage in *katG*

TTT	Phe	13	(1.8)	**TCT**	**Ser**	**13**	**(1.8)**	TAT	Tyr	4	(0.6)	TGT Cys 0 (0.0)
TTC	**Phe**	**23**	**(3.2)**	TCC	Ser	5	(0.7)	**TAC**	**Tyr**	**11**	**(1.5)**	**TGC Cys 1 (0.1)**
TTA	Leu	4	(0.6)	TCA	Ser	4	(0.6)	TAA	End	1	(0.1)	TGA End 0 (0.0)
TTG	Leu	5	(0.7)	TCG	Ser	4	(0.6)	TAG	End	0	(0.0)	TGG Trp 22 (3.1)
CTT	Leu	3	(0.4)	**CCT***	**Pro**	**2**	**(0.3)**	CAT	His	6	(0.8)	**CGT*** Arg 28 (3.9)
CTC	**Leu**	**4**	**(0.6)**	CCC*	Pro	0	(0.0)	**CAC**	**His**	**7**	**(1.0)**	CGC* Arg 12 (1.7)
→CTA	Leu	1	(0.1)	CCA	Pro	6	(0.8)	CAA	Gln	7	(1.0)	→CGA Arg 0 (0.0)
CTG	Leu	47	(6.5)	CCG	Pro	31	(4.3)	CAG	Gln	17	(2.3)	→CGG Arg 0 (0.0)
ATT	Ile	11	(1.5)	ACT	Thr	8	(1.1)	AAT*	Asn	4	(0.6)	AGT Ser 4 (0.6)
ATC*	**Ile**	**16**	**(2.2)**	**ACC**	**Thr**	**19**	**(2.6)**	AAC*	**Asn**	**24**	**(3.3)**	**AGC Ser 14 (1.9)**
→ATA	Ile	1	(0.1)	ACA	Thr	5	(0.7)	AAA	Lys	31	(4.3)	→AGA Arg 0 (0.0)
ATG	Met	13	(1.8)	ACG	Thr	7	(1.0)	AAG	Lys	4	(0.6)	→AGG Arg 1 (0.1)
GTT	**Val**	**14**	**(1.9)**	**GCT***	**Ala**	**8**	**(1.1)**	GAT	Asp	29	(4.0)	**GGT*** Gly 36 (5.0)
GTC	Val	4	(0.6)	GCC*	Ala	24	(3.3)	**GAC**	**Asp**	**26**	**(3.6)**	GGC* Gly 24 (3.3)
GTA	Val	9	(1.2)	GCA	Ala	16	(2.2)	GAA	Glu	26	(3.6)	→GGA Gly 1 (0.1)
GTG	Val	16	(2.2)	GCG	Ala	29	(4.0)	GAG	Glu	18	(2.5)	→GGG Gly 4 (0.6)

The codon pairs which are in boxes are those which show a bias in usage in highly expressed genes. The boldfaced codons are the ones which are used most frequently in highly expressed genes. Arrows denote codons that are rarely used in highly expressed genes. Asterisks mark codon pairs where the bias in codon usage in highly expressed genes is thought to optimize codon-anticodon interaction energies. The codon usage bias used in this analysis was based on the study of Grosjean and Fiers (1982), where the codon usage of 23 abundantly expressed proteins and 4 proteins of low abundance were compared.

human catalase (Quan et al., 1986; Bell et al., 1986), yeast catalase T (Hartig and Ruis, 1986), PVC (*Penicillium vitale* catalase)(Vainshtein et al., 1986), and human myeloperoxidase (Johnson et al., 1987) both manually and by computer. The best possible computer alignments between the HPI protein and the other proteins showed little homology. Homology was manually searched for around residues which were potential ligands for the heme group (ie. His and Tyr). His and Tyr residues were found by x-ray crystallography of BLC (Fita and Rossmann, 1985b) and PVC (Vainshtein et al., 1986) to be intimately involved in heme binding and catalysis. Essential residues included the BLC residues predicted to play a role in catalysis, His-74, Ser-113, and Asn-147 as well as several BLC residues thought to interact with the heme group (Pro-335; Arg-353; Tyr-357; Val-73; Thr-114; Phe-152). Of the heme binding residues, the region around Tyr-357, the fifth ligand of the heme Fe was the most likely region to be conserved. The regions around Tyr and His (potential ligand replacement) residues were thoroughly searched for homologies, but the homologies were weak and did not extend to any extent into the rest of the protein. All of the catalytic residues and most of the heme binding residues had been conserved in all of the previously sequenced catalases, including yeast catalase T (Hartig and Ruis, 1986) and PVC (Vainshtein et al., 1986), but not in HPI. Consequently, the HPI protein appears to be a novel catalase with no homology to any previously sequenced catalases or peroxidases.

The larger size of the HPI catalase was comparable to that of the 670 amino acid PVC which has a flavodoxin-like domain occupying the latter portion of the protein (Vainshtein et al., 1986). The flavodoxin-like domain of the PVC may give it a second function. The larger size of the HPI catalase may also be related to its dual function as both a

catalase and a peroxidase.

Hydrophilicity plots of the amino acid residues of HPI and BLC are shown in Figure 25. A comparison of the plots showed some common features in the two proteins. In both proteins the largest hydrophobic regions were found approximately 300 residues into the protein. This large hydrophobic region in BLC was part of the β-barrel domain, extending from residue 56 to 376. The fifth ligand Tyr-357 of BLC was located just beyond the large hydrophobic region. Tyr-336 of HPI, located near to the end of the hydrophobic region, showed the best (although weak) homology to the Tyr-357 region of BLC. The homology did not extend in

```
BLC  PDKMLQGRLFAYPDTHRHRLGPNYLQIPVNCP
HPI  WSWYFFENLFKYEWVQTRSPAGAIQFEAVDAP
              ** *                * *
```

either direction beyond this. It is possible that the Tyr-336 is a ligand of the HPI heme, although the surrounding homology is weak.

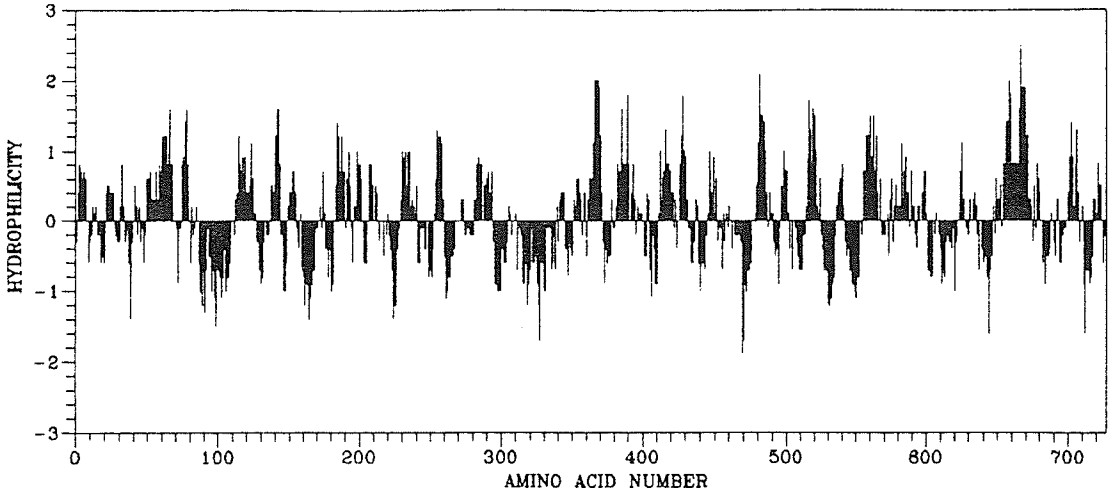
The hydrophilicity plot showed a large number of hydrophilic residues near the C-terminus of the HPI protein. The hydrophilic portions of a protein are usually the antigenic determinants (Hopp and Woods, 1981). The absence of the coding region for the C-terminal portion of the HPI protein in pBT5 (*SphI* clone) may explain the absence of immunoprecipitable protein in crude protein preparations from UM262/pBT5.

4.10.8 Cloning and Sequencing of Potential *katG* Mutant Promoters

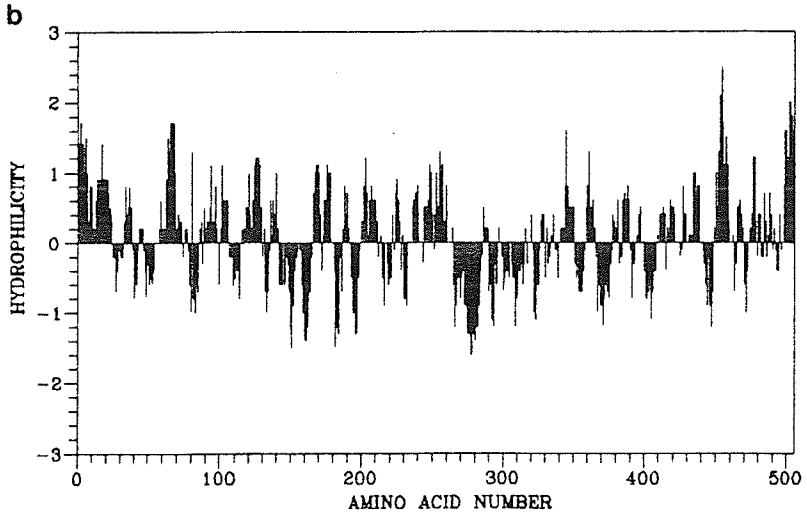
Studies of the *katG* mutants UM1(*katG14*), UM2(*katG15*), and UM56-64 (*katG16*) implicated the *katG* mutations in UM2 and UM56-64 as potential regulatory mutants. Although all of these mutants had low levels of catalase, 84 kd proteins could be immunoprecipitated from crude extracts

Figure 25. Hydrophilicity plots of a) the *E.coli* HPI catalase and b) bovine liver catalase. Hydrophilic amino acids are represented by positive values while hydrophobic amino acids are represented by negative values. A point has been plotted for each amino acid using the hydrophilicity predictions generated by computer (Microgenie). The computer algorithm on which the prediction was based has been described by Hopp and Woods (1981).

a


HYDROPHILICITY
3
2
1
0
-1
-2
-3
0
100
200
300
400
500
600
700
AMINO ACID NUMBER

b


HYDROPHILICITY
3
2
1
0
-1
-2
-3
0
100
200
300
400
500
AMINO ACID NUMBER

of these mutants with HPI antisera. Each of the mutants displayed a different response to ascorbate induction. UM1, which had peroxidase but not catalase activity, did show an increase in peroxidase activity upon ascorbate induction. UM2 had only a small increase in the production of immunoprecipitable protein, lower than would be expected in a normal ascorbate induction. In contrast, UM56-64 showed an abnormally high catalase induction of 25 fold by ascorbate, suggesting that it too had abnormal regulation (Loewen et al., 1985b). These genes were cloned into pAT153 to create the plasmids pMS*katG14*, pMS*katG15*, and pMS*katG16* and characterized (M. Smolenski, 1986).

*Bgl*II fragments containing the promoter regions from each of the three mutant *katG* clones, pMS*katG14*, pMS*katG15*, and pMS*katG16* were cloned and sequenced. The promoter sequences were identical to the wild type promoter, showing that none of the abnormal regulatory features of these mutant genes were the result of the promoters. A positive regulator of *katG*, the *oxyR* protein (Christman et al., 1985) has been shown to interact with the promoter region of *katG* (Bruce Ames, personal communication) and it is likely that the *katG15*(UM2), and *katG16*(UM56-64) strains contain *oxyR* mutations in addition to *katG* mutations and are not regulatory mutants.

4.10.9 Beyond *katG*

An 884 bp sequence beyond the apparent termination of the *katG* mRNA has been included in Figure 26 although the entire region has not been sequenced in both directions. This sequence has been compared by computer to the data base of sequences called GenBank (1986). As no homologous sequences were found, it is likely that this sequence has not been identified and sequenced nor is it similar to any other known genes. This sequence did

Figure 26. The DNA and predicted amino acid sequences of the region sequenced beyond the termination sequence of *katG*. The nucleotides are numbered on the right hand side of the page and the amino acids are numbered on the left hand side of the page. The nucleotide sequences recognized by restriction sites mapped in pBT22 are italicized and labelled.

```
                                                                    2423
                    TTCTTTACCAGCGTATAGTGGGCGAACGAAAACTACACACTGGATCTCTCATGTC

                                                                    2495
TGCCGCAGGAAAGAGCAACCCACTGGCAATCAGTGGCCTGGTTGTGCTCACACTTATCTGGAGTTATAGCTG

                                                                    2567
GATTTCATGAAGCAAGTCACCAGTTACATCGGTGCCTTCGACTTTACCGCCTTACGCTGCATTTTCGGCGCT
       M  K  Q  V  T  S  Y  I  G  A  F  D  F  T  A  L  R  C  I  F  G  A

                                                        DraI        2639
CTCGTTTTATTCATCGTCCTTTTATTACGTGGTCGCGGAATGCGCCCGACACCGTTTAAATACACCTTAGCC
 L  V  L  F  I  V  L  L  L  R  G  R  G  M  R  P  T  P  F  K  Y  T  L  A
23

                                                                    2711
ATTGCCCTGTTACAAACCTGCGGGATGGTTGGTCTGGCGCAGTGGGCGTTGGTCAGCGGAGGTGCGGGGAAG
 I  A  L  L  Q  T  C  G  M  V  G  L  A  Q  W  A  L  V  S  G  G  A  G  K
47

                                                                    2783
GTGGCGATCCTGAGCTATACCATGCCGTTCTGGGTGGTGATTTTCGCCGCGTTGTTTCTCGGTGAACGCCTG
 V  A  I  L  S  Y  T  M  P  F  W  V  V  I  F  A  A  L  F  L  G  E  R  L
71

            SspI NruI                                               2855
CGACGTGGGCAATATTTCGCGATTCTGATTGCCGCTTTCCGCTTAATTTTTGGTGTTGCAGCCGTGGCAACT
 R  R  G  Q  Y  F  A  I  L  I  A  A  F  R  L  I  F  G  V  A  A  V  A  T
95

                                                                    2927
CGATTTCTCTTCGATGAAAAGTGCCATGCTGGCAATCCTCTCCGGCGTCAGTTGGGGGGCGAGCGCGATTCG
 R  F  L  F  D  E  K  C  H  A  G  N  P  L  R  R  Q  L  G  G  E  R  D  S
119

                                                        HincII       2999
ATTGCTAAACGTCTGTTATGCCCGTCATCCGCGCGTGGATTTATTGTCGTTAACATCCTGGCAGATGCTGAC
 I  A  K  R  L  L  C  P  S  S  A  R  G  F  I  V  V  N  I  L  A  D  A  D
143

                                                                    3071
GCGGCGCTGGCGATGAGTGTGGTCGCTTTTACTGGTGCCGCAACGTGAAATTGACTGGCAGCCCACCGTGTT
 A  A  L  A  M  S  V  V  A  F  T  G  A  A  T End
167

                                                                    3143
CTGGGCGCTTGGCCTACAGTGCGATTATGGCGACAGGCAXXTGGCGTGGAGCTTATGGXGGXGXAXGAAAAA

                                                                    3215
TTAGCCAGCCAGTAATGCCAGTTTAAGCATATTGGCGAAACGTTGCGGAGTGACGGXXXXCXGGXGGCXGCX

XGGCGAGAAXXCGAGGGGXCGXXGAAGGGACXGGXXA
```

however contain another open reading frame that extended for 181 amino acids and was preceded by a potential Shine-Dalgarno sequence, GGA. The termination codon may not be valid because the sequence at this point was deteriorating, and there was only a poor potential terminator, a 5 bp hyphenated dyad. Upstream from the open reading frame were potential promoter sequences, although none had extremely good homology to the consensus sequence.

DISCUSSION

5.0 DISCUSSION

Three transposon Tn*10* mutations affecting catalase activity in *E. coli* were identified and characterized. Two of the mutants had low catalase levels on agar plates, were ascorbate inducible, and lacked HPII bands on polyacrylamide gels stained for catalase. These characteristics were typical of mutations previously mapped as the loci *katE* (Loewen, 1984) and *katF* (Loewen and Triggs, 1984). The mutant loci were mapped by Hfr crosses and P1 transduction as *katE12*::Tn*10* at 37.8 min and *katF13*::Tn*10* at 59 min on the *E. coli* chromosome. The third mutation, *kat17*::Tn*10*, was selected for its low catalase activity in a strain that already lacked HPII (*katE*). *kat17*::Tn*10* was no longer ascorbate inducible, and it lacked the bifunctional HPI-A and HPI-B bands on polyacrylamide gels stained for catalase or peroxidase. This mutation was identified by P1 transduction to be *katG17*::Tn*10* and mapped to 89.2 min on the *E. coli* chromosome between the markers *argH* and *metB*.

The three catalase loci, *katE*, *katF*, and *katG* are widely separated on the *E. coli* chromosome and obviously do not form an operon. Loewen et al. (1985b) suggested that the *katG* and *katE* genes which are separated by about 50 min on the *E. coli* chromosome map may have resulted from a duplication of the *E. coli* chromosome early in evolution. It is believed that two entire duplications of the *E. coli* chromosome may have placed related loci 25 or 50 minutes away from each other. Indeed other duplicated genes, such as *pfkA* and *pfkB* are clustered 25 or 50 min away from each other on the *E. coli* chromosome (Riley and Anilionis, 1978). Recently, *katF* has been identified as a regulatory protein required for the production of HPII catalase (M. Mulvey, unpublished

results), suggesting that *katE* is the structural gene for HPII catalase. If this is the case, then the theory that the two catalase genes, *katE* and *katG* may have evolved from one ancestral gene is plausible. However, under stringent hybridization conditions, *katG* did not hybridize with any other *E. coli* genes, showing that the relationship between HPI and HPII may be so distant that any homology may be evident only after comparison of the protein sequences of these two catalases.

No other loci affecting catalase activity in *E. coli* were identified in our lab, although *katA-D* have been identified in *S. typhimurium* by Levine (1977). Of these loci, only one, *katC*, was mapped by P1 transduction to 7 min, and at present no equivalent locus has been identified in *E. coli*.

HPI (KatG protein) has recently been identified as a part of an oxidative regulon controlled by the positive regulatory protein, OxyR (Christman et al., 1985). This regulon can be induced by a number of oxidative chemicals, including H_2O_2, accounting for the ascorbate induction of the transposon mutants which still contained HPI catalase. The *katE* and *katF* loci affecting HPII catalase in *E. coli* are regulated independently. HPII is not ascorbate inducible, nor is it under the control of the positive regulator OxyR. Normally, HPI is constitutively produced during log phase growth, but it is inducible by H_2O_2 to higher levels (Loewen et al., 1985a). Morgan et al. (1986) have shown that the induction of the HPI protein occurs at the level of transcription, with *oxyR1* mutants (mutants where OxyR protein is overproduced) having approximately 50 fold more *katG* mRNA and 50 fold more catalase than normal. As discussed in the historical, glucose also affects the *katG* mRNA levels, but the mechanism of action remains obscure since the

available evidence suggests that it is not via classical catabolite repression. This information indicates that *katG* is under both positive and negative control and is inducible to levels at least 50 fold higher than normal non-induced catalase levels.

The *oxyR* locus has been mapped to 88 min in *S. typhimurium*, near to the *katG* gene. A deletion in *S. typhimurium* extending from *metB* to *argH* eliminated the HPI-A and HPI-B proteins but left the *oxyR* protein (Christman et al., 1985), showing that these genes do not form an operon. However, the mapping of the *E. coli oxyR* gene in our lab (unpublished results) suggested that *oxyR* was closer to *katG* than either *metB* or *argH*. Another difference in the *oxyR* regulon of *E. coli* and *S. typhimurium* has already been identified, that is the requirement of both *oxyR* and *htpR* for the induction of the novel alkylhydroperoxidase (*ahp*) by heat shock in *S. typhimurium* but not in *E. coli* (Morgan et al., 1986). Consequently, the placement of *oxyR* in a position closer to *katG* on the *E. coli* chromosome can not be eliminated.

The tetracycline resistance conferred by the Tn*10* mutations in the three catalase genes simplified the construction of strains carrying specific catalase mutations because it provided a positive selection system for the catalase mutation. The transfer of these mutations into various strains allowed the study of the independent induction of the HPI and HPII catalases (Loewen et al., 1985a) as well as the construction of other catalase mutants for specific purposes, including the cloning of *katG*, *katE*, and *katF*. The construction of mutants lacking catalase, but of known genotype, was necessary to be certain that the catalase gene of interest would complement the mutation in the strain if it was cloned. Besides the certainty of the catalase genotype, the transposon mutants

had other advantages over the traditionally isolated nitrosoguanidine mutants. The low frequency of insertion of transposon Tn*10*, approximately 10^{-5} to 10^{-6} transpositions per input phage genome (Foster et al., 1981), allowed the isolation of single mutations with otherwise unaltered genetic backgrounds. In contrast, NG mutants often have multiple mutations, so the genetic make up of the strain was always uncertain. The insertion of the transposon in the case of *katG* prevented the production of any immunoprecipitable HPI protein (Loewen et al., 1985b), allowing experiments without HPI or the inactive protein to be performed.

The plasmid pLC36-19 was selected from the Clarke and Carbon plasmid bank (1976) because it conferred a higher level of catalase on its host. pLC36-19 complemented the catalase deficiencies of catalase mutants, and it contained the structural gene(s) for the HPI-A and HPI-B isoenzyme pair (Loewen et al., 1983). However, because only 2.5 kb of the 19.2 kb *E. coli* genomic DNA insert in pLC36-19 was needed to code for the 84 000 dalton HPI protein, it was necessary to identify the location of the catalase gene and subclone it for a more detailed characterization.

In order to localize the portion of pLC36-19 required for the synthesis of the HPI protein, fourteen restriction enzymes were mapped in pLC36-19 and transposon Tn*5* insertions were made in the plasmid. Although Tn*5* insertions have successfully been used for the correlation of plasmid function and location in other instances (DeBruijn and Lupski, 1984), only four Tn*5* insertions mapping at different locations in the plasmid were isolated, and none of these disrupted the catalase activity conferred by the plasmid. These insertions did indicate where the HPI gene was not located and making use of this information and the new

restriction sites provided by the transposon, *Hin*dIII deletion mutants of the transposon carrying plasmid, pC1, were isolated. Two plasmids, pBT1 and pBT2 were isolated and *Hin*dIII restriction digests of the plasmids were found to differ by only one restriction fragment. pBT2 which contained a 3.8 kb *Hin*dIII fragment not present in pBT1 was also capable of complementing the catalase deficiency of a mutant, UM53, while pBT1 was not. The orientation of the 3.8 kb *Hin*dIII fragment in the deletion plasmid was confirmed to be the same as in pLC36-19. These results localized the coding region in part or in total to the 3.8 kb *Hin*dIII fragment of pLC36-19.

The next obvious step was to subclone the 3.8 kb *Hin*dIII fragment into the higher copy number vector, pAT153. This gave the catalase encoding plasmid pBT22, and the same insert in the opposite orientation gave the catalase encoding plasmid, pBT54. Both plasmids conferred high catalase levels to the catalase deficient mutant, UM53, which confirmed that the entire gene was contained on the 3.8 kb *Hin*dIII fragment and that catalase expression was not being influenced by the promoter of the plasmid tetracycline resistance gene. Sites for 14 restriction enzymes were mapped in pBT22 to provide markers for further analysis. The same enzyme sites are presumed to be present in pBT54, but in the opposite orientation.

The boundaries of the catalase gene were further defined by deletion analysis. The 0 kb end of the *Hin*dIII insert in pBT22 was defined by a specific deletion of a 340 bp *Eco*RI fragment. The deletion of this fragment disrupted catalase activity, a result which places one of the ends of the HPI gene within 300 bp of the 0 kb end of the insert. The initial attempt to define the opposite end of the gene was made using

pBT5, a clone of the 7550 bp *Sph*I fragment from pLC36-19 which contained most of the *Hin*dIII 3.8 kb fragment and extended beyond the 0 kb end into pLC36-19. Because pBT5 did not have catalase activity, the opposite end of the gene must lie beyond the *Sph*I site. BAL31 deletions were used to further define the 3.8 kb end of the gene. A BAL31 mutant which retained catalase activity, pBT28, had a deletion which extended through the *Hin*cII site at the 3.8 kb end of the insert but stopped before the *Nru*I site at 2.8 kb. The deletions in two catalase deficient BAL31 mutants, pBT29 and pBT30, extended slightly beyond the *Nru*I site at the 3.8 kb end of the insert, but ended before the *Sph*I site. This result confined the HPI gene to a region of approximately 2.5 kb extending from within 300 bp of the 0 kb end to a point between the *Nru*I and *Sph*I sites at 2.0 and 2.8 kb respectively. The small deletions of pBT29 and pBT30 which destroyed the catalase activity suggested that the end of the gene was closer to the *Nru*I than the *Sph*I site. This was later confirmed by the DNA sequence.

The isolated deletion mutants were used to define the HPI promoter region. Crude protein was first isolated from catalase deficient strains bearing the mutant plasmids pBT24 and pBT5 with the hope that a deletion in the promoter end would prevent the production of immunoprecipitable protein, while a deletion at the C-terminal end would permit the production of immunoprecipitable protein. However, neither mutant plasmid produced any immunoprecipitable protein, suggesting that the absence of the carboxyl terminus of the protein either prevented it from being recognized by the antisera or made it so unstable that it was degraded before it could be isolated. The next approach was to use maxicell analysis to examine the plasmid encoded polypeptides of the isolated deletion plasmids. The identification of HPI proteins synthesized from pBT29 and

pBT30 which were less stable, but nearly as large as the HPI proteins produced from pBT22, pBT54, and pBT28, suggested that they had small deletions at the 3.8 kb end and that this was in the C-terminal end of the gene. A deletion at the opposite end, as in pBT24, produced a 45 000 dalton protein, much smaller than would be expected if the deletion were in the C-terminus of the protein. Therefore the promoter appeared to be situated at the 0 kb end of the 3.8 kb *Hin*dIII fragment. This conclusion was confirmed by cloning a 320 bp *Bgl*II fragment from the 0 kb end into the promoter cloning vector pKK232-8. This fragment directed the transcription of a promoterless chloramphenicol acetyltransferase gene, showing that the promoter was at the 0 kb end. A potential promoter region has since been identified in this same fragment by sequencing (see below), and the positive regulator of the gene, *oxyR*, has also been confirmed to interact with this *Bgl*II fragment (Gisela Storz, personal communication).

Although the gene on pBT22 was known to encode the structural gene for HPI catalase and the *katG* gene that had been previously mapped at 89.2 min on the *E. coli* chromosome was suspected to be the structural gene, there was no proof that *katG* was the structural gene. By hybridizing ^{32}P-labelled pBT22 to genomic digests of wild type *E. coli* DNA and *E. coli* DNA containing a Tn*10* insertion in *katG*, the structural gene on pBT22 was confirmed to be the same gene in which a Tn*10* had been inserted and mapped as *katG*. This was evident from the change in hybridization pattern observed between *Hin*cII digests of *E. coli* genomic DNA with and without a Tn*10* insertion in *katG*.

After characterizing the HPI gene and identifying it as *katG*, fragments were subcloned into M13mp18 and M13mp19, and the entire gene

was sequenced using dideoxy chain termination sequencing. An open reading frame of 2178 bp coding for a 726 amino acid polypeptide was identified in the sequence. This sequence also had appropriately placed Shine-Dalgarno, -10 and -35 sequences upstream to it which showed relatively good homology to the *E. coli* consensus sequence. The -35 sequence exhibited the poorest consensus to the *E. coli* sequence, possibly reflecting the fact that *katG* is a positively regulated gene, as positively regulated genes usually have poor homology to the *E. coli* -35 consensus sequence (Harley and Reynolds, 1987).

It is known that OxyR protein binds in the promoter region of *katG* (Gisela Storz, personal communication), so there may be a consensus sequence like that found in the heat shock genes and alkylation response genes which allows the recognition by the positive regulators HtpR and Ada (Gottesman, 1984; Teo et al., 1986). No homology to the heat shock promoter consensus sequence (Watson et al., 1987) or the *ada* promoter consensus sequence (Teo et al., 1986) was identified in *katG*, supporting the findings of Christman et al. (1985) that the HPI-A and HPI-B proteins are not part of the oxidative regulon that was also induced by other stresses such as heat shock. However, a region covering the -10 sequence of the promoter was identified that showed marked homology to the *lexA* binding region consensus sequence (Watson et al., 1987).

```
taCTGTatata-a-aCAGta    lexA binding consensus sequence
  **** ***    ** ***
CGCTGTGTATCGTAACGGTA    katG sequence starting at base 102
```

The capital letters represent bases in the *lexA* consensus sequence that are highly conserved while the bases represented by small letters are weakly conserved. Although there is no evidence that *katG* is repressed by LexA protein (a negative regulator of the SOS response), this homology

may indicate that a relationship with the oxidative response should be investigated further. The promoter region of *katG* and its interaction with the positive regulator OxyR is presently being investigated by Gisela Storz.

A potential termination sequence was also located in an appropriate location downstream from the termination codon. The most commonly used termination codon, UAA (Watson et al., 1987), was found 8 bp upstream from a 14 bp hyphenated dyad. The hyphenated dyad is capable of base pairing to form a long stem and loop structure with a single base mismatch. This loop is extremely stable with a $\Delta G=-21.4$ Kcal. This stem and loop structure in combination with the following 5 uracil residues should allow rho independent termination to occur, although the polyU region is shorter than that normally present in rho independent terminators. This may be compensated for by the high stability of the stem and loop structure.

As a single base inclusion or exclusion in a DNA sequence shifts the inferred protein sequence out of frame, it is necessary to confirm that the predicted sequence is correct by sequencing a portion of the protein. Unfortunately, the N-terminus of the HPI protein was blocked and could not be sequenced, like that of the bovine liver catalase. Schroeder et al. (1964) originally suggested that the N-terminus of bovine liver catalase may be blocked by an acetyl group, but this was not confirmed by x-ray crystallography or positive and negative ion mass spectroscopy (Schroeder et al., 1982a). In bacteria, N-α-acetylated proteins are rare, and at present the only known acetlyated proteins are the *E. coli* ribosomal proteins, L12, S5, S18, tu, and L7 (Tsuansawa and Sakiyama, 1984). Other groups commonly modifying the N-terminus of

proteins include formyl and pyroglutamyl groups (Price and Stevens, 1982).

The HPI protein must be modified in some manner because two "isoenzyme" forms are produced by the *katG* gene. This may be the result of some type of N-terminal modification that results in a charge variant. Moss (1982) said that secondary isoenzymes (isoenzymes that are not encoded by separate genes), may result from the covalent modification of the protein or by the covalent attachment of small molecules or radicals that could alter their properties. It is possible that the covalent attachment of cellular H_2O_2 to the HPI active site results in one of the secondary isoenzyme forms. There is no evidence that the two forms of HPI, HPI-A and HPI-B, result from the cleavage of a signal peptide from a preHPI protein. There appeared to be no signal peptide in the HPI predicted amino acid sequence, and the same sample that produced two bands on a native polyacrylamide gel produced only one protein band on an SDS-PAGE gel. Furthermore, studies in our lab (unpublished results), showed that HPI was neither a secreted nor a membrane-contained protein, although Kranz et al. (1984) have suggested that HPI may be a minor membrane protein. Kranz suggests that his findings maybe the result of a small amount of cytoplasmic protein contamination in the membrane preparation.

Since the N-terminus was blocked, cyanogen bromide was used to produce peptides for sequencing. Because cyanogen bromide cleaves at methionine residues, sequences succeeding methionine residues were searched for in the sequenced cyanogen bromide peptides. Claiborne and Fridovich (1979) examined the peptides produced by cyanogen bromide cleavage of HPI and found the protein contained 13 to 14 methionine

residues. The HPI predicted sequence contained 13 methionine residues, and seven of the sequenced cyanogen bromide cleaved peptides were identified in the HPI sequence. Although there were some discrepencies between the HPI predicted sequence and the determined amino acid sequence, the presence of all the peptides in the predicted sequence and the continued matching of amino acid residues even after a contradictory residue showed that the protein was in the correct reading frame.

The HPI predicted amino acid sequence gave a molecular weight of 80 049 daltons, similar to the apparent SDS/PAGE molecular weights for HPI of 78 000 (Claiborne and Fridovich, 1979) and 84 000 and the gel filtration molecular weight of 81 000 (Loewen and Switala, 1986). This reconfirmed that the protein was HPI and that the protein reading frame was correct. The amino acid composition compared favourably with that determined by Claiborne and Fridovich (1979). There was disparity in the amino acid contents of the predicted sequence and the analyzed amino acid contents in the serine, threonine, and tryptophan residues. This was not unexpected since serine, threonine, and tryptophan contents were analyzed independently from the other amino acids. The most unusual feature about the amino acid composition of HPI was the presence of only one cysteine residue in such a large protein. The lack of cysteine was substantiated by the chemically determined amino acid compositions. These results show that disulfide linkages are not important in the stabilization of the HPI protein. The similarity between the amino acid compositions of BLC and HPI was surprising considering that there is no homology between these two proteins. The amino acid content could however reflect the fact that both are cytoplasmic proteins with a similar function.

At this point the HPI amino acid sequence was compared to a number of other catalases whose protein sequences have been determined or predicted. These included BLC (Schroeder et al., 1982), human catalase (Quan et al., 1986; Bell et al., 1986), yeast T-catalase (Hartig and Ruis, 1986), and PVC (*Penicillium vitale* catalase; Vainshtein et al., 1986). Since the PVC amino acid sequence was predicted only on the basis of the x-ray structure, most of the attention was directed to the alignment of HPI with bovine liver and yeast-T catalase. All of the previously characterized catalases show considerable homology; human and bovine liver catalases are very similar and yeast-T catalase and PVC showed good homology in the β-barrel or heme binding domain, although not in the NAD(P)H binding domain. Both yeast T-catalase and PVC are larger than bovine liver or human catalases by 36 aa and 144 aa respectively. The extension of the PVC has the structure of a flavodoxin binding domain. HPI is also considerably bigger (200 aa) than BLC but no function has been assigned to the extra amino acids. The dual function of HPI may require more amino acids to form a second functional domain.

The HPI protein showed no homology to the previously sequenced catalases. The only potential homology was identified around Tyr-336 which had homology with four of the amino acids surrounding the Tyr-357 of BLC which acts as the fifth ligand of the Fe of the heme group. However, the homology did not extend in either direction around this region, suggesting that if Tyr-336 is a ligand of Fe, none of the other features of the heme binding domain have been conserved.

HPI's lack of homology with other catalases identifies it as a unique catalase with many properties that differ from previously characterized catalases. As described in the historical, the HPI enzyme

is not only a catalase but a broad spectrum peroxidase, capable of using much larger substrates than typical catalases in the peroxidatic mode. This would demand a much different active site from that which has been determined to be present in BLC and PVC by x-ray crystallography. Chance (1949) found that the rate of oxidation of compounds by BLC was reduced by an order of magnitude for each additional carbon atom in the R group of straight chain alcohols, probably because of the limited diameter of the channel accessing the heme. The deep channel accessing the active site is so narrow that entrance is limited to molecules with a van der Waal's diameter of less than 3.5 A (Fita and Rossmann, 1985b), preventing large substrates from reaching the active site. HPI would need a wider channel or a separate active site to allow access to the large substrate molecules used in the peroxidatic mode. The HPI active centre may be more like that of cytochrome-C-peroxidase where the heme environment is polar and partially accessible to solvent. Cytochrome-C-peroxidase also has different essential active site residues and instead of the heme itself forming a radical during the reaction, this function is performed by an amino acid side chain (Fita and Rossmann, 1985b).

Unlike typical catalases, HPI was resistant to inhibition by AT, a compound that interacts with His-74 of the active site of BLC (Margoliash et al., 1960) and inhibits it. This also supports the concept that HPI is a unique catalase.

Another characteristic exemplifying the difference between the active sites of HPI and other characterized catalases was the broad pH maxima exhibited by typical catalases while HPI catalase has a narrow pH optimum at pH 6.8. Murthy et al. (1981) have suggested that the broad pH optimum of BLC may be a result of the unoccupied sixth ligand of the

Fe. Other heme proteins which do not exhibit pH independence, such as metmyoglobin and cytochrome-C-peroxidase, have water occupying the sixth ligand of Fe. The narrow pH optimum of HPI catalase may indicate that the sixth ligand of Fe in this protein is occupied by water or another amino acid side chain.

The codon usage of *katG* was generally like that of a highly expressed gene. Exceptions to this included the alanine and aspartate codons, where the codon bias was not typical of highly expressed genes. As previously suggested, the choice of the less optimal codons in *katG* for alanine may be a method of translational control for the fine tuning of *katG* expression as *katG* has a high alanine content. Codons with low abundance isoaccepting tRNA's were clearly avoided. Alternatively, *katG* could be considered just a moderately expressed gene; the codon usage in the yeast T-catalase was typical of a moderately expressed gene (Hartig and Ruis, 1986). It was expected that the *katG* gene would exhibit some codon bias to allow the efficient translation observed when *katG* is induced to higher levels.

Three promoters of mutant catalase genes, *katG14* (UM1), *katG15* (UM2), and *katG16* (UM56-64) were cloned and sequenced. Two of these mutants, *katG15* and *katG16* had been partially characterized (Loewen et al., 1985b; Mark Smolenski, 1986) and were suspected promoter mutants. However, the sequences of all the promoters were identical to the wild type sequence, showing that these mutants were not *katG* regulatory mutants. The altered regulatory characteristics may have resulted from mutations in the positive regulator of *katG*, *oxyR*. The multiple mutations typical of nitrosoguanidine mutants may explain the various phenotypes exhibited by the *katG15* and *katG16* mutants, all having lower

catalase activity and abnormal regulation of this activity.

The relevance of the 884 bp sequence beyond the termination sequence of *katG* is at present unknown. It does not contain the *oxyR* sequence or any other sequences presently recorded in Genbank (1986), although there is a potential open reading frame. This region could contain some other oxygen regulated loci, as studies by Alibadi et al., (1986), have revealed six oxygen inducible operon fusions, two of these, *oxiC* and *oxiE*, which map close to *oxyR* in *S. typhimurium*.

The *katG* gene product HPI was physically characterized and revealed as a unique catalase, unlike any other previously sequenced catalases. This conclusion was supported by the chemical properties differentiating HPI from typical catalases. In contrast to eucaryotic catalases, bacterial catalases appear to be a very diverse family of enzymes which have evolved to deal with their particular niches. The detailed characterization of HPI supported the concept that bacterial enzymes are very diverse. This has been further corroborated by the inability of *katG* to hybridize to genomic DNA from either *Acinetobacter anitratum* or *Pseudomonas aeruginosa* under non-stringent hybridization conditions (Pam Sorbey, unpublished results). Another example of an *E. coli* gene which is not highly conserved is the *ppc* gene which did not hybridize with DNA from certain other bacteria (Sabe, 1984), further demonstrating the diversity of bacterial evolution. It may be that similar catalases will only be found among the very closely related Enterobacteriaceae, as even *Proteus mirabilis*, a member of the Enterobacteriaceae family does not have a similar catalase with a broad spectrum peroxidase activity (Jouve et al., 1984).

Once the sequence of the HPII catalase has been predicted from the

DNA sequence, the conserved regions of the HPI protein may be identified by alignment. In yeast, although the catalases A and T do not hybridize to each other (Cohen et al., 1985), they both show homology to BLC. If this is the case in *E. coli*, the sequence of HPII may lead to the identification of the conserved regions of HPI which are important for heme binding and catalysis.

Now that the HPI protein has been purified and sequenced, the next step in understanding the significance of the protein to *E. coli* will be to identify the residues playing important roles in the structure and function of the enzyme. This may be done by determining residues which appear to be important to the active site by x-ray crystallography or by comparison to other similar catalase proteins. Attempts are now being made to crystallize the HPI catalase for x-ray crystallography. The regulation of HPI is also an interesting field of investigation and an understanding of the regulation of *katG* could lead to a better understanding of its function. The fact that high levels of HPI can efficiently lower the rates of spontaneous mutation in *Salmonella typhimurium* (Sies et al., in press) suggests that catalase has a much more important role in the cell than had previously been presumed. Further investigation of this unique *E. coli* catalase, hydroperoxidase I, is necessary to provide a better understanding of how organisms cope with toxic oxygen species.

REFERENCES

6.0 REFERENCES

Aliabadi, Z., F. Warren, S. Mya, and J. W. Foster. 1986. Oxygen-regulated stimulons of *Salmonella typhimurium* identified by Mud (Ap lac) operon fusions. J. Bacteriol. 165:780-786.

Allgood, G. S. and J. J. Perry. 1986. Characterization of a manganese-containing catalase from the obligate thermophile *Thermoleophilum album*. J. Bacteriol. 168:563-567.

Ananthaswamy, H. N. and A. Eisenstark. 1977. Repair of hydrogen peroxide-induced single-strand breaks in *Escherichia coli* deoxyribonucleic acid. J. Bacteriol. 130:187-191.

Armel, P. R., G. F. Strniste, and S. S. Wallace. 1977. Studies on *Escherichia coli* x-ray endonuclease specificity. Roles of hydroxyl and reducing radicals in the production of DNA lesions. Radiat. Res. 69:328-338.

Atassi, M. Z., and A. F. S. A. Habeeb. 1972. Reactions of proteins with citraconic anhydride. Methods Enzymol. 25:546-553.

Bachmann, B. J. 1983. Linkage Map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. 47:180-230.

Barbado, C., M. Ramirez, M. Angel-Blanco, J. Lopez-Barea, and C. Pueyo. 1983. Mutants of *Escherichia coli* sensitive to hydrogen peroxide. Curr. Microbiol. 8:251-253.

Barnes, W. M. 1977. Plasmid detection and sizing in single colony lysates. Science 195:393-394.

Beauchamp, C., and I. Fridovich. 1970. A mechanism for the production of ethylene from methional. J. Biol. Chem. 245:4641-4646.

Bell, G. I., R. C. Najarian, G. T. Mullenbach, and R. A. Hallewell. 1986. cDNA sequence coding for human kidney catalase. Nucl. Acids Res. 14:5561-5562.

Bennetzen, J. L., and B. D. Hall. 1982. Codon selection in yeast. J. Biol. Chem. 257:3026-3031.

Bilinski, T., Z. Krawiec, A. Liczmanski, and J. Litwinska. 1985. Is hydroxyl radical generated by the Fenton reaction in vivo? Biochem. Biophys. Res. Commum. 130:533-539.

Birnboim, H. C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. 100:243-255.

Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7:1513-1523.

Bochner, B. R., P. C. Lee, S. W. Wilson, C. W. Cutler, and B. N. Ames. 1984. AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. Cell 37: 225-232.

Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer. 1977. Construction and characterization of new cloning vehicles II. A multipurpose cloning system. Gene 2:95-113.

Borer, P. N., B. Dengler, I. Tinoco Jr., and O. C. Uhlenbeck. 1974. Stability of ribonucleic acid double-stranded helices. J. Mol. Biol. 86:843-853.

Boulnois, G. J. and K. N. Timmis. 1984. Synthesis of plasmid encoded polypeptides in maxicells. In Advanced Molecular Genetics. (A. Puhler and K. N. Timmis, eds.) Springer-Verlag, Berlin, Germany. pp. 204-211.

Boyer, H. W. and D. Roulland-Dussoix. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. 41:459-472.

Breimer, L. H. and T. Lindahl. 1985. Enzymatic excision of DNA bases damaged by exposure to ionizing radiation or oxidizing agents. Mutat. Res. 150:85-89.

Brosius, J. 1984. Plasmid vectors for the selection of promoters. Gene 27:151-160.

Carlsson, J. and W. S. Carpenter. 1980. The *recA*$^+$ gene product is more important than catalase and superoxide dismutase in protecting *Escherichia coli* against hydrogen peroxide toxicity. J. Bacteriol. 142:319-321.

Chance, B. 1949. The primary and secondary compounds of catalase and methyl or ethyl hydrogen peroxide. J. Biol. Chem. 179:1341-1369.

Christman, M. F., R. W. Morgan, F. S. Jacobson, and B. N. Ames. 1985. Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in *Salmonella typhimurium*. Cell 41:753-762.

Claiborne, A. and I. Fridovich. 1979. Purification of the o-dianisidine peroxidase from *Escherichia coli* B. J. Biol. Chem. 254:4245-4252.

Claiborne, A., D. P. Malinowski, and I. Fridovich. 1979. Purification and characterization of hydroperoxidase II of *Escherichia coli* B. J. Biol. Chem. 254:11664-11668.

Claiborne, A. 1978. Relationship between catalase and dianisidine peroxidase in *Escherichia coli* B. Fed. Proc. 37:1513.

Clarke, L. and J. Carbon. 1976. A colony bank containing synthetic colE1 hybrid plasmids representative of the entire *E. coli* genome. Cell 9:91-99.

Clayton, R. K. 1959. Purified catalase from *Rhodopseudomonas spheroides*. Biochim. Biophys. Acta 36:40-47.

Clewell, D. B. 1972. Nature of ColE1 plasmid replication in *Escherichia coli* in the presence of chloramphenicol. J. Bacteriol. 110:667-676.

Cohen, G., F. Fessl, A. Traczyk, J. Rytka, and H. Ruis. 1985. Isolation of the catalase A gene of *Saccharomyces cerevisiae* by complementation of the *cta1* mutation. Mol. Gen. Genet. 200:74-79.

Dabbs, E. R. 1981. The gene for ribosomal protein L31, *rpmE* is located at 88.5 minutes on the *Escherichia coli* chromosomal linkage map. J. Bacteriol. 148:379-382.

Davis, B. J. 1964. Disc electrophoresis. II. Method and application to human serum proteins. Ann. N. Y. Acad. Sci. 121:404-427.

De Bruijn, F. J., and J. R. Lupski. 1984. The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids -a review. Gene 27:131-149.

Demple, B. and J. Halbrook. 1983. Inducible repair of oxidative damage in *Escherichia coli*. Nature 304:466-468.

Demple, B., J. Halbrook, and S. Linn. 1983. *Esherichia coli xth* mutants are hypersensitive to hydrogen peroxide. J. Bacteriol. 153:1079-1082.

Dickerson, R. E., H. B. Gray, and G. B. Haight. 1979. Chemical Principles. The Benjamin/Cummings Publishing Company Inc., Don Mills, Ontario.

Dingman, D. W. and D. P. Stahly. 1984. Protection of *Bacillus larvae* from oxygen toxicity with emphasis on the role of catalase. Appl. Environ. Micro. 47:1228-1237.

Eisenstark, A. and G. Perrot. 1987. Catalase has only a minor role in protection against near-ultraviolet radiation damage in bacteria. Mol. Gen. Genet. 207:68-72.

Farr, S., R. D'Ari, and D. Touati. 1986. Proc. Natl. Acad. Sci. USA, in press.

Farr, S. B., D. O. Natvig, and T. Kogoma. 1985. Toxicity and mutagenicity of plumbagin and the induction of a possible new DNA repair pathway in *Escherichia coli*. J. Bacteriol. 164: 1309-1316.

Fee, J. A. 1982. Is superoxide important in oxygen poisoning? Trends Biochem. Sci. 7:84-86.

Fields, P. I., R. V. Swanson, C. G. Haidaris, and F. Heffron. 1986. Mutants of *Salmonella typhimurium* that cannot survive within the macrophage are avirulent. Proc. Natl. Acad. Sci. USA 83:5189-5193.

Finn, G. J. and S. Condon. 1975. Regulation of catalase synthesis in *Salmonella typhimurium*. J. Bacteriol. 123:570-579.

Fita, I. and M. G. Rossmann. 1985a. The NADPH binding site on beef liver catalase. Proc. Natl. Acad. Sci. USA 82:1604-1608.

Fita, I. and M. G. Rossmann. 1985b. The active center of catalase. J. Mol. Biol. 185:21-37.

Foster, T. J., V. Lundblad, S. Hanley-Way, S. M. Halling, and N. Kleckner. 1981. Three Tn*10*-associated excision events: relationship to transposition and role of direct and inverted repeats. Cell 23:215-227.

Frey, J., M. Bagdasarian, and K. N. Timmis. 1984. Generation of deletion mutations in vitro with the BAL31 exonuclease. In Advanced Molecular Genetics. (A.Puhler, and K. N. Timmis eds.) Springer-Verlag, Berlin, Germany. pp. 141-150.

Fridovich, I. 1986. Biological effects of the superoxide radical. Arch. Biochem. Biophys. 247:1-11.

Fridovich, I. 1978. The biology of oxygen radicals. Science 201:875-880.

Fridovich, I. 1977. Oxygen is toxic! Bioscience 27:462-466.

Friesen, J. D., J. Parker, R. J. Watson, N. P. Fiil, S. Pedersen, and F.S. Pedersen. 1976. Isolation of lambda transducing bacteriophage carrying the *relA* gene of *Escherichia coli*. J. Bacteriol. 127:917-922.

Furuta, S., H. Hayashi, M. Hijikata, S. Miyazawa, T. Osumi, and T. Hashimoto. 1986. Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liver catalase. Proc. Natl. Acad Sci. 83:313-317.

Ganoza, M. C., E. C. Kofoid, P. Marliere, and B. G. Louis. 1987. Potential secondary structure at translation-initiation sites. Nucl. Acids Res. 15:345-359.

Glover, D. M. 1984. Gene Cloning. Chapman and Hall, N.Y., N.Y.

Goff, S. A., and A. L. Goldberg. 1985. Production of abnormal proteins in *E. coli* stimulates transcription of *lon* and other heat

shock genes. Cell 41:587-595.

Gregory, E. M., and I. Fridovich. 1974. Visualization of catalase on acrylamide gels. Anal. Biochem. 58:57-62.

Gregory, E. M., B. J. Vetri, D. L. Wagner, and T. D. Wilkins. 1977. Carbohydrate repression of catalase synthesis in *Bacteriodes fragilis*. J. Bacteriol. 129:534-535.

Grosjean, H., and W. Fiers. 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene 18:199-209.

Gottesman, S. 1984. Bacterial regulation: global regulatory networks. Ann. Rev. Genet. 18:415-441.

Haddock, B. A., and C. W. Jones. 1977. Bacterial Respiration. Bacteriol. Rev. 41:47-99.

Hagensee, M. E. and R. E. Moses. 1986. Repair response of *Escherichia coli* to hydrogen peroxide DNA damage. J. Bacteriol. 168: 1059-1065.

Halliwell, B. and J. M. C. Gutteridge. 1985. Free Radicals in Biology and Medicine. Oxford University Press, New York.

Halliwell, B. D. A. Rowley, and J. M. C. Gutteridge. 1983. Transition metal catalysis and oxygen radical reactions. In Life Chemistry Reports Supplement 2 (Oxidative damage and Related Enzymes-EMBO Workshop 1983) pp 8-14.

Halliwell, B. 1982. Superoxide and superoxide-dependent formation of hydroxyl radicals are important in oxygen toxicity. Trends Biochem. Sci. 7:270-272.

Harley, C. B. and R. P. Reynolds. 1987. Analysis of *E. coli* promoter sequences. Nucl. Acids Res. 15:2343-2361.

Hartig, A. and H. Ruis. 1986. Nucleotide sequence of *Saccharomyces cerevisiae* CTT1 gene and deduced amino-acid sequence of yeast catalase T. Eur. J. Biochem. 160:487-490.

Hassan, H. M. and I. Fridovich. 1978. Regulation of the synthesis of catalase and peroxidase in *Escherichia coli*. J Biol. Chem. 253:6445-6450.

Hassan, H. S. and I. Fridovich. 1980. Mechanism of the antibiotic action of pyocyanine. J. Bacteriol. 141:156-163.

Herbert, D. and J. Pinsent. 1948. Crystalline bacterial catalase. Biochemical J. 43:193-202.

Hershfield, V., H. W. Boyer, C. Yanofsky, M. A. Lovett, and D. R.

Helinski. 1974. Plasmid colE1 as a molecular vehicle for cloning and amplification of DNA. Proc. Natl. Acad. Sci. USA 71:3455-3459.

Himelbloom, B. H. and H. M. Hassan. 1986. Effects of cysteine on growth, protease production, and catalase activity of *Pseudomonas fluorescens*. Appl. Environ. Microbiol. 51:418-421.

Hindley, J. and R. Staden. 1983. DNA Sequencing. (T. S. Work, and R. H. Burdon, eds.) Elsevier Biomedical Press, N.Y., N.Y.

Hopp, T. P. and K. R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78:3824-3828.

Ikemura, T. and H. Ozeki. 1982. Codon usage and transfer RNA contents: organism-specific condon-choice patterns in reference to the isoacceptor contents. Cold Spring Harbor Symp. Quant. Biol. 47:1087-1097.

Imlay, J. A., and S. Linn. 1986. Bimodal pattern of killing of DNA-repair-defective or anoxically grown *Escherichia coli* by hydrogen peroxide. J. Bacteriol. 166:519-527.

Jacob, G. S., and W. H. Orme-Johnson. 1979. Catalase of *Neurospora crassa*. 1. Induction, purification, and physical properties. Biochemistry 18:2967-2975.

Joenje, H., P. Van Der Valk, J. J. P. Gille, F. Arwert, E. H. A. Poll, M. L. Kwee, D. Lindhout, A. S. Eriksson, and J. Van Rijn. 1983. Oxygen-dependent genetic damage. In Life Chemistry Reports Supplement 2 (Oxidative Damage and Related Enzymes-EMBO Workshop 1983) p.309-317.

Johnson, K. R., W. M. Nauseef, A. Care, M. J. Wheelock, S. Shane, S. Hudson, H. P. Koeffler, M. Selsted, C. Miller, and G. Rovera. 1987. Characterization of cDNA clones for human myeloperoxidase: predicted amino acid sequence and evidence for multiple mRNA species. Nucl. Acids Res. 15:2013-2028.

Jones-Mortimer, M. C. 1973. Mapping of structural genes for the enzymes of cysteine biosynthesis in *Escherichia coli K12* and *Salmonella typhimurium* LT2. Heredity 31:213-221.

Jones-Mortimer, M. C. 1968. Positive control of sulphate reduction in *Escherichia coli*. Biochem. J. 110:589-595.

Jouve, H. M., J. Gaillard, and J. Pelmont. 1984. Characterization and spectral properties of *Proteus mirabilis* PR catalase. Can. J. Biochem. Cell Biol. 62:935-944.

Kirkman, H. N., and G. F. Gaetani. 1984. Catalase: a tetrameric enzyme with four tightly bound molecules of NADPH. Proc. Natl. Acad. Sci. USA 81:4343-4347.

Kono, Y., and I. Fridovich. 1983a. Functional significance of Mn-catalase in *Lactobacillus plantarum*. J. Bacteriol. 155:742-746.

Kono, Y., and I. Fridovich. 1983b. Isolation and characterization of the pseudocatalase of *Lactobacillus plantarum*: a new manganese-containing enzyme. J. Biol. Chem. 258:6015-6019.

Kramer G. F., and B. N. Ames. 1987. Oxidative mechanisms of toxicity of low-intensity near UV-light in *Salmonella typhimurium*. J. Bacteriol. 169:2259-2266.

Kranz, R. G., C. A. Barassi, and R. B. Gennis. 1984. Immunological analysis of the heme proteins present in aerobically grown *Escherichia coli*. J. Bacteriol. 158:1191-1194.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.

Layne, E. 1957. Spectrophotometric and turbidimetric methods for measuring proteins. Methods Enzymol. 3:447-454.

Lee, P. C., B. R. Bochner, and B. N. Ames. 1983. AppppA, heat-shock, stress, and cell oxidation. Proc. Natl. Acad. Sci. USA 80: 7496-7500.

Levin, D. E., M. Hollstein, M. F. Christman, E. A. Schwiers, and B. N. Ames. 1982. A new *Salmonella* tester strain (TA102) with A·T base pairs at the site of mutation detects oxidative mutagens. Proc. Natl. Acad. Sci. USA 79:7445-7449.

Levine S. A. 1977. Isolation and characterization of catalase deficient mutants of *Salmonella typhimurium*. Molec. Gen. Genet. 150: 205-209.

Loewen, P. C., and J. Switala. 1986. Purification and characterization of catalase HPII from *Escherichia coli* K12. Biochem. Cell Biol. 64:638-646.

Loewen, P. C., J. Switala, and B. L. Triggs-Raine. 1985a. Catalases HPI and HPII in *Escherichia coli* are induced independently. Arch. Biochem. Biophys. 243:144-149.

Loewen, P. C., B. L. Triggs, C. S. George, and B. E. Hrabarchuk. 1985b. Genetic mapping of *katG* a locus that affects synthesis of the bifunctional catalase-peroxidase hydroperoxidase I in *Escherichia coli*. J. Bacteriol. 162:661-667.

Loewen, P. C. 1984. Isolation of catalase-deficient *Escherichia coli* mutants and genetic mapping of *katE*, a locus that affects catalase activity. J. Bacteriol. 157:622-626.

Loewen, P. C., and B. L. Triggs. 1984. Genetic mapping of *katF*, a locus

that with *katE* affects the synthesis of a second catalase species in *Escherichia coli*. J. Bacteriol. 160:668-675.

Loewen, P. C., B. L. Triggs, G. R. Klassen, and J. H. Weiner. 1983. Identification and physical characterization of a ColE1 hybrid plasmid containing a catalase gene of *Escherichia coli*. Can. J. Biochem. Cell Biol. 61:1315-1321.

Maloy, S. R., and W. D. Nunn. 1981. Selection for loss of tetracycline resistance by *Escherichia coli*. J. Bacteriol. 145:1110-1112.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Margoliash, E., A. Novogrodsky, and A. Schejter. 1960. Irreversible reaction of 3-amino-1:2:4:-triazole and related inhibitors with the protein of catalase. 74:339-348.

Mayer, B. K., and J. O. Kalkinbam III. 1986. Catalase (EC.1.11.1.6) activity and its heat inactivation for differentiation of *Mycobacterium avium*, *Mycobacterium intracellulare* and *Mycobacterium scrofulaceum*. Int. J. System. Bacteriol. 36: 207-212.

Mayer, H., R. Grosschedl, H. Schütte, and G. Hobom. 1981. *Cla*I, a new restriction endonuclease from *Caryophanon latum* L. Nucl. Acids Res. 9:4833-4845.

McCormick, J. P., J. R. Fischer, J. P. Pachlatko, and A. Eisenstark. 1976. Characterization of a cell-lethal product from the photo-oxidation of tryptophan: hydrogen peroxide. Science 191:468-469.

Meir, E., and E. Yagil. 1985. Further characterization of the two catalases in *Escherichia coli*. Curr. Microbiol. 12:315-320.

Meir, E., and E. Yagil. 1984. Catalase-negative mutants of *Escherichia coli*. Curr. Microbiol. 10:13-18.

Melik-Adamyan, W. R., V. V. Barynin, A. A. Vagin, V. V. Borisov, B. K. Vainshtein, I. Fita, M. R. N. Murthy, and M. G. Rossmann. 1986. Comparison of beef liver and *Penicillium vitale* catalases. J. Mol. Biol. 188:63-72.

Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.

Mitchell, P. 1979. Keilin's respiratory chain concept and its chemiosmotic consequences. Science 206:1148-1159.

Moody, C. S., amd H. M. Hassan. 1982. Mutagenicity of oxygen free radicals. Proc. Natl. Acad. Sci. USA 79:2855-2859.

Morgan, A. R., R. L. Cone and T. M. Elgert. 1976. The mechanism of DNA

strand breakage by vitamin C and superoxide and the protective roles of catalase and superoxide dismutase. Nucl. Acids Res. 3:1139-1149.

Morgan, R. W., M. F. Christman, F. S. Jacobson, G. Storz, and B. N. Ames. 1986. Hydrogen peroxide-inducible proteins in *Salmonella typhimurium* overlap with heat shock and other stress proteins. Proc. Natl. Acad. Sci. USA 83:8059-8063.

Murthy, M. R. N., T. J. ReidIII, A. Sicignano, N. Tanaka, and M. G. Rossmann. 1981. Structure of beef liver catalase. J. Mol. Biol. 152:465-499.

Moss, D. W. 1982. Isoenzymes. Chapman and Hall, New York, N. Y.

Nadler, V., I. Goldberg, and A. Hochman. 1986. Comparative study of bacterial catalases. Biochim. Biophys. Acta 882:234-241.

Nahmias, J. A. and G. C. Bewley. 1984. Characterization of catalase purified from *Drosophila melanogaster* by hydrophobic interaction chromatography. Comp. Biochem. Physiol. 77B:355-364.

Nies, D., and H. G. Schlegel. 1982. Catalase from *Comamonas compransoris*. J. Gen. Appl. Microbiol. 28:311-319.

Ono, M., and M. Kuwano. 1978. Mutation affecting the thermolability of the 50s ribosomal subunit in *Escherichia coli*. J. Bacteriol. 134:677-679.

Osumi, T., H. Ozasa, S. Miyazawa, and T. Hashimoto. 1984. Molecular cloning of cDNA for rat liver catalase. Biochem. Biophys. Res. Commun. 122:831-837.

Ota, A., 1986. Properties of catalase activity in vegetative and sporulating cells of yeast *Saccharomyces cerevisiae*. J. Cell. Biochem. 30:331-339.

Padgett, P. J., and N. R. King. 1986. Factors relating to the aerotolerance of *Spirillum volutans*. Can J. Micro. 32:548-552.

Patterson, L. K. 1981. Studies of radiation induced peroxidation in fatty acid micelles. In Oxygen and Oxy-radicals in chemistry and Biology. (M. A. J. Rodgers and E. L. Powers, eds.) Academic Press: New York.

Pearson, M. L. 1972. The role of adenosine 3'5'-cyclic monophosphate in the growth of bacteriophage lambda. Virology 49:605-609.

Platt, T., and D. G. Bear. 1983. Role of RNA polymerase, factor, and ribosomes in transcription termination. In Gene Function in Procaryotes. (J. Beckwith, J. Davies, and J. A Gallant, eds,), Cold Spring Harbor, New York. pp 123-161.

Price, N. C., and L. Stevens. 1982. Fundamentals of Enzymology. Oxford

University Press, New York.

Quan, F., R. G. Korneluk, M. B. Tropak, and R. A. Gravel. 1986. Isolation and characterization of the human catalase gene. Nucl. Acids Res. 14:5321-5333.

Richter, H. E., and P. C. Loewen. 1982. Catalase synthesis in *Escherichia coli* is not controlled by catabolite repression. Arch. Biochem. Biophys. 215:72-77.

Richter, H. E., and P. C. Loewen. 1981. Induction of catalase in *Escherichia coli* by ascorbic acid involves hydrogen peroxide. Biochem. Biophys. Res. Commun. 100:1039-1046.

Riley, M., and A. Anilionis. 1978. Evolution of the bacterial genome. Ann Rev. Microbiol. 32:519-560.

Rørth, M., and P. K. Jensen. 1967. Determination of catalase activity by means of the Clark oxygen electrode. Biochim. Biophys. Acta. 139:171-173.

Rosenberg, M. and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Ann. Rev. Genet. 13:319-353.

Rupp, W. D., C. E. Wilde III, D. L. Reno, and P. Howard-Flanders. 1971. Exchanges between DNA strands in ultraviolet-irradiated *Escherichia coli* J. Mol. Biol. 61:25-44.

Sabe, H., T. Miwa, T. Kodaki, K. Izui, S. Hiraga, and H. Katsuki. 1984. Molecular cloning of the phophoenolpyruvate carboxylase gene, *ppc*, of *Escherichia coli*. Gene 31:279-283.

Sammartano, L. J., R. W. Tuveson, and R. Davenport. 1986. Control of sensitivity to inactivation by H_2O_2 and broad-spectrum near-UV radiation by the *Escherichia coli katF* locus. J. Bacteriol. 168:13-21.

Sammartano, L. J., and R. W. Tuveson. 1984. The effect of exogenous catalase on broad-spectrum near-UV (300-400 nm) treated *Escherichia coli* cells. Photochem. Photobiol. 40:607-612.

Sancar, A., A.M. Hack, and W. D. Rupp. 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. 137:692-693.

Sato. S., Y. Nakada, and K. Nakazawa-Tomizawa. 1987. Amino-acid sequence of a tetrameric, manganese superoxide dismutase from *Thermus thermophilus* HB8. Biochim. Biophys. Acta 912:178-184.

Schonbaum, G. R., and B. Chance. 1976. Catalase. In The Enzymes (P. D. Boyer, ed.) Academic Press, N. Y. 13:363-408.

Schreir, P. H., and R. Cortese. 1979. A fast and simple method for

sequencing DNA cloned in the single-stranded bacteriophage M13. J. Mol. Biol. 129:169-172.

Schroeder, W. A., J. R. Shelton, J. B. Shelton, B. Robberson, G. Apell, R. S. Fang, and J. Bonaventura. 1982a. The complete amino acid sequence of bovine liver catalase and the partial sequence of bovine erythrocyte catalase. Arch. Biochem. Biophys. 214:397-421.

Schroeder, W. A., J. R. Shelton, J. B. Shelton, G. Apell, L. Evans, J. Bonaventura, and R. S. Fang. 1982b. The partial amino acid sequence of human erythrocyte catalase. Arch. Biochem. Biophys. 214:422-424.

Schroeder, W. A., J. R. Shelton, J. B. Shelton, and B. M. Olson. 1964. Some amino acid sequences in bovine liver catalase. Biochem. Biophys. Acta 89:47-65.

Seah, T. C. M., A. R. Bhatti, and J. G. Kaplan. 1973. Novel catalatic proteins of bakers' yeast I. an atypical catalase. Can. J. Biochem. 51:1551-1555.

Seah, T. C. M., and J. G. Kaplan. 1973. Purification and properties of the catalase of baker's yeast. J. Biol. Chem. 248: 2889-2893.

Selker, E., K. Brown, and C. Yanofsky. 1977. Mitomycin C-induced expression of *trpA* of *Salmonella typhimurium* inserted into the plasmid colE1. J. Bacteriol. 129:388-394.

Shpaer, E. G. 1986. Constraints on codon context in Escherichia coli genes. Their possible role in modulating the efficiency of translation. J. Mol. Biol. 188:555-564.

Sies H., M. F. Christman, G. Storz, and B. N. Ames. 1987. Role of the *oxyR* regulon in spontaneous mutagenesis. In press.

Silhavy, T. J., M. L. Berman, and L. W. Enquist. 1984. Experiments with Gene Fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Skadsen, R. W., and J. G. Scandalios. 1986. Evidence for processing of maize catalase 2 and purification of its messenger RNA aided by translation of antibody-bound polysomes. Biochemistry 25:2027-2032.

Smolenski, M. 1986. Cloning and Partial Characterization of Three *katG* Mutants from *Escherichia coli*. MSc Thesis. University of Manitoba.

Steers, E., G. R. Craven, and C. B. Anfinsen. 1965. Evidence for nonidentical chains in the -galactosidase of *Escherichia coli* K12. J. Biol. Chem. 240:2478-2484.

Stephens, J. C., S. W. Artz, and B. N. Ames. 1975. Guanosine 5'-

diphosphate 3'-diphosphate (ppGpp): positive effector for histidine operon transcription nad general signal for amino acid deficiency. Proc. Natl. Acad. Sci. USA 72:4389-4393.

Sulebele, G. A., and D. V. Rege. 1968. The nature of the glucose effect on the induced synthesis of catalase in *Saccharomyces cerevisiae.* Enzymologia 36:321-334.

Sutcliffe, J. G. 1978. pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. 5:2721-2728.

Teo, I., B. Sedgwick, M. W. Kilpatrick, T. V. McCarthy, and T. Lindahl. 1986. The intracellular signal for induction of resistance to alkylating agents in *E. coli.* Cell 45:315-324.

Touati, D. 1987. in press.

Triggs-Raine, B. L., and P. C. Loewen. 1987. Physical characterization of *katG* encoding catalase HPI of *Escherichia coli.* Gene 52:121-128.

Tsunasawa, S., and F. Sakiyama. 1984. Amino-terminal acetylation of proteins: an overview. Methods in Enzymol. 106:165-170.

Tuveson, R. W., and L. J. Sammartano. 1986. Sensitivity of *hemA* mutant *Escherichia coli* cells to inactivation by near-uv light depends on the level of supplementation with δ-aminolevulinic acid. Photochem. Photobiol. 43:621-626.

Tuveson, R. W. 1981. The interaction of a gene (*nur*) controlling near-UV sensitivity and the *polA1* gene in strains of *E. coli* K12. Photochem. Photobiol. 33:919-923.

Twigg, A. J., and D. Sherratt. 1980. Trans-complementable copy-number mutants of plasmid colE1. Nature 283:216-218.

Tyrrell, R. M. 1985. A common pathway for protection of bacteria against damage by solar UVA (334 nm, 365 nm) and an oxidizing agent (H_2O_2). Mutat. Res. 145:129-136.

Vainshtein, B. K., W. R. Melik-Adamyan, V. V. Barynin, A. A. Vagin, A. I. Grebenko, V. V. Borisov, K. S. Bartels, I. Fita, and M. G. Rossmann. 1986. Three-dimensional structure of catalase from *Penicillium vitale* at 2.0 A resolution. J. Mol. Biol. 188:49-61.

Van Hemmen, J. J., and W. J. A. Meuling. 1975. Inactivation of biologically active DNA by γ-ray-induced superoxide radicals and their dismutation products singlet molecular oxygen and hydrogen peroxide. Biochim. Biophys. Acta 402:133-141.

Vassilyadi, M., and Archibald, F. 1985. Catalase, superoxide dismutase, and the production of O_2-sensitive mutants of *Bacillus coagulans.*

Can. J. Microbiol. 31:994-999.

Vogelstein, B., and D. Gillespie. 1979. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad Sci., USA 76:615-619.

Way, J. C., M. A. Davis, D. Morisato, D. E. Roberts, and N. Kleckner. 1984. New Tn*10* derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene 32: 369-379.

Watson, J. D., N. H. Hopkins, J. W. Roberts, J. A. Steitz, and A. M. Weiner. 1987. Molecular Biology of the Gene. The Benjamin/Cummings Publishing Company, Inc. Don Mills, Ontario.

Winquist, L., U. Rannug, A. Rannag, and C. Ramel. 1984. Protection from toxic and mutagenic effects of hydrogen peroxide by catalase induction in *Salmonella typhimurium*. Mutat. Res. 141:145-147.

Winterbourn, C. C. 1983. Biological mechanisms of hydroxyl radical production. In Life Chemistry Reports Supplement 2 (Oxidative Damage and Related Enzymes-EMBO Workshop 1983) pp. 1-7.

Worcel, A., and E. Burgi. 1974. Properties of a membrane-attached form of the folded chromosome of Escherichia coli. J. Mol. Biol. 82:91-105.

Wu, T. T. 1966. A model for three point analysis of random general transduction. Genetics 54:405-410.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.

Yonei, S., R. Yokota, and Y. Sato. 1987. The distinct role of catalase and DNA repair systems in protection against hydrogen peroxide in *Escherichia coli*. Biochem. Biophys. Res. Commun. 143: 638-644.

Yoshpe-Purer, Y., Y. Henis, and J. Yashphe. 1977. Regulation of catalase level in *Escherichia coli* K12. Can J. Microbiol. 23:84-91.

Yoshpe-Purer, Y., and Y. Henis. 1976. Factors affecting catalase level and sensitivity to hydrogen peroxide in *Escherichia coli*. Appl. Environ. Micro. 32:465-469.

Yousten, A. A., J. L. Johnson, and M. Salin. 1975. Oxygen metabolism of catalase-negative and catalase-positive strains of *Lactobacillus plantarum*. J. Bacteriol. 123:242-247.