The Cloning and Analysis of *Rhizobium* Dehalogenase Genes

SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE DEPARTMENT OF BIOCHEMISTRY

UNIVERSITY OF LEICESTER

STEPHEN SCOTT CAIRNS BSc (Hons)

April 1994

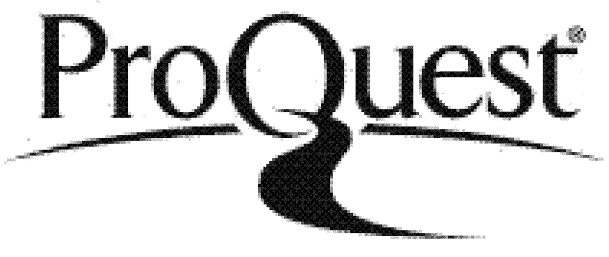
UMI Number: U056891

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U056891
Published by ProQuest LLC 2015. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

7502091808

Abstract

The Cloning and Analysis of *Rhizobium* Dehalogenase Genes

Stephen. S. Cairns

A *Rhizobium* genomic library was constructed in *Escherichia coli* NM522 and the library was screened for the ability to grow on the halogenated compound 2-chloropropionate (2CP). Two positive clones were identified and the plasmids designated pSC1 and pSC530. Both clones allowed *E.coli* to grow on 2CP at growth rates slower than that seen for the *Rhizobium*.. Investigation of the dehalogenase activities of the clones showed that extracts from cells containing pSC1 were active against D-, L- and D/L-2CP and extracts from cells containing pSC530 were also active against dichloropropionate. Restriction enzyme mapping of the clones indicated that they were unique regions of the genome that conferred the new growth ability on *E.coli*. Southern analysis of the *Rhizobium* genomic DNA using insert DNA as probes confirmed that both of the clones were from the *Rhizobium* and that the clones were not contiguous on the genome. Subcloning of pSC1 resulted showed that the insert encoded two stereospecific dehalogenase genes and two plasmids were constructed by subcloning that were each able to express each of the two dehalogenases. The stereospecific dehalogenases (HadD and HadL) were purified and the N-terminal amino acid sequences were determined. HadL was also purified from the *Rhizobium* and the N-terminal sequence was shown to be the same as the cloned HadL. The complete nucleotide sequences of the *hadD* and *hadL* genes were determined and showed little similarity to each other or to other known dehalogenases. HadD and HadL were used, both *in vivo* and *in vitro*, to resolve a racemic solution of 2CP, an analagous situation to possible industrial applications of dehalogenases.

Declaration

This thesis, submitted for the degree of Doctor of Philosophy entitled: The cloning and analysis of *Rhizobium* dehalogenase genes, is based upon work conducted by the author in the Department of Biochemistry between October 1st 1990 and September 30th 1993.

All of the work recorded in this thesis is original unless otherwise acknowledged in the text or by references.

None of this work has been submitted for another degree in this or any other University.

Signed: Date:

Acknowledgements

I would firstly like to thank Dr. R. A. Cooper for his helpful advice and encouragement throughout my time in his lab and during the course of my work. My thanks also go to Dr. A. Cornish and Dr. J. Sampson or their useful advice during my research. I would also like to thank the many people I have worked with in Lab 108, especially Neil Robson and Abby Hasan, without whom I am sure things wouldn't have gone so well. Next I would like to thank all those numerous people in the Biochemistry department that provided the answers to most of my questions.

I feel I have to say thankyou to Steve Heald, Neil Wyborn and Tim Seddon who have made living in Leicester an experience that I won't forget in a hurry, as well as to Jim for being pompous and to Steve for hauling on the rope when necessary. Finally a special thank you to my Mum, who has supported me throughout and to Jackie, who has stuck by me and encouraged me for all this time.

Abbreviations

Amp	Ampicillin
bp	Base pairs
CTAB	Hexadecyltrimethyl ammonium bromide
DCA	Dichloroacetic acid
DCP	Dichloropropionic acid
2CP	D/L-2-chloropropionic acid
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast Protein Liquid Chromatography
IPTG	Isopropyl-ß-D-thiogalactoside
kbp	Kilo base pairs
kDa	Kilo Daltons
λ	Wavelength
MCA	Monochloroacetic acid
MOPS	3-[N-morpolino] propanesulphonic acid
NAD(H)	Nicotine adenine dinucleotide (reduced)
NTA	Nitrilotriacetic acid?
OD_{680}	Optical density at 680 nm
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
SDS	Sodium dodecylsulphate
TEMED	N,N,N',N'-tetramethylethyldiamine
Tris	Trishydroxymethylaminomethane

Table of Contents

Chapter 1

Introduction

1. Introduction

1.1 Occurrence of naturally halogenated organic compounds

Naturally occurring organic compounds containing covalently bound halogens are found widely throughout the environment, having been isolated from both prokaryotes and eukaryotes, over 700 compounds being identified (Wackett 1991). Several plant species have been discovered that synthesise organofluorine compounds, with a well documented case being the synthesis of fluoroacetic acid by *Dichapetalum cymosum* (Marais 1944). This compound is converted *in vivo* to fluoroacetyl CoA and this gives rise to the subsequent 'lethal synthesis' of fluorocitrate which is a potent inhibitor of aconitase, an enzyme found in the tricarboxylic acid cycle. Since this discovery, fluoroacetate has been found in a number of other higher plants including tea plants, although at sublethal concentrations (Suida and DeBernardis 1973).

By far the most common halogen substituent found naturally is chlorine, with most of the chlorinated compounds isolated having fungal or lichen origins. In the latter case the metabolite is often synthesised by the fungal partner. Many of these compounds have an aromatic nature and show biological activity, especially as antimicrobial or antitumor agents. Well known examples of antibiotics are compounds such as chloramphenicol from *Streptomyces venezuelae* , chlortetracycline from *St. aureofaciens* and griseofulvin from *Penicillium grisofulvin*. Many other chlorinated compounds have been isolated from fungi with functions ranging from pigments to plant auxin-like properties.

Brominated compounds are found in many species of marine organisms, and again many show antimicrobial activities or have a pigment role in the organism. Indeed, the use of a brominated pigment from sea

molluscs stretches back to the time of the Phoenicians. They used the compound 6,6'-dibromoindigotin, commonly known as Tyrian purple to dye the robes worn by royalty. Although iodinated compounds occur with less frequency, they are potentially the most important to man as they play an important role in human physiology. The thyroid gland contains high levels of the iodo-substituted amino acids monoiodotyrosine and diiodotyrosine. Whilst these have no biological activity in themselves, they are the precursors to the biological hormones 3,3',5,5'tetraiodothyronine and 3,3',5-triiodothyronine, which are responsible for growth and development.

In the light of the fact that there is such a large number of halogenated organic compounds and they have such a diversity of structures, it comes as little surprise to learn that there are pathways in a variety of organisms that have evolved to metabolise halogenated compounds.

1.2 Occurrence and use of synthetic halogenated compounds.

Increasingly since the 1950's, the chemical industry has developed large numbers of organic compounds with single or multiple halogen substitutions for a wide variety of purposes. These compounds are now used extensively as drugs, aerosol propellants, pesticides, solvents, paints and refrigerants. Due to the widespread use of these compounds, and the nature of their use, their presence in the environment has become a cause for concern as they now are posing a significant problem as pollutants. The compounds show a detrimental effect on mammals for three reasons. Firstly, uptake of the compounds can cause toxicity or cancer. Secondly, due to their persistent nature the chances for long-term exposure are greatly increased and finally, the compounds are able to bioaccumulate and increase in organisms higher up the food chain.

The reason for the recalcitrant nature of many of the synthetic organohalides is a consequence of the chemistry of the carbon-halogen bond. For example the fluorocarbons are particularly inert, with a bond dissociation energy of 105 kcal/mol for the carbon-fluorine bond, compared with 78 kcal/mol for the carbon-chlorine bond (Goldman 1972). This inert nature has resulted in the persistence of fluorocarbons in the atmosphere, where they undergo ultraviolet photodecomposition forming free-radicals which then degrade the ozone layer, a matter of international concern. Chlorinated compounds vary in their environmental persistence. Aryl and alkenyl chlorides decompose much more slowly then alkyl chlorides. This is because the latter compounds undergo hydrolytic and photolytic cleavage of the carbon-halogen bond much more readily then the former compounds (Wackett 1991).

One of the major uses for synthetic chlorinated organic compounds is as herbicides and pesticides. A huge variety of pesticides are now available, differing greatly in their structures, effectiveness and their environmental persistence. The insecticide DDT (1,1,1-trichloro-2,2-di-[4-chlorophenyl]ethane) and the cyclodiene insecticides are widely thought to be amongst the most recalcitrant compounds in the environment. However, it has been shown that even these compounds do undergo degradation in the environment, both abiotic and biotic (Neilson 1990).

A wide variety of herbicides, which differ greatly in their structure and function are used globally, but one of the simplest groups of compounds are the substituted aliphatic acids. Only a few of the many possible structures appear active and the compounds with the most widespread use are trichloroacetic acid (TCA) and 2,2-dichloropropionic acid (DCP). Other aliphatic acid herbicides include 2,2,3-trichloropropionic acid, 2,2-dichlorobutyric acid and 2,3-dichloroisobutyric acid. In 1953 the Dow

Chemical Company began marketing a herbicide, the main active ingredient of which was DCP. This herbicide was called Dalapon and went into widespread use for the control of Graminaceae species of weeds.

1.3 Properties and uses of halogenated aliphatic acids

As has been said, these compounds are amongst the simplest herbicides. The halogenated aliphatic acids (HAA's) are carboxylic acid derivatives with one or more of the aliphatic chain hydrogens replaced by a halogen, which results in an increased degree of ionisation of the carboxyl group due to the electronegativity of the halogen. This means that the substituted aliphatic acids are stronger acids then the parent non-halogenated acids. The position of the halogen substitution has an effect on the degree of ionisation, the closer to the carboxyl group, the greater the inductive effect and the number of substitutions also has an effect, the more substitutions, the greater the degree of induction.

The short chain HAA's are all water soluble liquids or crystalline solids, but their solubility decreases with an increase in chain length. All decompose slowly at room temperature but decompose more quickly upon heating. During boiling the compounds undergo dehalogenation due to a nucleophilic attack on the substituted carbon atom by hydroxyl ions. This occurs because the carbon atom is rendered susceptible due to the electronegative effect of the halogen. This susceptibility to nucleophilic attack is thought to be responsible for the biological dehalogenation of these compounds and this will be discussed in more detail later.

HAA's were first patented as herbicides in the early 1950's (Leasure 1964) and a number have been used for this purpose. There is a definite structure-activity relationship. Firstly, the aliphatic acids must be substituted with a halogen, all unsubstituted acids show no herbicidal

activity. Only α-chlorinated aliphatic acids show activity, and acids substituted at other positions must also be substituted at the α-carbon. Indeed, substitutions at positions other than the α-carbon may have the effect of diminishing the herbicidal effect of the HAA's. Increasing the chain length of the HAA has the effect of weakening the herbicidal activity, to the extent that 2,2-dichlorovaleric acid (5-carbon) is only weakly active and 2,2-dichlorohexanoic acid (6-carbon) is completely inactive (Leasure 1964). This may be due to the decreased solubility of the longer HAA's, which may decrease uptake in the plant and the cells, so reducing the phytotoxic effect. Substitution of halogens other than chlorine result in a decrease in activity, possibly due to the lower electronegativity of bromine or iodine, resulting in a weaker acid.

The majority of studies into the effect and fate of HAA's have concentrated on TCA and DCP. These have shown that the compounds are absorbed into the plant through the roots and shoots and then translocated throughout the plant. However, higher plants seem unable to metabolise the HAA's once they are present the plant, indeed it has been shown that DCP remains in plants for at least three generations, showing phytotoxic effects (Ashton and Crafts 1973). However, the effect of HAA's on plants is quite marked and wide-ranging. They cause formative effects, growth inhibition, leaf chlorosis and necrosis and eventually the death of the plant if applied at high enough concentrations (Ashton and Crafts 1973). The results of many studies indicated that the effect of HAA's was not on the production of metabolic energy but rather in the utilisation of that energy. DCP has been reported to alter carbohydrate, lipid and nitrogen metabolism, but these responses are more likely to be as a consequence of the secondary action of the herbicide, rather than a result of the direct action of DCP. It is known that DCP is able to precipitate proteins and that halogenated propionates and acetates are able to alkylate the sulphydryl or

amino groups of proteins (Ashton and Crafts 1973), and so it may be these properties that are responsible for the *in vivo* activity of HAA's. At low concentrations they may cause conformational changes in enzymes resulting in a wide range of disrupted metabolic functions. This may be the reason that the effects of HAA's are seen in a number of metabolic pathways.

A second use for HAA's and one which is now of considerable interest is in the manufacture of more complex herbicides. HAA's with a single substitution at the 2-carbon (2-HAA's) can be used as intermediates in the preparation of pyridyloxy-phenoxy substituted alkanoic acids, which are herbicides active against grass species (Taylor 1984). These herbicides can exist in two enantiomeric forms, but the D- form of the herbicide is the only isomer to show biological activity, and it is possible to produce this form from the L- isomer of 2-HAA's after removal of the HAA D- isomer by optical isomer separation. However, as will be described later, there is now a biological approach to the problem of resolution of the isomers that may aid greatly in this process.

With the exception of fluoroacetate and other long chain fluoroalkanoic acids, the presence in the biosphere of HAA's has been due to their production and use by man. As such they are classed as a group of compounds termed xenobiotics. It may be reasoned that because these compounds have only recently been introduced into the environment and have very few naturally occurring close analogues, the metabolic pathways necessary to deal with their degradation would not exist and so these compounds would persist. However, investigations into the persistence of HAA's in soil and identification of the organisms responsible for their degradation have shown this is not the case.

1.4 Microbial degradation of halogenated aliphatic acids

Shortly after their first use as herbicides investigations were begun into the fate of the HAA's once they had come into contact with the soil. Since this research was started, many species of microorganism have been reported that are capable of degrading HAA's, indeed they are some of the least persistent xenobiotics. The first report of organisms able to degrade HAA's was made in 1957 by Jensen. Using enrichment techniques with monochloroacetate (MCA) as the selecting carbon source he was able to identify five strains of bacteria, all of which proved similar. All five strains were able to grow solely on MCA accompanied with the release of Cl^-. Growth was greatly enhanced by the addition of 0.05% yeast extract and a similar effect was seen upon the addition of peptone or numerous amino acids. Monobromoacetate was decomposed as readily as MCA by two of the strains, although at concentrations above 0.05M, MBA had a growth inhibitory effect. Dichloroacetate (DCA), 2-chloropropionate (2CP) and DCP were also degraded, but to a much lesser extent then MCA. TCA and 3-chloropropionate (3CP) were not decomposed (Jensen 1957). All five strains were tentatively identified as non-fluorescent, non-proteolytic pseudomonads. Using DCP as the selective carbon source, Jensen was again able to isolate five strains, which were essentially similar and these were identified as *Agrobacterium*. They grew readily on DCP in basal medium enriched with yeast extract or peptone, with DCA and 2CP also being attacked, but not MCA or TCA. Jensen attributed the dehalogenation of the HAA's to a substrate-induced enzyme, which is present in only a limited number of soil bacteria (Jensen 1957).

A further six organisms, tentatively assigned to either or both *Agrobacterium* and *Alcaligenes* ,were isolated by Magee and Colmer (1959) using DCP as the selective carbon source. Hirsch and Alexander (1960)

used enrichment techniques, with DCP as the carbon source, to isolate eight organisms, five strains of *Nocardia* and three strains of *Pseudomonas*. One representative each of these species was chosen for further study. Both of the organisms were able to grow on minimal medium with DCP alone as the sole carbon and energy source, and the addition of yeast extract and calcium carbonate served to increase the growth rate and the extent of dehalogenation. The change in growth rate was attributable to the yeast extract, whilst the increase in the extent of dehalogenation was a consequence of the addition of calcium chloride, most probably due to its role in the prevention of acid accumulation. The use of radiolabelled ^{14}C-DCP, showed that there was labelling of respiratory carbon dioxide and cellular materials, indicating that the carbon skeleton of DCP was metabolised further by the bacteria. The substrate specificity of the organisms was investigated and the *Nocardia* sp. was shown to be active against all monohaloacetates except fluoroacetate, whilst the *Pseudomonas* sp. was able only to metabolise these compounds with difficulty, if at all. TCA was degraded only by the *Pseudomonas* sp., but tribromoacetate was degraded by both of the organisms. The degradation of 3-monohalopropionates occurred at a slower rate than the 2-substituted alkanoic acids. This finding is in contrast to those of Kaufman (unpublished, reported by Kearney *et al.*, 1965), who isolated nine species of bacteria based on their ability to metabolise DCP, and found that eight of these isolates were able to dehalogenate 3-chloropropionate at a faster rate than that seen for 2CP. Two of these organisms were also shown to be able to degrade 2,3-dichloroisobutyrate more rapidly than 2,2-dichloroisobutyrate.

A trichloroacetate-degrading organism was isolated by Jensen (1957). The organism was unable to degrade TCA in minimal medium even after the addition of yeast extract, but grew well on soil medium (soil, water,

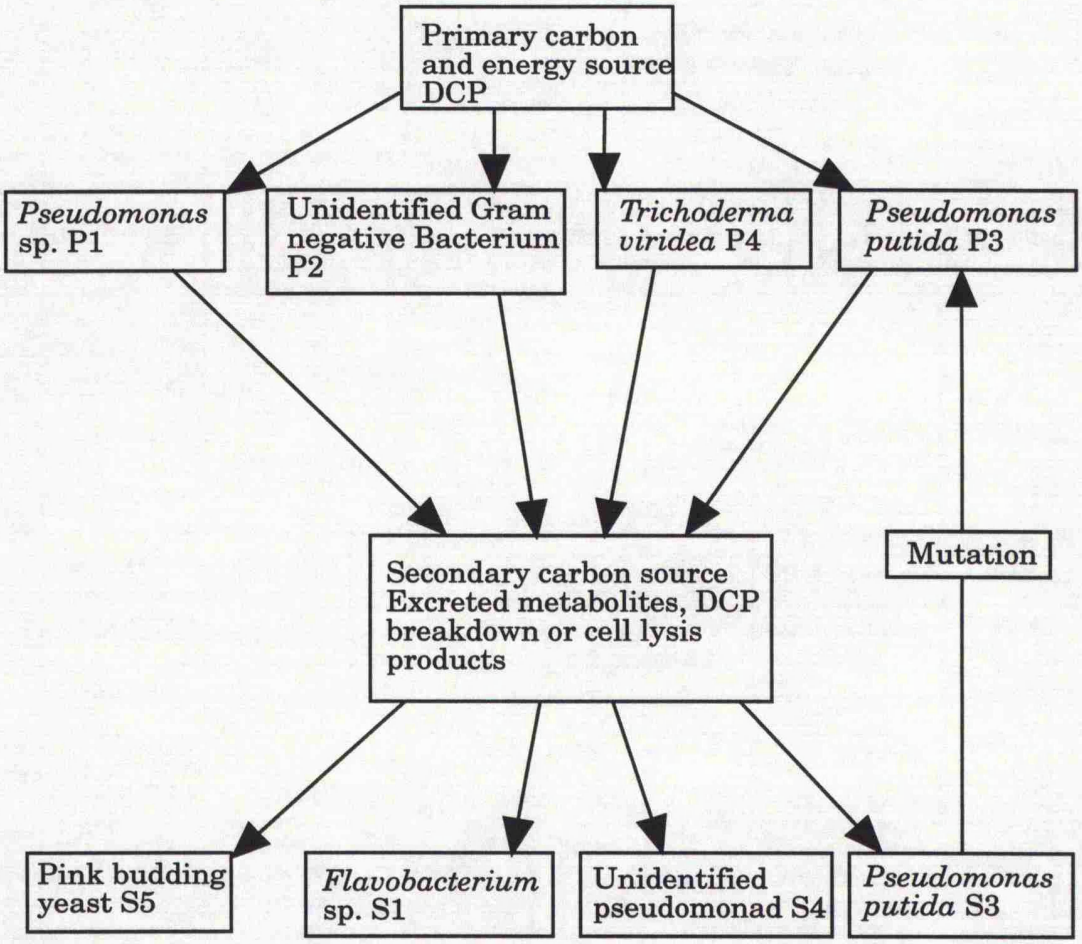
salts ± agar) supplemented with TCA below concentrations of 0.05M. The only other HAA that was degraded was DCA, but at a very reduced rate. Jensen went on to isolate a further eight strains similar to those previously reported (Jensen 1960) and these were identified as *Arthrobacter* sp. Weightman and co-workers isolated a two-membered microbial community able to dehalogenate TCA, but not to utilise it as the sole carbon and energy source(Weightman *et al.*, 1992).

Although many species have now been reported able to utilise 2-halogenated alkanoic acids, very few have been isolated on the ability to utilise 3-chloropropionic acid as the sole carbon and energy source. An isolate able to achieve this was obtained by Bollag and Alexander (1971) and identified as *Micrococcus denitrificans*. Investigations into the substrate specificity of cell-free extracts from this organisms showed that the dehalogenase activity was specific for halogen substitutions at the ß-position. Halogens at the α-position were not attacked, and where substitutions were at both positions, e.g.. 2,3-dichloropropionic acid, only the ß-position was attacked. A pseudomonad was isolated from soil by Castro and Bartinicki (1965) which was able to metabolise 3-bromopropanol via 3-bromopropionic acid and 3-hydroxypropionic acid to acetate, carbon dioxide and free bromide ion, when grown in minimal medium containing glucose, yeast extract and 3-bromopropanol. The rate-limiting step of the pathway was shown to be the ß-oxidation of the carbinol moiety, with the dehalogenation step being quite rapid. An organism able to grow on both α-chlorinated aliphatic acids and ß-chlorinated four carbon aliphatic acids has been isolated and identified as *Alcaligenes* sp. strain CC1(Kohler-Staub and Kohler 1989) A variety of organisms have also been isolated by their ability to utilise haloalkanes as their carbon and energy source. These include *Xanthobacter* (Janssen *et al.*, 1985), *Ancylobacter* (Van den Wijngaard *et al.*,

1992), *Pseudomonas* and *Arthrobacter* (Van den Wijngaard *et al.*, 1989) and *Corynebacterium* (Yokota *et al.*, 1987).

With the exception of Weightman *et al.*, (1992), all of the described research has been of organisms in monoculture, and not of mixed populations, a situation more analagous to that seen in the environment. Jensen (1957) reported that dehalogenation occurred much more rapidly in soil than in monoculture and speculated upon the causes of this increase, including the various nutritional and physical interactions between the organisms present. In order to investigate a microbial community more closely Senior *et al.*, (1976), used continuous flow enrichment techniques to isolate and maintain a community capable of growth on DCP. This community was isolated from four different soil samples on separate occasions and had the same structure in terms of component organisms, which suggests that the community is very stable and a good representative of a natural community. The structure of the community is shown in Figure 1.1. The community consisted of three primary users, which grew on the DCP. These were a *Pseudomonas* sp., an unidentified gram-negative bacterium and *Trichoderma viridae*. There were four secondary users in the community, that is they grew not on the DCP but upon metabolites directly produced from the catabolism of DCP or excreted from the primary users, or upon compounds released into the medium as a result of cell death and lysis. These primary users were a *Flavobacterium* sp., *Pseudomonas putida* S3, an unidentified pseudomonad and a pink budding yeast. With the exception of the yeast, which was lost at dilution rates greater than $0.2h^{-1}$, the community remained stable for over 13,500h (Senior *et al.*, 1976). The most notable alteration in the structure of the community came after 2,900h of growth, when a fourth primary DCP utiliser appeared. This was identified as *P. putida* and was identical in all other respects, except for its ability to grow on DCP, to the secondary utiliser *P. putida* S3. The

Figure 1.1



Structure of the microbial community growing on 2,2-dichloropropionic acid (Senior *et al* 1976).

organism was designated *P. putida* P3. This organism was shown in monoculture to grow well on DCP as the sole carbon and energy source, unlike the parent strain. The change from strain S3 to P3 was achievable in continuous culture by growing strain S3 in the presence of DCP and after 1200h of growth, strain S3 was completely displaced by the new strain P3. *P.putida* P3 was later shown to have arisen from *P.putida* S3 by a mutation that considerably elevated the dehalogenase activities of the organism. Strain S3 possessed dehalogenase activities considered too low for growth, and so the mutation allowed an up-regulation in the level of expression of the dehalogenase activity giving rise to an organism capable of growth on DCP and 2CP.

1.5 Enzymology of halogenated aliphatic acid degradation

In order for HAA's to be metabolised by an organism, the essential step is the removal of the halogen substitution(s) prior to the channelling of the carbon skeleton into the central metabolic pathways. Whether this activity is carried out abiotically, for example by photo-degradation, or by enzymes breaking the carbon-halogen bond depends upon the properties of the organic compound. Organisms are known to induce specific enzymes that catalyse the cleavage of the carbon-halogen bond and are active against a wide variety of xenobiotic compounds. Enzymes that carry out this reaction by design rather than as a fortuitous reaction are termed dehalogenases (Jensen 1957).

There are three main reactions whereby the halogen may be removed from the carbon backbone, namely dehydrodehalogenation, reductive dehalogenation and hydrolytic dehalogenation. Where dehalogenation occurs spontaneously due to the modification of the carbon skeleton, the elimination of the halide is thought to be via one of the mechanisms listed. There are several alternative mechanisms other than the three mentioned

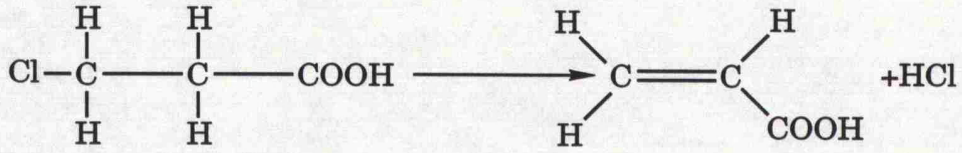
by which dehalogenation can occur. These alternatives will also be outlined.

Dehydrodehalogenation is the mechanism whereby the halogen and a hydrogen are removed simultaneously resulting in a carbon-carbon double bond being formed. This is believed to be the mechanism responsible for the dehalogenation of 3-chloropropionic acid by extracts of *Micrococcus denitrificans* (Bollag and Alexander 1971). After studying the products of the dehalogenation of ^{14}C labelled 3-chloropropionate using thin-layer chromatography they determined that dehalogenation produced acrylic acid, which could only be formed by a dehydrodehalogenation mechanism. NADP was seen to be stimulatory to the reaction, enhancing the rate of 3-CP dehalogenation by 30%. Whilst this is not a significant enhancement, indicating that NADP is probably not a co-factor, the increase in the rate may possibly be due to its role as an proton acceptor. Dehydrodehalogenation is also seen to occur in the metabolism of chlorophenoxyacetates. Extracts from a *Pseudomonas* sp. were able to catalyse the dehalogenation of 4-chloro-2-methyl muconic acid. The product, 2-methyl-4-carboxymethylene-but-2-enolide, was formed by a lactonisation reaction as a consequence of the dehydrodehalogenation. These reactions are shown in Figure 1.2a and Figure 1.2b. (Gaunt and Evans 1971).

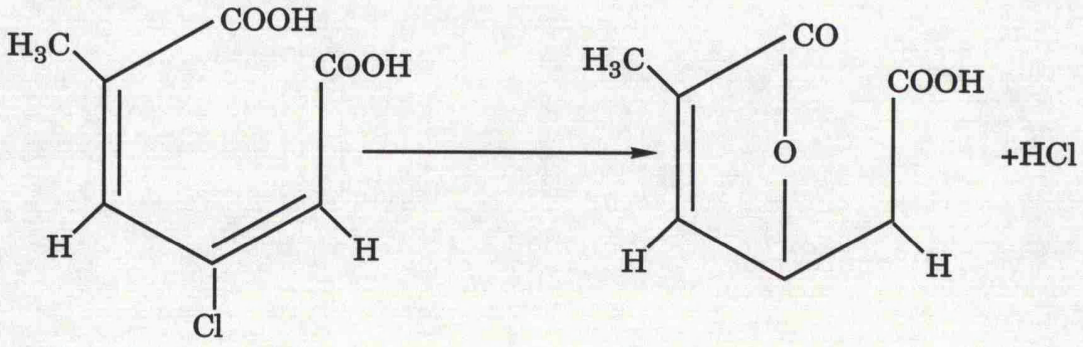
Reductive dehalogenation is either the direct replacement of a halogen substitution with a hydrogen or the removal of two halogens from adjacent carbon atoms with the formation of an additional bond between the carbon atoms (Castro and Belser 1968). An example of the first reaction is the conversion of DDT to dichlorodiphenyldichloroethane (DDD), a process which is seen to occur anaerobically, the aerobic dehalogenation of DDT being by a dehydrodehalogenation mechanism. Reductive

Figure 1.2

(a)



(b)



(a) Dehydrodehalogenation of 3-chloropropionic acid by *Micrococcus denitrificans* (Bollag and Alexander 1971).

(b) Dehydrodehalogenation of 4-chloro 2-methyl muconic acid by a *Pseudomonas* sp. (Gaunt and Evans 1971).

dehalogenation of vicinal dibromides by soil-water cultures has been reported (Castro and Belser 1968), but the organisms responsible for the reaction were never isolated and so the presence of specific reductive dehalogenases was not confirmed. Reductive dehalogenation of short chain HAA's has not been previously reported, possibly due to the fact that the reactions take place in anaerobic conditions and there have been very few investigations under these conditions.

By far the most widely reported dehalogenation mechanism is that of hydrolytic dehalogenation. This involves the substitution of a hydroxyl group for the halogen (Goldman 1972). The removal of the halogen is often the first step in the metabolism of halogenated aliphatic compounds. The predicted products of this reaction are hydroxy-acids for the monohalogenated compounds and oxo-acids for the dihalogenated compounds (Jensen 1957). These predicted products have been confirmed by several investigations. The enzymes responsible for the hydrolytic dehalogenation are termed halidohydrolases, but due to their abundance in the biosphere compared to other dehalogenating enzymes they are commonly called dehalogenases. The many studies that have been carried out so far, which will be described in more detail later, have defined several specific characteristics of dehalogenase enzymes. The mechanisms by which the dehalogenases act will be discussed here, the biochemical and genetic characterisation of the dehalogenase enzymes will be discussed later (Section 1.6).

The most widely studied hydrolytic dehalogenases are those that are active against haloalkanoic acids and these can be subdivided into two groups. The first group is the haloacetate dehalogenases, (EC 3.8.1.3), which are only active against haloacetates and the second group are the 2-haloacid dehalogenases, (EC 3.8.1.2), which act upon haloacetates, 2-

halopropionates and some act on 2-haloalkanoic acids with greater carbon chain lengths. Each of these groups can then be subdivided further based upon substrate specificity. The haloacetate dehalogenases can be subdivided into two groups based around their reactivity towards fluoroacetate. These groups are termed A1 for enzymes that will attack the C-F bond and A2 for enzymes that are unable to (Leisinger and Bader 1993). The first group are represented by the H-1 dehalogenase from a *Moraxella* sp. (Kawasaki *et al.*, 1981a). This enzyme is active against fluoroacetate and other halogenated haloacetates and has been shown to be plasmid-encoded. A representative of the second subgroup, A2, is the H-2 dehalogenase from the same *Moraxella* sp. This is active against all haloacetates except monofluoroacetate and is encoded on the same plasmid as H-1. The biochemical and genetic relationships between the two dehalogenases will be discussed later. (Section 1.6.1, 1.6.2)

The second group of haloalkanoic acid dehalogenases, the 2-haloacid dehalogenases show a greater degree of diversity. They, too, can be subdivided on the basis of their substrate specificity, although this is based upon chiral selectivity rather than the nature of the C-halogen bond. They can also be divided upon the basis of their substrate/product configuration, i.e. whether the substrate and product have the same optical orientation or if a change of configuration has accompanied the dehalogenation reaction. Therefore the 2-haloacid dehalogenases can be subdivided as follows. The first group, B1, are active against L-2-haloacids and cause an inversion of configuration. The second group, B2, are active against D- and L-2-haloacids and cause an inversion of configuration, whilst the group B3 are active against both isomers and retain the configuration between substrate and product. The final group B4 are active against D-2-haloacids and cause an inversion of the configuration (Leisinger and Bader 1993). The mechanisms by which the reactions are carried out have not been

completely elucidated but several proposals have been made. A representative of group B1 is the dehalogenase from *P.dehalogenans* described by Little and Williams (1971). The enzyme only attacked L-2CP with the resulting product being D-lactate and was seen to be insensitive to sulphydryl blocking agents. The proposed reaction mechanism for the inversion of configuration is shown in Figure 1.3. This generalised base catalysed reaction was proposed involving a catalytic histidine residue that acts as an electron donating group activating a water molecule which attacks the substrate. This mechanism is applicable to groups B2 and B4, where there is also an inversion of configuration. Some of the dehalogenases that invert the configuration of the product are sensitive to the action of sulphydryl blocking reagents. A mechanism has been proposed that accounts for this with a reactive thiol group on the enzyme attacking the substrate, leading to an inversion of the configuration. This mechanism is shown in Figure 1.6. The group B3 is distinctive in that it retains the configuration of the product. The mechanism for this reaction is proposed to be a double inversion (Weightman *et al.*, 1982), resulting in an overall retention of the configuration, via a thioether intermediate. A representative of this group of dehalogenases is Fraction II from *P.putida*PP3. This enzyme is susceptible to sulphydryl blocking agents, which would be consistent with the proposed mechanism. The reaction mechanism for the group B3 is shown in Figure 1.4.

The dehalogenation for dihalogenated alkanoates is believed to be slightly different. The first halogen is removed by a enzyme-catalysed nucleophilic substitution, equivalent to the dehalogenation reactions seen for the monohalogenated acids. The second halogen is believed to be released spontaneously as the intermediate, a hydroxy-haloacid, is unstable. This reaction is shown in Figure 1.5.

Figure 1.3

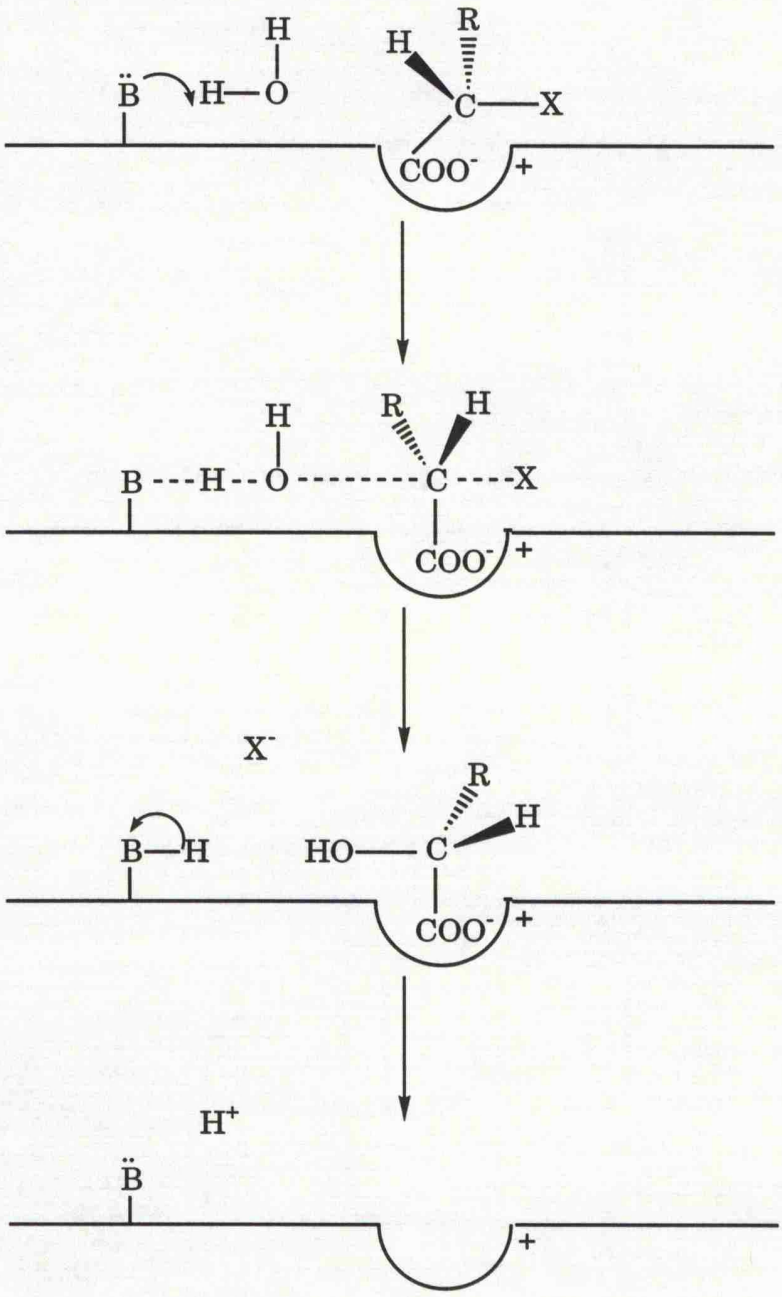


Figure 1.3

Dehalogenation mechanism resulting in an inversion of configuration. X=halogen, R= alkyl group. [After Hardman (1991)].

Figure 1.4

Dehalogenation reaction involving a retention of configuration via a double inversion and thioether intermediate. R=alkyl group, X=halogen. [After Hardman (1991)].

Figure 1.4

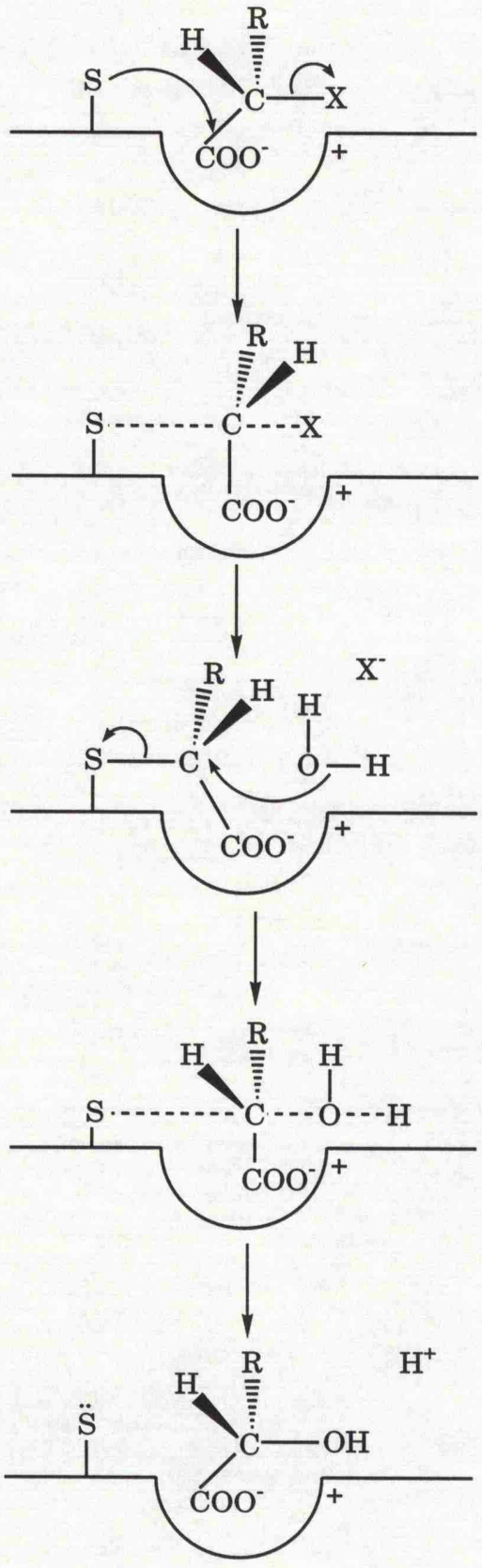


Figure 1.5

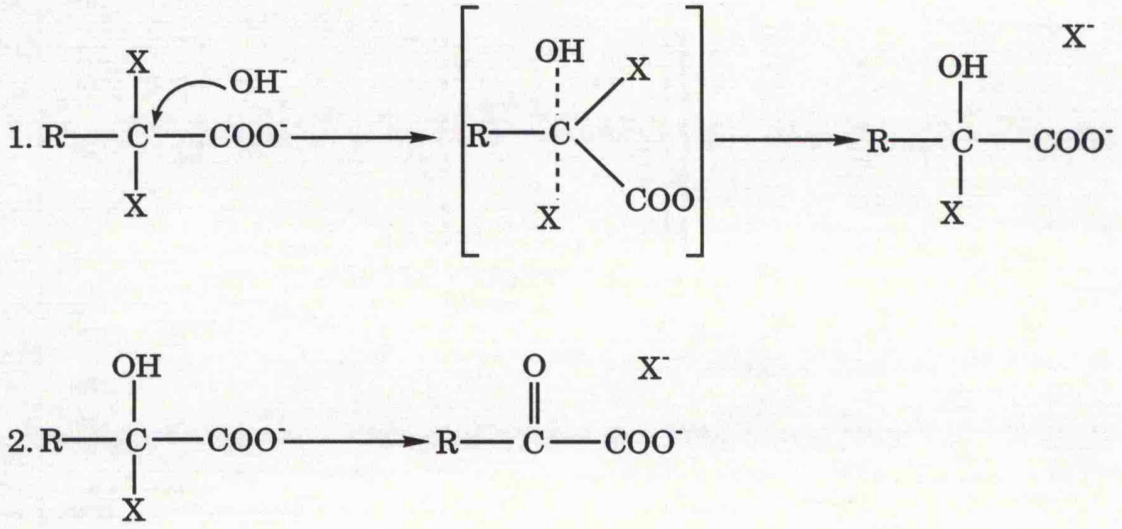


Figure 1.5

Mechanism for the dehalogenation of dihaloalkanoic acids based upon the dehalogenation of DCP by an *Arthrobacter* sp. R=alkyl group, X=halogen. Step1 is an enzymic reaction, step 2 is spontaneous. [After Hardman (1991)].

Figure 1.6

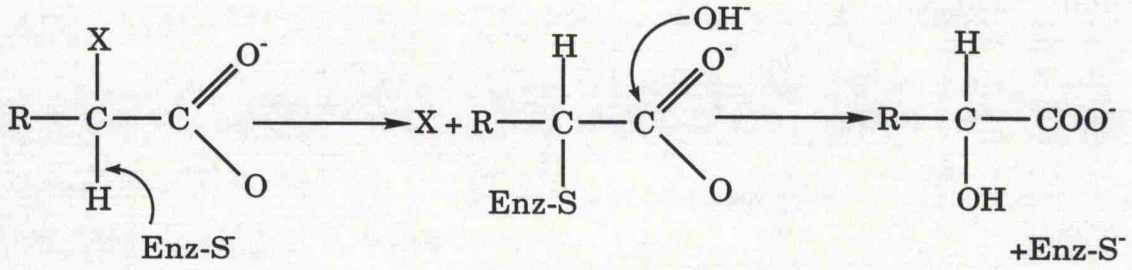


Figure 1.6

Mechanism proposed by Goldman (1968) involving a sulphydryl group and leading to an inversion of configuration. [After Hardman (1991)].

As mentioned earlier there are other less common reactions that have been described. A dehalogenation reaction involving the co-factor glutathione has been reported for the degradation of dichloromethane by *Hyphomicrobium* sp. strain DM2. The reaction results in the formation of formaldehyde and inorganic chloride by the following mechanism. A chloromethyl-glutathione conjugate is formed enzymatically which then undergoes nonenzymatic hydrolysis to yield hydroxymethyl glutathione, which is then broken down to form formaldehyde and reduced glutathione (Kohler-Staub and Leisinger 1985). The reaction is shown in Figure 1.7a. The dehalogenation of haloalcohols has been studied and has been shown to be carried out by enzymes termed halohydrin hydrogen-halide lyases, which have been purified and characterised (Leisinger and Bader 1993). These enzymes catalyse the dehalogenation of mono- and dihalogenated alcohols to form their corresponding epoxides. The general equation for this reaction is shown in Figure 1.7b. An investigation into the anaerobic degradation of DCP showed that the removal of DCP was associated with the formation of methane, but the mechanism by which this process occurred was not elucidated (Shanker and Robinson 1991). A study of TCA degradation by a mixed culture revealed that the product of TCA dehalogenation was not oxalate, which would be expected for the hydrolytic reaction, but carbon dioxide. Each mole of CO_2 released was accompanied with the release of 3 moles of Cl^-. It was thought unlikely that following dehalogenation the carbon skeleton was immediately stoichiometrically converted to CO_2, and so it was believed that the dehalogenation process caused the concomitant cleavage of the carbon-carbon bond to yield CO_2. However, a reaction that does not involve a specific dehalogenase is also possible. If TCA is decarboxylated then it spontaneously forms dichlorocarbene, which is highly reactive and will spontaneously hydrolyse to form formic acid and/or CO with HCl. The enzymes responsible for this reaction were not isolated

Figure 1.7

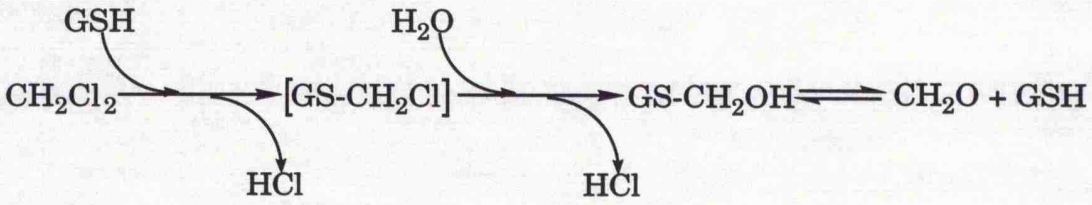
(a) Dehalogenation of dichloromethane via a glutathione-dependent dehalogenase from *Hyphomicrobium* sp. strain DM2. GSH and GS represent free and bound glutathione respectively. (Kohler-Staub and Leisinger 1985)

(b) General equation for the dehalogenation of haloalcohols. R=H;X, X=halogen. (Leisinger and Bader 1993).

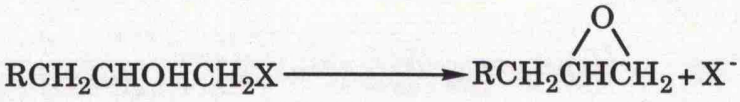
(c) Postualted degradation of TCA via a decarboxylation reaction (Weightman *et al* 1992)

Figure 1.7

(a)



(b)



(c)

$CCl_3COOH + OH^- \longrightarrow CCl^-_3 + CO_2 + H_20 \longrightarrow :CCl_2 + HCl$

$:CCl_2 + 2H_2O \longrightarrow HCOOH + 2HCl$ or $:CCl_2 + H_2O \longrightarrow CO + 2HCl$

and so the biochemistry remains to be elucidated. The reaction is outlined in Figure 1.7c (Weightman *et al.*, 1992).

1.6 Characteristics of dehalogenase enzymes

Many dehalogenases have now been isolated and characterised and this section will deal with the properties of the reported enzymes. The first section will comprise of the biochemical data known about the reported dehalogenases, whilst the second will deal with the genetic information known, an area that has expanded more recently.

Hardman and Slater (1981) were able to isolate sixteen bacterial strains from soil that were able to utilise 2CP or MCA as their sole carbon and energy source. These isolates could be subdivided into five groups on the basis of their differences in dehalogenase activity. Group A had a dominant MCA dehalogenase activity, group B had DCA dehalogenase activities 1.5-2.5 times greater than their MCA activity. Group C contained a single isolate that had an unusually low DCP activity, group D also contained a single isolate that was unusual in that it had a high 2CP and DCA activity whilst group E comprised of one organism with an intermediate level of DCA activity and an overall dehalogenase activity ratio similar to that of group A organisms. The dehalogenase activities were analysed by non-denaturing polyacrylamide electrophoresis and a total of four electrophoretically distinct dehalogenases were identified. Three of the isolates contained one, three and four dehalogenases respectively, whilst the remainder contained various combinations of two dehalogenases. In some cases electrophoretically identical dehalogenases seemed to be represented in several strains, in others enzymes with the same electrophoretic mobility had different substrate specificities. Two possibilities were postulated for the common occurrence of multiple dehalogenases. Firstly, as the dehalogenases represent the rate-limiting

enzyme, the concentration of dehalogenase in the cell will have a direct influence on the organisms growth rate, so selection on the basis of growth would favour organism's with multiple dehalogenase genes or the acquisition of other dehalogenase genes for other organisms. Secondly, the presence of a number of dehalogenases with broadly similar activities may represent organisms in an intermediate stage of evolution.

1.6.1 Biochemical characterisation of dehalogenases

Although the first reports of the dehalogenation of HAA's were in 1957, it was later that the first enzymes responsible for the dehalogenation reaction were isolated. Goldman (1965) isolated a soil pseudomonad that showed the ability to dehalogenate fluoroacetate to form glycollate and free fluoride ion, shown in the equation below:

$$FCH_2COO^- + OH^- \longrightarrow HOCH_2COO^- + F^-$$

The enzyme responsible was partially purified and shown also to be active against monochloroacetate and monoiodoacetate, albeit at a much reduced rate compared to the activity seen for the enzyme against fluoroacetate, one sixth and one twohundreth of the rate, respectively. The enzyme was not able to dehalogenate difluoro- and trifluoroacetate, or 2-fluoropropionate and 3-fluoropropionate. The enzyme was found to have an alkaline pH optimum of between pH9.0 and pH9.5, and was susceptible to sulphydryl-blocking agents. A reaction mechanism was postulated that proceeded via and alkyl-enzyme intermediate. The first step, that is the formation of the of the enzyme-substrate intermediate would be rate limiting and the rate would be dependant on the nature of the halogen substituent. Goldman *et al.*, (1968), isolated a soil pseudomonad that had dehalogenase activity. The organism was able to utilise monochloroacetate and dichloroacetate as the sole sources of carbon and energy and was shown to synthesis two

dehalogenases, the ratio of which varied depending on the carbon source for growth. Cell-free extracts from dichloroacetate-grown cells were used for the purification of the two enzymes, which were readily separable by gel-filtration and termed Halidohydrolase I and Halidohydrolase II. Halidohydrolase I was the only enzyme when the cells were grown on monochloroacetate whilst Halidohydrolase II was the most abundant in cells grown on dichloroacetate.. The enzymes were different in several respects. Firstly, although they shared common substrates, the rates of dehalogenation were quite different. Halidohydrolase I was nine times more active towards monochloroacetate than towards dichloroacetate, whilst Halidohydrolase II was twice as active against dichloroacetate than monochloroacetate. The two enzymes differed in their thermal stability, Halidohydrolase I being less stable. Both enzymes showed a broad pH optimum in the alkaline region, like that seen for the haloacetate dehalogenase studied previously (Goldman 1965), but differed from this enzyme in their sensitivity to sulphydryl-blocking agents and there inability to dehalogenate monofluoroacetate. Both enzymes acted on 2-chloropropionate and 2-chlorobutyrate but only on the L-isomers and the products were shown to have undergone an inversion of configuration, i.e. the products were in the D- configuration. There were two mechanisms postulated for the reaction, the first a simple nucleophilic displacement of the halogen by a hydroxyl ion from water, the second a nucleophilic displacement in which a residue of the enzyme is involved in donating electrons, the reaction mechanism shown in Figure 1.3.

The first dehalogenase able to act on dichloropropionate was partially purified by Kearney *et al.*, (1964). This enzyme was isolated from an *Arthrobacter* sp. that was able to utilise DCP as the sole carbon and energy source. The dehalogenase was active against DCP and also dichloroacetate, 2-chloropropionate and dichlorobutyrate. Goldman (1972) suggested that

the enzyme acted in a similar fashion to Halidohydrolases I and II, the difference in substrate specificity being due to the enzyme's ability to accommodate bulkier side groups at its active site. The mechanism for this reaction has already been explained in Figure 1.5. The dehalogenase properties of a strain of *Pseudomonas dehalogenans* isolated by Jensen (1960) were studied briefly by Davis and Evans (1962). Cell-free extracts were prepared from cells grown on glucose with monochloroacetate present and the dehalogenase was found to be active against monochloro-, monobromo- and monoiodoacetate, dehalogenating each at a similar rate. Dichloroacetate was dehalogenated at a reduced rate and trichloroacetate, monofluoroacetate and chloro-propionates were not dehalogenated at all. The enzyme also showed an alkaline pH optimum and was sensitive to sulphydryl-blocking agents.

Another *P.dehalogenans* strain (NCIB 9061), isolated by Jensen (1960) had its dehalogenase activity examined in some detail by Little and Williams (1971). The enzyme was prepared from cells grown on glucose with monochloroacetate present and differed from the previously studied *P.dehalogenans* dehalogenase (Davis and Evans 1962) in its substrate specificity. The monohalogenated acetates were all substrates for the dehalogenase with the exception of monofluroacetates, but the enzyme was also able to dehalogenate other HAA's. Dichloroacetate, L-2-chloropropionate and dichloropropionate were also dehalogenated by the enzyme, at a reduced rate compared to that of MCA. Trichloroacetate and 3-chloropropionate were not dehalogenated. Kinetic studies were carried out and apparent Michaelis constants were derived for all of the substrates and were in the range 0.09-0.30 mM. The enzyme was purified and the native molecular weight was determined by gel-filtration to be 15 kDa. The enzyme was shown not to be sensitive to sulphydryl-blocking reagents. As there was no activity towards chloroacetamide and iodoacetamide, which

resemble the corresponding acids in size and reactivity, Little and Williams (1971) suggested that a carboxyl group was necessary on the substrate for reactivity, possibly to allow the substrate to bind at a positively charged binding site. The results led to the proposal of the mechanism in Figure 1.3.

As described earlier, the work of Senior *et al.*, (1976) had led to the evolution of a strain of *Pseudomonas putida* that had acquired dehalogenase activities not seen in the parent strain. *P.putida* P3, (later termed PP3), has been the attention of much research into both the biochemical and genetic aspects of dehalogenation. Slater *et al.*, (1979) studied the growth characteristics and dehalogenase activity of *P.putida* P3. Although it was originally isolated from a culture growing in DCP as the sole carbon and energy source (Senior *et al.*, 1976), the organism was seen to grow more rapidly on 2-chloropropionate (Slater *et al.*, 1979). These two halogenated compounds represented the only chlorinated aliphatic acids capable of supporting growth and when grown in medium containing both a halogenated and non-halogenated carbon source there was no sign of diauxic growth. Both DCP and 2CP were able readily to induce dehalogenase activity when added to cultures growing exponentially on succinate, with little effect on the growth rate. This was in contrast to other chlorinated aliphatic acids that where poor inducers of dehalogenase activity with some also causing major disturbances in the growth of the organism. Cell-free extracts from cells grown on 2CP were able to dehalogenate a number of 2-chlorinated aliphatic acids, the major substrates being monochloroacetate, dichloroacetate, 2-chloropropionate and dichloropropionate. Halogenation in other positions resulted in little or no dehalogenation, but several of these compounds were inducers. When P3 was grown under different conditions of induction there were variations in both the specific and relative dehalogenase activities. For example, the

DCA:2CP activity ratio varied from 1.1:1.0 to 6.8:1.0 for induction with 3CP and DCA, respectively. These results, along with thermal denaturation studies led Slater *et al.* (1979) to propose the presence of more than one dehalogenase in *P.putida* P3. This proposal was confirmed by the work of Weightman *et al.* (1979). Cells were grown with 2CP as the sole carbon and energy source and the dehalogenases were purified by a process of ammonium sulphate fractionation and ion-exchange chromatography. This resulted in the separation of the dehalogenases into two distinct peaks, each of which exhibited different substrate specificities. The majority of the activity towards 2CP and DCP was associated with the peak I fraction, whilst the peak II fraction showed activity towards DCA and trichloroacetate. There was MCA activity associated evenly between both peaks. Weightman *et al.* (1979) suggested that *P.putida* P3 contained two dehalogenases, one of which is associated with the dehalogenation of chloropropionates, whilst the other has a far greater specificity towards chloroacetates. The strain P3 was then used in continuous flow experiments which resulted in the selection of two mutants with altered dehalogenase activities (Weightman and Slater 1980). The cells were grown on 2CP in the presence of 2-monochlorobutanoic acid (2MCBA) with the intention of selecting strains that were able to grow solely on 2MCBA. The mutants obtained, *P.putida* PP309 and *P.putida* PP310 differed from *P.putida* P3 in that they showed elevated levels of fraction I dehalogenase, with a corresponding increase in MCA activity and an alteration in the ratios between MCA and DCA dehalogenation. The growth of both mutants was seen to be inhibited by the presence of 2CP at a concentration greater than 0.3g carbon l^{-1}, whereas the parent strain was able to grow at a concentration of 1.0g carbon l^{-1}. Neither of the mutants were able to grow on 2MCBA as the sole carbon and energy source as they were unable to grow on the dehalogenation product 2-hydroxybutanoic acid. The mutation

leading to this change in activities was thought to involve a gene multiplication of both the dehalogenase gene and the uptake system, the latter thought to be responsible for the increased sensitivity to 2CP of the mutant strains. Work was then carried out to investigate the specificities of the two dehalogenases from *P.putida* P3 (Weightman *et al.*, 1982). The organism was seen to grow at the same rate on the individual stereoisomers of 2CP as on the racemic mixture. When the two dehalogenases were purified and their stereospecificities determined there were notable differences. Fraction I dehalogenase was active against both D- and L-2CP and the dehalogenation product, lactate, retained the configuration of the substrate. When fraction II dehalogenase was investigated it too was active towards both the stereoisomers of D/L-2CP, however the product formed was then in the opposite optical configuration to the substrate. Fraction I dehalogenase was much more susceptible to inhibition by sulphydryl-blocking agents than fraction II, the presence of DTT helping to protect the activity of fraction I. These stereospecificities make these enzymes markedly different to those previously described as they were able to dehalogenate both isomers of D/L-2CP, whereas the dehalogenases studied by Little and Williams (1971) and Goldman *et al.* (1968) where only able to dehalogenate the L-isomer. Even though the substrate specificities are the same, the enzymes are probably not related as they show such different reaction mechanisms. As fraction II dehalogenase inverts the configuration of the product, whereas fraction I dehalogenase maintains the configuration they must act by different mechanisms. Weightman (1982) postulated that fraction II operated by a base catalysed nucleophilic attack, as shown in Figure 1.3. The mechanism proposed for fraction II was the one shown in Figure 1.4, a double inversion via a thioether intermediate, although a possible fault in this mechanism is the inherent stability of the thioether bond (Little and Williams 1971).

As there seemed to be a link between the higher levels of gene expression and sensitivity to halogenated alkanoic acids, this area was investigated by the creation of further mutants. Growth inhibition of *P.putida* PP3 was seen when MCA or DCA were added to exponentially growing cultures with succinate as the carbon source. The extent of the inhibition depended on the structure and concentration of the inhibitor, with the most powerful inhibitor being monobromoacetate. Mutants that were resistant to MCA or DCA were readily isolated by plating cells onto medium containing succinate and either MCA or DCA. Half of the DCA^r mutants were unable to utilise 2CP for growth and a greater proportion of MCA^r mutants were also 2CP negative. All MCA^r mutants were DCA^r, and conversely, all DCA^r mutants were MCA^r (Weightman *et al.*, 1985). Further characterisation of the mutants showed four classes, which varied in their production of the two dehalogenases. Where there were lower levels of expression, or only one dehalogenase was expressed, growth inhibition was less pronounced. The model proposed to explain these results required a specific haloalkanoate permease as well as a dehalogenase to be present for growth to occur. Mutants were isolated that showed equivalent levels of dehalogenase activity but differed in their growth rates. The reason for this was thought to be a mutation in the uptake system, so transport was reduced, accompanied by a reduction in growth rate. Another interesting feature of the mutants was the very low detection of revertants. Mutants had been selected at high frequencies but were very stable. This would indicate that the mutation had involved a deletion of DNA that was associated with the genes of the dehalogenase system. This could be explained if the genes were encoded on a transposable element. Further investigation into this question will be discussed in section 1.6.2.

An organism able to utilise both D- and L-2CP was isolated and the dehalogenase activity characterised (Motosugi *et al.*, 1982a and 1982b). Cell-free extracts from the organism, *Pseudomonas* sp. strain 113, were found to have maximal dehalogenase activity towards L-2CP, with the activity towards D-2CP being approximately 70% that towards the L-isomer. There was also dehalogenase activity towards MCA, DCA, DCP and MCBA, but at much reduced rates. The dehalogenase activity was purified and found to contain a single non-specific dehalogenase, similar to the fraction II dehalogenase from *P.putida* P3 (Weightman *et al.*, 1979). The dehalogenase inverted the configuration of the product relative to the substrate, thus putting it into the B2 group of haloacid dehalogenases. The dehalogenase showed activity towards bromo- and iodoalkanoates and again showed a requirement for a free carboxyl group, as postulated by Little and Williams (1971). The pH optimum was at pH9.5 and the dehalogenase was resistant to the effects of sulphydryl-blocking reagents. The molecular weight of the dehalogenase was determined for the native protein and the subunit and was found to be 68 kDa and 35 kDa respectively, indicating that the protein had a dimeric native conformation.

A strain of *Pseudomonas* sp.CBS3 was isolated due to its ability to degrade 4-chlorobenzoate and was found also to contain two 2-haloacid dehalogenases (Klages *et al.*, 1983, Schneider *et al.*, 1991). Both of the dehalogenases, DehCI and DehCII were studied further. The dehalogenases were induced by chloroacetate but were not induced by 4-chlorobenzoate, indicating that the dehalogenase may not be part of the metabolism of haloaromatics, but has a separate metabolic function. DehCI was primarily active against haloacetates, except fluoroacetate. Dichloroacetate and dichloropropionate were dehalogenated to a limited extent and only the L-isomer of monochloropropionate was dehalogenated, with an inversion of configuration to form D-lactate. The dehalogenase is

therefore a representative of the B1 group of 2-haloacid dehalogenases. The dehalogenase had the characteristic alkaline pH optimum, the actual value being pH9.5 and there was no effect of sulphydryl-blocking reagents on activity, so the reaction mechanism is probably the generalised base catalysis reaction shown in Figure 1.3. The native molecular weight of the dehalogenase was determined by gel-filtration to be 41 kDa, whilst the subunit molecular weight was determined to be 28 kDa, so the native protein was thought to have a dimeric confirmation. The differences in the measured native molecular weight and that predicted for a dimer from the subunit molecular weight may be due to anomalous migration of the protein. DehCII was also active against haloacetates and was able to dehalogenate L-2CP, with an inversion of configuration (Mörsberger *et al.* 1991). The pH optimum of DehCII was pH9.5, in accordance to previously published findings and the native molecular weight was determined to be 64 000 Da with a subunit molecular weight of 29 000 Da, indicating that the native enzyme has a dimeric confirmation. The genes for both of these dehalogenases have been cloned and will be described later (Schneider *et al.*, 1991).

A further example of an organism containing dehalogenase activities for different haloorganic compounds is reported by Van Der Ploeg *et al.*, (1991). An organism, *Xanthobacter autotrophicus*, was isolated on the basis of growth on 1,2-dichloroethane, but was found to also contain a haloacid dehalogenase. The enzyme was purified and characterised. The dehalogenase acted on mono- and dihaloacetates and also on 2CP, but only on the L-isomer, forming D-lactate, so this dehalogenase is placed in the B1 class. The pH optimum was pH9.5 and there was little inhibition from sulphydryl-blocking agents. The gene encoding the haloacid dehalogenase was cloned and sequenced. This will be discussed in section 1.6.2.

A dehalogenase similar to fraction I dehalogenase from *P.putida* P3 (Weightman 1979) was isolated from three soil pseudomonads (Kocabiyik 1988). The isolates were selected by their ability to grow on 2CP as the sole carbon and energy source and their dehalogenase activity characterised in terms of substrate specificity and electrophoretic mobility. Cell-free extracts from all three isolates were most active against DCA, with activity also measured against MCA and 2CP. In all cases the activity towards DCA was approximately 2.5 times greater than the activity towards MCA. When non-denaturing polyacrylamide gels were run and the gels stained for dehalogenase activity, a common band was seen for all three isolates and this band had an R_f value identical to that of a known fraction I type dehalogenase from a soil isolate E4 (Hardman and Slater 1980). Further characterisation of the dehalogenases was not carried out. Kocabiyik and Türkoglu (1989) reported that *Pseudomonas* sp. strain 19S, originally isolated through 2,4-dichlorophenoxyacetic acid soil enrichment procedures was able rapidly to degrade MCA. The specificity of the extracts was determined and the dehalogenase was found to act on MCA and iodoacetate. DCA was dehalogenated to a much lesser extent, and TCA and DCP were not substrates. The enzyme showed the usual pH optimum, with maximal activity being obtained at approximately pH9.8. The most notable feature of this dehalogenase is the thermal stability. The optimum temperature was 60°C and if the enzyme was incubated at this temperature for 15 minutes, 56% of the activity remained. The effect of sulphydryl-blocking agents and activity towards other substrates was not reported.

An organism isolated using enrichment techniques with DCP as the carbon source was shown to have a single dehalogenase active against DCP (Busto *et al.*, 1992). The organism was identified as *Pseudomonas alcaligenes* and cell-free extracts where shown to be active against 2CP, MCA and DCA as well as DCP. The dehalogenase had a pH optimum of

pH9.5, in line with the pH optimum of other reported 2-haloacid dehalogenases and showed only a single band after activity stains of native polyacrylamide gels using a range of substrates. No attempt was made to determine the mechanism or the stereospecific activities of the dehalogenase.

A notable piece of work is the characterisation of the dehalogenases of *Pseudomonas putida* AJ1/23. This organism was isolated from soil exposed to chloropropionate and is of interest because of the dehalogenase activity towards D-2CP. Two dehalogenase enzymes were isolated from the organism and characterised further (Smith *et al.*, 1990, Jones *et al.*, 1992). The two dehalogenases were shown to be stereospecific for the D- and L-isomers of 2CP. The D-2-haloacid dehalogenase was purified from cells grown on 2CP. The dehalogenase had a native molecular weight of 135 kDa and a subunit molecular weight of 31.8 kDa, indicating a tetrameric native conformation. The pH optimum was pH9.5, with the enzyme stable over the pH range pH6-9. When the substrate specificity was determined, the dehalogenase was shown to be active against 2-chloro- and 2-bromopropionate and monochloro- and monobromoacetate. There was some activity towards DCP and monohalobutyrates. The stereospecificity was unusual in that the dehalogenase only attacked the D-isomer, with the formation of L-lactate, so the dehalogenase caused an inversion of configuration. As such the dehalogenase is a member of the B4 group of 2-haloacid dehalogenases. The D-2CP specific dehalogenase showed little susceptibility to sulphydryl-blocking agents and the reaction mechanism was postulated to be that proposed involving base catalysis. The L-2CP specific dehalogenase from *Pseudomonas putida* AJ1/23 was purified after the gene coding for the enzyme had been cloned and expressed (Jones *et al.*, 1992). The dehalogenase was shown to have a native molecular weight of 79 kDa, with a subunit molecular weight of 26 kDa. The authors suggest

that the protein exists as a tetramer, although they do not exclude the possibility of a trimer as the native conformation. Studies of the stereospecificity of the dehalogenase showed that the enzyme attacked only L-2CP with the formation of D-lactate. The mechanism by which the reaction takes place is presumably the same as that of the D-2CP specific dehalogenase, although the susceptibility of the L-2CP specific dehalogenase to sulphydryl-blocking reagents has not been reported.

Berry *et al.* (1979) reported the isolation of a *Rhizobium* sp. capable of growth on HAA's. This organism was isolated using an enrichment procedure with DCP as the selective carbon source. The organism was also able to grow on 2CP but was unable to grow on MCA or DCA. Growth on DCP, 2CP and on lactate in the presence of DCA and 3-chloropropionate induced dehalogenase activity. These compounds with the exception of 3CP were all substrates, as was MCA. The 2-haloacid dehalogenases expressed by this organism were isolated and characterised. Initially two dehalogenases were identified (Allison *et al.*, 1983) and a third was subsequently purified (Leigh *et al.*, 1988). The dehalogenases were termed Dehalogenase I, II and III. Dehalogenase I was an L-isomer specific dehalogenase that caused an inversion of configuration to form D-lactate. The enzyme also acted on DCA but not MCA. It had a pH optimum of pH8.4 and showed some sensitivity to sulphydryl-blocking reagents. Dehalogenase II was a non-stereospecific dehalogenase that acted on 2CP, DCP, MCA and DCA. The dehalogenase caused an inversion of configuration when chiral substrates were used. The pH optimum of this dehalogenase is unusual in that there was no discernible optimum over the pH range pH6.1-10.5. This dehalogenase showed the most sensitivity to thiol-binding reagents. Dehalogenase III was a D-2CP specific dehalogenase that caused the inversion of configuration of the substrate. The dehalogenase also acted on MCA. The pH optimum was shown to be

quite broad, at pH 9.1-10.5 and this dehalogenase showed the least susceptibility to sulphydryl-blocking reagents. This was the first D-2CP specific dehalogenase to be reported.

A 2-haloacid dehalogenase from *Pseudomonas cepacia* was characterised in some detail by Tsang *et al.* (1988). The isolate from which the dehalogenase was purified had been obtained on the basis of growth on monobromoacetate, and the dehalogenase characterised was most active against this compound. After purification some of the characteristics of the dehalogenases were determined. The dehalogenase had a native molecular weight of 45 kDa as determined by gel-filtration and cross-linking and a subunit molecular weight of 23 kDa, indicating a dimeric conformation. The dehalogenase was active against MCA and DCA, as well as mono-halopropionates, although only the L-isomer of 2CP was dehalogenated, however, the orientation of the product was not determined. The pH optimum was found in the usual alkaline range, with the optimum being at pH9.4. There was little effect of sulphydryl-blocking agents on the dehalogenase activity. This enzyme can be placed in the B1 class of dehalogenases, although the mechanism of reaction and configuration of the product has not been confirmed. The gene encoding this dehalogenase has been cloned and the cloned protein subjected to protein engineering by random and site-directed mutagenesis (Asmara *et al.*, 1993). The results obtained implicated two residues as key to the catalytic activity of the dehalogenase, His-20 and Arg-42. It was also shown that Arg-18 may have a role in the functioning of the enzyme, possibly in the positioning of His-20. These residues have been located to highly conserved regions amongst L-isomer specific dehalogenases (Section 1.6.2 and Figure 5.7).

As described earlier, a *Xanthobacter autotrophicus* was isolated that contained both a haloacid and haloalkane dehalogenase. The haloalkane dehalogenase was also purified and characterised (Keuning *et al.*, 1985). The dehalogenase was active against chloro-, bromo- and iodo-*n*-alkanes with chain lengths of less than four carbons. There was no activity against haloacetates. The dehalogenase was a monomer of 36 kDa with a pH optimum of pH8.2, somewhat lower than the pH optimum for haloalkanoate dehalogenases. The enzyme was very sensitive to sulphydryl-blocking reagents, implying a catalytic group such as a cysteine. The dehalogenase was crystallised and the structure determined (Franken *et al.*, 1991). The dehalogenase was determined to be a spherical molecule composed of two domains, the first with an α/ß type structure with a central region of ß-sheet and the second domain is arranged on top of the first and is formed of α–helices connected by loops. The active site is a hydrophobic cavity located between the two domains, with Asp-124 being the residue essential for catalysis, rather than a cysteine. The inhibition seen by thiol reagents was thought to be due to possible conformational changes that disturb or destroy the substrate binding site. The structure resembles that of dienelactone hydrolase even though there is no significant sequence homology (Franken *et al.*, 1991).

The organism *Hyphomicrobium* sp. strain DM2 expresses an inducible dichloromethane dehalogenase which is glutathione-dependant (Kohler-Staub and Leisinger, 1985). This dehalogenase was purified and characterised. The pure enzyme had a native molecular weight of 195 kDa from gel-filtration and a subunit molecular weight of 33 kDa, so the dehalogenase has a hexameric structure. This result was confirmed by cross-linking experiments. The dehalogenase had a pH optimum of pH 8.5, similar to that of the haloalkane dehalogenase from *Xanthobacter*

autotrophicus (Keuning *et al.*, 1985) and was specific for dihalomethanes and was inhibited by haloalkanes.

1.6.2 Genetic characterisation of dehalogenases

In recent years a number of dehalogenase genes have been cloned, expressed and sequenced. This has allowed a much more detailed understanding of dehalogenases in terms of their relatedness and also has allowed experiments such as the protein engineering previously described to take place.

Some of the first work concerning the genetics of the dehalogenases was carried out by Hardman et *al.* (1986). They were concerned with the location of dehalogenase genes, whether the genes were chromosomal or located on plasmid DNA. Six isolates, four *Pseudomonas* sp. and two *Alcaligenes* sp., were obtained from soil with the ability to grow on haloalkanoic acids. They were examined for plasmids and shown to contain one of five plasmids, with molecular weights ranging from 98,000 to 190,000 Da. When the organisms were cultured on non-selective medium the ability to utilise HAA's was lost. Once the plasmid had been lost, the ability to utilise HAA's was irretrievable, indicating the presence of an unstable catabolic plasmid. Resistance to one or more heavy metals could also be correlated to the presence or absence of the plasmids. Transfer of the plasmids to other hosts could not be demonstrated although the reason for this was not clear. However, a dehalogenase gene associated with one of the plasmids was readily transferable to the plasmid R68-45, possibly as part of a transposable element.

A report of plasmid-determined dehalogenation had been made previous to this work by Kawasaki *et al.* (1981a). A strain of *Moraxella* was isolated that was able to grow on monofluroacetates and expressed two

dehalogenases, DehH1, active against MFA and MCA and DehH2, active against MCA but not MFA. A plasmid was seen to be present in the strain and upon curing the strain of the plasmid two types of cured strain were seen. The first (type I) had lost DehH2 and the second (type II) had lost bothDehH1 and DehH2. The type I strain still had a plasmid present but the new plasmid was smaller than the original and the type II strain had no plasmid DNA. The plasmid from the wild-type could be transferred to the type II strain with complete reversion to the wild-type phenotype. The interpretation of these results was that both of the dehalogenase genes being encoded on the same plasmid. The plasmid was isolated and termed pUO1. This plasmid would undergo spontaneous deletion to give a plasmid termed pUO11 with a phenotype consistent with the type I cured strain, i.e. no DehH2 dehalogenase activity (Kawasaki *et al.*, 1981b). The sizes of pUO1 and pUO11 were determined to be approximately 44 MDa and 40MDa respectively by electron microscopy. Restriction enzyme digests gave patterns that were analogous, showing that pUO11 had been derived from pUO1. The size of the deletion was deduced to be approximately 4 MDa (6.0 kbp) and this fragment was responsible for encoding the *dehH2* gene. The two dehalogenase genes were cloned separately and their nucleotide sequences determined (Kawasaki *et al.*, (1992). The genes were known to be located close to each other from Southern hybridisation experiments, (Kawasaki and Tonomura 1989) and also the fragment containing the *dehH1* gene hybridised to the fragment containing the *dehH2* gene, so the expectation was that the genes would be significantly similar. The open-reading frames of the genes were assigned and the genes' sequences and predicted amino acid sequences were aligned but no common sequences of significant length were seen, with the identity at the amino acid level was 19% (Leisinger and Bader 1993). The previously noted hybridisation of the two fragments was due to flanking regions of high

similarity. Comparison of the *dehH2* gene to the dehalogenase gene from *P.putida* No.109 was also carried out. This dehalogenase showed some similarities in substrate specificity to DehH2 and both enzymes were sensitive to sulphydryl-blocking agents. The comparison revealed that the two genes were 56% identical at the nucleotide level and the predicted amino-acid sequences were 51% identical and this was thought to indicate a common ancestor. Comparison of the sequence of DehH1 to those in the database revealed that there was significant similarity in the N-terminal region between this enzyme, the haloalkane dehalogenase DhlA from *X.autotrophicus* and three hydrolases from *P.putida*, 2-hydroxymuconic semialdehyde hydrolase, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase and tropinesterase, indicating a possible common ancestry (Kawasaki *et al.*, 1992).

The presence of a plasmid encoded dehalogenase in the organism *Alcaligenes xylosoxidans* has also been reported (Brokamp and Schmidt 1991). In experiments using a soil microcosm innoculated with the organism, the researchers were able to show that a 60 kbp plasmid harbouring the gene for a DCP dehalogenase could be transferred to *Pseudomonas fluorescens*. Subsequent cloning of this gene isolated it to a 6.9 kbp region of the plasmid. Horizontal gene transfer of the plasmid to indigenous bacteria in the microcosm was also demonstrated, indicating a possible reason for the widespread nature of dehalogenase genes in the environment.

The investigation into the dehalogenases of *P.putida* PP3 also examined the enzymes from a genetic aspect. As has been said, mutants that varied in their dehalogenase activity were easy to obtain and seemed to be due to a deletion of DNA rather than point mutations (Weightman *et al.*, 1986). The mutant phenotypes were a result of the loss of a combination of

dehalogenase and or haloalkanoate permease. *P.putida* PP3 was shown to contain two dehalogenases and a permease associated with each enzyme, the loss of permease activity resulting in resistance to MCA and DCA because the rate of uptake was such that there was no build-up of toxic concentrations of the compounds in the cell. The model proposed to account for the mutation frequency, influence of environmental conditions and loss of more than one function in four out of five of the classes of mutants involved the genes being located on transposable elements (Slater *et al.*, 1986). Three transposons were proposed; Tn-*dehA* carrying dehalogenase I (DehI) and permease I, Tn-*dehB* carrying dehalogenase II and Tn-*dehC* carrying permease II. Loss of one or a combination of these transposons could produce all of the mutant strains obtained. To test this model, transfer of the transposons was attempted. The only dehalogenase activity detected in transconjugants was dehalogenase I, a result consistent with the transfer of Tn-*dehA*. Further investigations into the properties of Tn-*dehA* (*DEH*) were undertaken. The transposon *DEH* demonstrated the ability to recombine with a range of target DNAs at high frequencies (Thomas *et al.*, 1992a). Expression of dehalogenase I was fully regulated when *DEH* was transposed onto plasmids that were mobilised into *P.putida*, *P.aeruginosa*, and *Escherichia coli*. When the transposon was inserted into unstable derivatives of the plasmid RP4 it was shown that the transposon would mobilise to insert into the chromosomes of *recA*$^-$ strains. Investigation into the nature of the insertion of *DEH* into the plasmids RP4 and pWWO showed that in the case of RP4 there were eight sites of integration and the size of the *DEH* insert varied from 6-13 kbp. There were only two sites of integration seen for pWWO and out of the twelve plasmids studied, eleven of them had an insertion at the same site, with the size of all of the inserts being 6 kbp. The insertion at the other site was only seen for the one plasmid and the insert was larger, being 9.5 kbp. The

genetic element *DEH* is unusual in that the size of the insert varies and so as such it cannot be described as a transposon in the classical sense, however it does show many of the characteristics of transposons in that it retains endonuclease restriction sites, shows high frequency recombination with a range of target DNAs and shows promiscuous insertion into RP4-based replicons. Characterisation of the structural and regulatory genes carried on *DEH* was undertaken. The region of *DEH* integration in pWWO was cloned and subsequent subcloning and insertion mutagenesis produced a map of the *dehI* structural gene and the $dehR_I$ regulator gene. *dehI* is under positive control of the regulator gene, with subclones of *DEH* lacking the regulator gene only allowing slow growth with 2CP as the carbon source (Thomas *et al.*,1992b). Upstream of the *dehI* gene is a -24/-12 type promoter and the gene is inactive in a *rpoN* mutant showing a dependency for RNA polymerase factor σ^{54}. The $DehR_I$ protein is a constitutively expressed *rpoN*-dependant activator which interacts with the inducers to activate *dehI* expression. The use of suicide plasmids containing *DEH* showed that *dehI* was stably inherited and that the element *DEH* contained all of the functions associated with recombination. A more detailed analysis of the mutations seen when *P.putida* PP3 was grown in the presence of HAA's was also carried out (Thomas *et al.*, 1993). The selection of DCA-resistant mutants was accompanied with DNA rearrangements that caused cryptification of the dehalogenase genes. The genes were retained by the organism but expression ceased. When mutants expressing no dehalogenase activity were grown in the presence of 2CP the ability to utilise 2CP returned via a mechanism of DNA rearrangement and therefore decryptification of the dehalogenase genes (Thomas 1993).

To investigate the genetics of the *Rhizobium* a series of mutant strains were produced by chemical mutagenesis (Leigh *et al* 1986a). Initially, a primary mutant (Type A) was made that lacked the ability to

utilise either DCP or D/L-2CP as a sole carbon or energy source. Enzyme assays, polyacrylamide gel electrophoresis (PAGE) and activity staining of the resultant gels showed that the mutant had lost the ability to produce any of the dehalogenases. Secondary mutants were then selected from this strain by plating them onto a selective media of either DCP or D/L-2CP. This gave rise to three types of mutant strains. When using DCP as the selective agent, two types of secondary mutants arose (Types 1 & 2). The first had regained inducible production of the dehalogenases, i.e. the wild-type phenotype, whilst the other strain constitutively produced Dehalogenase II. Using 2-CP as the selective agent gave rise to a mutant strain that constitutively produced Dehalogenases I and III (Type 3). These mutants are summarised in Table 1.1. From this data, it was possible to propose a model for the regulation of the dehalogenase genes (Figure 1.8). To obtain the Type A mutant strain requires a mutation in the regulator gene. This will cause the loss of production of the dehalogenases, provided all three genes are controlled by this regulator. To obtain the Type 1 mutant, which had the wild-type phenotype, requires either a reversion of the mutation in the regulator gene, or a suppressor mutation in the regulator gene which complements the first mutation and so restores the wild-type phenotype. The Type 2 mutant produced Dehalogenase II constitutively. For this to occur, a mutation in the promoter region of the gene, **P1**, would override the effects of the regulator gene mutation and so production would become constitutive. Lack of a mutation in the **P2** region would mean expression of Dehalogenases I and III would still be under control of the mutated regulator gene and therefore not expressed. The Type 3 mutant gave constitutive expression of Dehalogenases I and III. This is due to a similar reason to that for the Type 2 mutant, where there is a mutation in the promoter region, **P 2**, controlling expression of Dehalogenases I and III.

Table 1.1

Summary of *Rhizobium* dehalogenases and mutant strains, indicating which dehalogenases were expressed (Leigh *et al.*, 1986).

	Dehalogenase I	**Dehalogenase II**	**Dehalogenase III**
pH optimum	8.4	6.1 - 10.5	9.1 - 10.5
Substrate Specificity	L-2CP, DCA	2,2-DCP, DCA, MCA, L/D-2CP	D-2CP, MCA
Mutant Strains			
Type A	Absent	Absent	Absent
Type 1	Present	Present	Present
Type 2	Absent	Present	Absent
Type 3	Present	Absent	Present

Figure 1.8

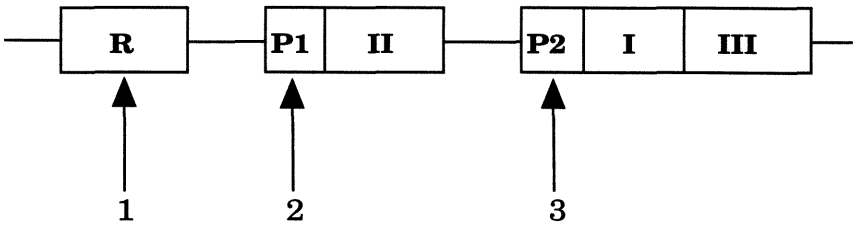


Figure 1.8

Proposed genetic structure and regulation for the *Rhizobium* sp. dehalogenase genes. **R**-regulator gene, **P1**, **P2**-promoter regions, **I**, **II**, **III**-structural genes for Dehalogenases I, II and III, respectively. Arrows indicate mutations causing; (1) No dehalogenases (Type A) and reversion to wild-type phenotype (Type 1) (2) Constitutive dehalogenase II only (Type 2) and (3) Constitutive dehalogenases I and III only (Type 3) (Leigh 1986). See Table 1.1 for mutant types.

Table 1.2

Summary table of data for published haloacid dehalogenases. Mech group refers to the classification according to Leisinger and Bader (1993). Sub/Prod config denotes the optical configuration of the product relative to substrate. S-H sensitivity refers to the effect of sulphydryl-blocking reagents on the activity of the dehalogenase. ND denotes not determined.

Table 1.2

Organism	Enzyme	Mech. group	Substrates	Sub/Prod config	pH optimum	S-H sensitivity
P.dehalogens NCIMB 9061	Halidohydrolase I	B1	MCA, DCA, L-2CP	ND	9.4	No
Pseudomonas sp.	Halidohydrolase I	B1	MCA, L-2CP	Inv	9.3	No
	Halidohydrolase II	B1	MCA, DCA, L-2CP	Inv	9.3	No
P.putida PP3	Halidohydrolase Fraction I	B3	MCA, DCA, D/L-2CP, DCP	Ret	9.0	Yes
	Halidohydrolase Fraction II	B2	MCA, DCA, D/L-2CP, DCP	Inv	9.0	No
Pseudomonas sp. 113	Halidohydrolase	B2	MCA, DCA, D/L-2CP, DCP	Inv	9.5	No
P.cepacia MBA4	Halidohydrolase IVa	B1	MCA, DCA, L-2CP	ND	9.4	No
Pseudomonas sp. CBS3	DehCI	B1	MCA, DCA, L-2CP, DCP	Inv	9.5	No
	DehCII	B1	MCA, DCA, L-2CP, DCP	Inv	9.5	ND
X.autotrophicus	Hdlb	B1	MCA, DCA, L-2CP	Inv	9.5	No
P.putida AJ1/23	HadD	B4	MCA, D-2CP	Inv	9.5	No
	HadL	B1	MCA, L-2CP	Inv	ND	ND

1.6.3 Nucletide sequencing of dehalogenase genes

The first reported 2-haloalkanoic acid dehalogenases to be sequenced were the *dehCI* and *dehCII* genes from *Pseudomonas* sp. CBS3. One of these dehalogenases DehCI has been described in section 1.6.1. The cloning of the genes has not been described but the two genes were cloned and subcloned to give two recombinant plasmids each able to express one of the dehalogenase genes. Expression of these genes was constitutive, implying that the genes were under negative control in the original organism. The structural gene for *dehCI* was 681 nucleotides long, encoding a 227 amino acid protein with a molecular weight of 25,401 Da, similar to the measured molecular weight of 28 kDa for the purified protein. There was a Shine-Dalgarno region close to the GTG start codon and possible -10/-35 sequences were detected. The *dehCII* structural gene was 687 nucleotides long and coded for a 229 amino acid protein with a molecular weight of 25,683, which again corresponded well to the measured molecular weight of approximately 28 kDa (Schneider *et al.*, 1991). There was a possible Shine-Dalgarno sequence upstream of the ATG start codon and putative -10/-35 regions were seen. Comparison of the nucleotide and deduced amino acid sequences of the two dehalogenases showed 45% identity at the nucleotide level and 37.5% identity at the amino acid level. There were several regions of highly conserved amino acids (See Figure 5.7). There was no similarity between these genes and the dichloromethane dehalogenase gene (La Roche and Leisinger, 1990) or the haloalkane dehalogenase gene (Janssen *et al.*, 1989).

The dehalogenases from *P.putida* AJ1 have also been cloned and sequenced (Barth *et al.*, 1992 and Jones *et al.*, 1992). A genomic library of AJ1 DNA was made in *E.coli* and then screened with an oligonucleotide probe based on the N-terminal sequence of the HadD protein. Eight positive clones were isolated, all of which were found to have both HadD and HadL activity. The

DNA was analysed and a series of subclones made until the two genes had been separated so that they were individually expressed. The expressed cloned dehalogenases were electrophoresed alongside purified protein from AJ1 and the proteins co-migrated. The sequence of the *hadD* gene is 903 nucleotides long and encodes a 300 amino acid protein with a molecular weight of 33,601 Da, close to the measured value of 31,800 Da for HadD. There is a Shine-Dalgarno sequence upstream of the *hadD* start but no -10/-35 sequences. The start of the *hadL* gene is 23 bp downstream of the *hadD* stop codon. The gene is 681 nucleotides in length and codes for a protein of 227 amino acids. The predicted molecular weight of this protein, 25,686 Da, is similar to that measured for the HadL dehalogenase. Again there is a ribosome binding site upstream of the start codon but no recognisable promoter sequences. Results suggest that the *hadD* and *hadL* are co-transcribed from a promoter upstream of *hadD*. Comparison of *hadD* with *hadL* at the nucleotide and amino acid level show very low levels of similarity, leading to the conclusion that the two dehalogenases are not related, even though they are located next to each other on the chromosome and are co-transcribed. However, comparison of the HadL protein to the DehCI and the DehCII dehalogenases shows three regions of very high similarity, those already identified from comparison of the DehCI and the DehCII dehalogenases (See Figure 5.7). It may be that these regions are fundamental to the activity of L-specific 2-haloacid dehalogenases.

As mentioned earlier, the gene for the HdlIVa dehalogenase from *P.cepacia* MBA4has been cloned and sequenced (Murdiyatmo *et al.*, 1992). This gene is 696 nucleotides long and is preceded by a good ribosome binding site. The molecular weight of the predicted protein is greater than the value measured for the HdlIVa dehalogenase but this may be due to anomalous migration of the protein in the gel. The predicted amino acid sequence showed the same regions of highly conserved amino acids seen for

the other L-specific dehalogenases, strengthening the argument that these regions are important for the dehalogenase activity.

The only other haloacid dehalogenase to be sequenced is from *X.autotrophicus* (Van Der Ploeg *et al.*, 1991). The *dhlB* gene encodes a 27,433 Da protein which is a similar molecular weight to that seen for the purified DhlB dehalogenase. There were no upstream sequences that showed similarity to the consensus -10/-35 sequences. The predicted amino acid sequence was very similar to the DehCI and the DehCII sequences, 61% and 60.5% identity, respectively, with the conserved regions already described being present. There was no similarity to the haloalkane dehalogenase gene (Van Der Ploeg *et al.*, 1991)

Other dehalogenase genes that have been cloned and sequenced include a dichloromethane dehalogenase (La Roche and Leisinger, 1990) and a haloalkane dehalogenase gene (Janssen *et al.*, 1989). The dichloromethane dehalogenase gene, *dcmA*, codes for a 287 amino acid protein with a predicted molecular weight of 37,340, in agreement with the measured value for the purified protein. Alignment of the *dcmA* gene with members of the glutathione S-transferase supergene family showed three regions of highly conserved amino acids, indicating that this gene is a member of the superfamily. The *dcmA* regulatory gene has been identified. The *dcmR* gene encodes a 30 kDa protein with a helix-turn-helix motif near the N-terminus. Investigations into the regulation of the *dcmA* and *dcmR* showed that expression was negatively controlled at the transcriptional level by the DcmR protein (La Roche and Leisinger, 1991). The haloalkane dehalogenase gene *dhlA* was found to code for a 310 amino acid protein with a molecular weight close to that seen for the purified protein (Janssen *et al.*, 1989). There were good ribosome binding sequences and promoter sequences upstream of the start codon. Comparison of this gene to a

nucleotide database revealed a short region that showed significant similarity to human and rabbit epoxide hydrolases. The gene was found to be associated with a 200 kbp plasmid, pXAU1, (Tardif *et al.*, 1991). The haloacid dehalogenase gene *dhlB* was not associated with this plasmid.

1.7 Aims of the present study

The aim of this study was to obtain clones of the *Rhizobium* sp. dehalogenase genes and to study their organisation in order to determine whether the model proposed by Leigh (Leigh 1986) is correct. Additionally nucleotide sequencing of the genes will be used to examine fine structural details to see if this may help explain the stereospecificity of Dehalogenases I and III and also help determine the evolutionary relatedness of the dehalogenases.

Chapter 2

Materials and Methods

2. Materials and Methods

2.1 Bacterial Strains and Vectors

Bacterial strains and plasmids used in this study are given in Tables 2.1a and 2.1b, respectively.

2.2 Growth media and conditions

Bacterial cells were grown in either complex or minimal media. Complex media (Luria broth (LB)) was as described by Miller (1972) and minimal media was as described by Hareland *et al* (1975). This medium contained $K_2HPO_4.3H_2O$ (4.25g), $NaH_2PO_4.2H_2O$ (1.0g), NH_4Cl (2.0g) and salts solution (10ml) per litre of distilled water. The salts solution was a 10x concentrate that contained NTA (1.0g), $MgSO_4$ (2.0g), $FeSO_4.7H_2O$ (100 mg), $MnSO_4.4H_2O$ (30 mg), $ZnSO_4.7H_2O$ (30mg) and $CoCl_2$ (10mg) per litre of distilled water. Cells prepared for measurement of enzyme activities were grown at 30°C whereas for DNA preparations they were grown at 37°C. Liquid cultures were incubated in a Gallenkamp orbital shaker at 200 r.p.m. Where appropriate, liquid media was solidified by the addition of 1% (w/v) Bacto-agar (Difco). Carbon sources (Sigma) and thiamine were sterilised separately and added aseptically to the media to final concentrations of 10mM for both halogenated and non-halogenated carbon sources and 10μg/ml for thiamine.

Liquid minimal cultures were also supplemented with LB to a final concentration of 0.025% (w/v) for growth of *E.coli* or yeast extract to a final concentration of 0.05% (w/v) for growth of *Rhizobium*.

Minimal medium cultures were routinely grown by inoculating flasks with 0.02 volumes of an overnight culture grown on LB. Luria broth

Table 2.1a Bacterial strains used in the course of this study

STRAIN	GENOTYPE	SOURCE
Rhizobium sp.	*hadD*$^+$, *hadL*$^+$, *hadX*$^+$	Berry (1979)
Escherichia coli		
NM522	*supE*, *thi*, *hsd5*, (*lac-proAB*), [F' *proAB lacI*^{q} ΔZM15]	Gough and Murray (1983)

Table 2.1b Plasmids used in the course of this study

PLASMIDS	RELEVANT FEATURES	REFERENCE
pUC18/pUC19	Amp^R, *lacZ'*	Yanisch-Perron *et al* (1985)
pSC1	*hadD*$^+$, *hadL*$^+$	This Study
pSC530	*hadX*$^+$	This Study
pSC2	*hadD*$^+$, *hadL*$^+$	This Study
pSC3	*hadD*$^+$	This Study
pSC4	*hadL*$^+$	This Study

cultures were inoculated from individual colonies/patches/plaques using a sterile loop or sterile toothpicks. Ampicillin was added as required to media to final concentration of 100μg/ml. Blue/white colony selection was used to detect the presence of inserts in the multiple cloning sites of pUC vectors. This monitors the inactivation of the α- fragment of the ß-galactosidase gene. Isopropyl-ß-D-thiogalacto-pyranoside (IPTG; 0.3mM) was incorporated into the media. The agar plates were also supplemented with 50μl of 3% (w/v) 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal;) in dimethylformamide, spread onto the surface of the plate immediately before the addition of the bacterial culture. Presence of white colonies indicates a disruption of the multiple cloning site.

2.3 Preparation of cell-free extracts

Cell-free extracts were prepared from cells in the mid-to-late exponential phase of growth. Bacteria from liquid culture were harvested by centrifugation at 10,000g for 10 minutes at 4°C. The cell pellets were washed in 0.1M Tris.Acetate buffer pH 7.6 containing 1mM EDTA and 10%(w/v) glycerol and routinely resuspended in 0.04 volumes of the same buffer. The cell suspension, maintained at 0°C, was then disrupted by ultrasonication in an MSE Soniprep 150W ultrasonic disintegrator at a peak amplitude of 8μm. Sonication of *Rhizobium* cell suspensions was generally carried out for 3x30 second periods, with 30 seconds between each sonication and for *E.coli* cell suspensions for 1x45 seconds. Unbroken cells and cell wall material were removed by centrifugation at 20,000g for 15 minutes at 4°C. Where the removal of membranous material was necessary extracts were ultracentrifuged at 120,000g for 90 minutes at 4°C.

2.4 Estimation of protein concentrations

The protein concentration of crude cell extracts was measured using the biuret method (Gornall *et al*, 1949). Protein concentrations of extracts prepared from cells grown in minimal media were usually in the range 1-3mg/ml. A standard curve of protein concentration was constructed using bovine serum albumin as reference. The concentration of protein in column fractions was estimated through measurement of the ultra-violet absorbance at 260/280nm (Layne 1959).

2.5 Enzyme Assays

Enzyme assays were conducted either continuously by the measurement of the organic product of dehalogenation or discontinuously by the measurement of released Cl^-. All assays were carried out at 30°C in disposable plastic cuvettes with a path length of 1cm and a total volume of 1 ml unless stated otherwise. For the discontinuous assay, halogen-free substrates had to be used.

2.5.1 Production of halide-free substrate

A 1M solution of the halogenated substrate (Sigma) was acidified by the addition of concentrated nitric acid until the pH was approximately pH1.0. The substrate was then extracted three times with ethyl acetate and the organic phases were pooled and dried over anhydrous sodium sulphate. The resulting solution was evaporated at 45°C in a rotary evaporator. The residue was then redissolved in distilled water and neutralised with 10M NaOH before being made up to the original volume. This preparation could then be used for enzyme assays involving Cl^- release or for activity staining electrophoresis gels.

2.5.2 Discontinuous assay

Measurement of free halide released during the dehalogenation reaction was carried out by an adaptation of the method of Bergman and Sanik (1957). The reaction mixture consisted of 10 mM halogenated substrate in 0.1M Tris.Acetate buffer pH 7.6 and the reaction was started by the addition of cell-free extract to a final volume of 1 ml. This reaction mixture was incubated at 30°C and aliquots taken at various time points. Once an aliquot was taken it was added to 100 µl of 0.25M ammonium ferrous sulphate in 9M nitric acid and mixed thoroughly. To this was added 100 µl of mercuric thiocyanate-saturated ethanol and the solution was vortexed. The colour was allowed to develop for 10 mins and measured at 460nm in a Pye-Unicam SP500 Series 2 spectrophotometer. Halide concentration was determined by comparison of the absorbance of the test against a standard curve of known concentrations of halide in the range of 0-1µmol ml^{-1} . The standard curve became non-linear beyond a concentration of 0.5µmol ml^{-1} and so a standard curve was also constructed for 0-0.2µmol ml^{-1} and assays were carried out in this range. Standard curves are shown in Figure 2.1a and 2.1b, respectively.

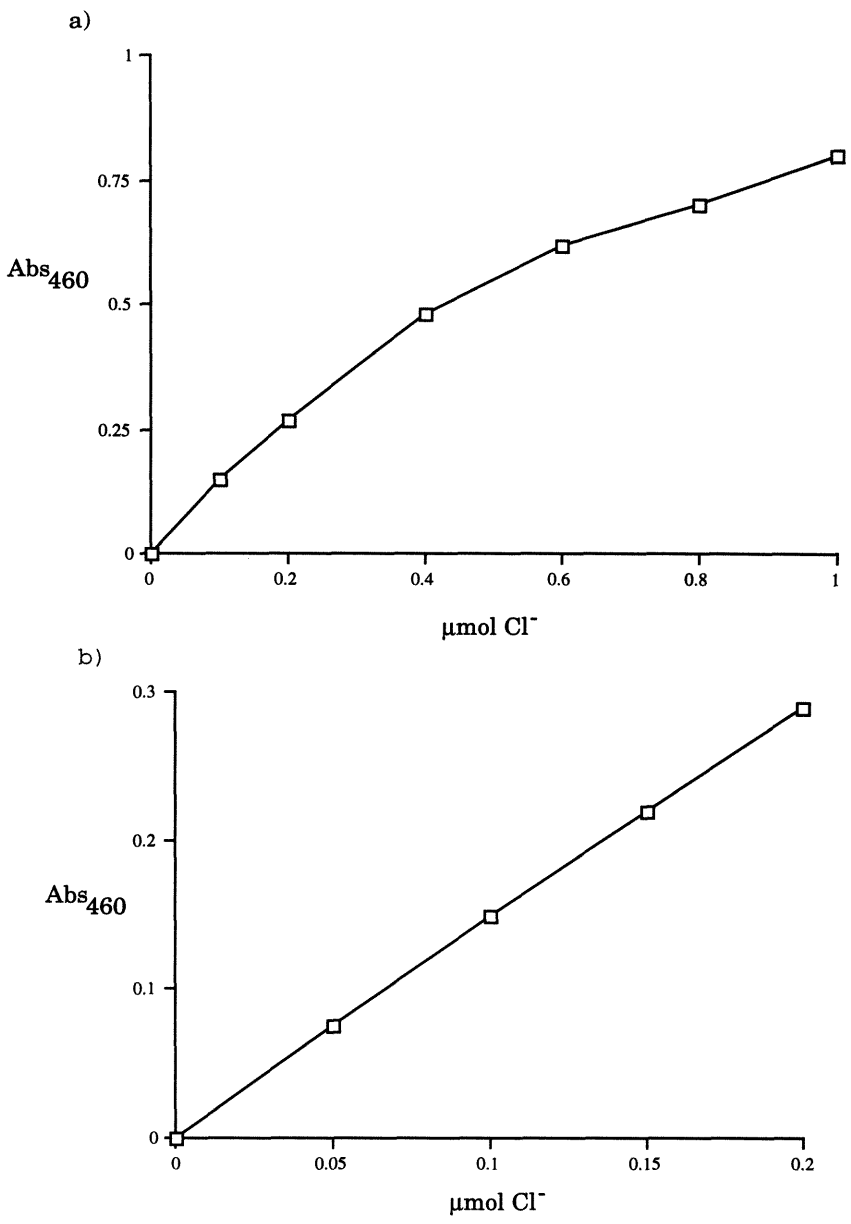
2.5.2 Continuous Assays

The organic product of dehalogenation was determined by continuous assay. This method was used to measure the production of pyruvate, the product of DCP dehalogenation and lactate, the product of 2CP dehalogenation.

Figure 2.1

Standard curve of Cl^- concentration against Abs_{460} for a) 0-1.0 μmol Cl^- and b) 0-0.2 μmol Cl^-. Standards were known concetrations of NaCl in 100mM Tris.Acetate buffer pH7.6, with colour developed as described in Materials and Methods (Section 2.5.2).

Figure 2.1


a)
1
0.75
Abs460
0.5
0.25
0
0
0.2
0.4
0.6
0.8
1
µmol Cl-
b)
0.3
0.2
Abs460
0.1
0
0
0.05
0.1
0.15
0.2
µmol Cl-

2.5.3 Determination of pyruvate

Pyruvate was determined by following the oxidation of NADH associated with the production of pyruvate by the dehalogenation of DCP. A reaction mixture was set up containing the following in 1ml:

100mM sodium phosphate buffer pH 7.6

160μM NADH

5 U Lactate Dehydrogenase

10 mM DCP

The reaction was followed at 340nm in a Pye-Unicam SP1800 recording spectrophotometer. Pyruvate contamination present in the DCP was determined and the reaction was allowed to continue until a steady base line was recorded. At this point, 50μl of ultracentrifuged cell-free extract was added to the reaction. The use of ultracentrifuged material was to eliminate the endogenous rate of the membrane bound NADH oxidase but the endogenous rate of NADH oxidation found in cell-free extracts was measured when necessary in the absence of DCP. The molar extinction coefficient of NADH was assumed to be 6220 $M^{-1}.cm^{-1}$.

2.5.4 Determination of lactate

Lactate was determined by measuring the reduction of NAD^+ consquent upon the formation of lactate accompanying the dehalogenation of 2CP. A reaction mixture was set up containing the following in 1ml:

400mM hydrazine/ 500mM glycine buffer pH9.0

1.025mM NAD^+

5U D-Lactate Dehydrogenase (LDH) / 19U L-LDH

10 mM 2CP

The use of D- or L-LDH allowed the stereospecific dehalogenase activities of HadD and HadL to be measured. The reaction was started by the addition of 50µl of cell-free extract and the reduction of NAD^+ was measured at 340nm. The molar extinction coefficient of NADH was assumed to be 6220 $M^{-1}.cm^{-1}$.

2.5.5 Enzyme Inhibition

Enzyme inhibition studies were conducted with crude cell-free extracts. Dehalogenase activity was measured after the extract had been pre-incubated for 5 minutes with the inhibitor and compared with the dehalogenase activity in the absence of the inhibitor.

2.6 Preparation of plasmid DNA

Plasmids were prepared using an alkaline-lysis method based on that described by Sambrook *et al* (1989). In general 5ml cultures were used for the screening of plasmids of interest and to establish basic stocks. Where the requirement for plasmid DNA was greater, 50ml cultures were processed.

Cells were harvested from an appropriate volume of an overnight Luria broth-antibiotic culture by centrifugation in a benchtop MSE MicroCentaur Microcentrifuge. The supernatant was carefully removed and the pellet resuspended in 0.015 volumes of ice-cold GTE solution (50mM glucose; 25mM Tris-Cl pH 8.0; 10mM EDTA), the mixture was allowed to stand at room temperature for 5 minutes. Next 0.03 volumes of freshly prepared alkaline lysis solution (1% SDS in 0.2M NaOH) was added and the contents mixed with three quick inversions. The mixture was incubated on ice for 5 minutes and then 0.023 volumes of ice-cold potassium acetate solution (3M with respect to potassium; 5M with respect to acetate) added. The tube was inverted and vortexed for 1 second before being incubated on ice for a further

5 minutes. Protein and precipitated chromosomal DNA were removed by centrifugation for 5 minutes in a microfuge at room temperature or for 20 minutes at 15,000g at 4°C depending on the scale of the preparation. The supernatant was then transferred to a fresh tube and was extracted twice with phenol/chloroform/isoamyl alcohol (IAA) (25:24:1) equilibrated with 1M Tris-Cl pH 7.5, to remove any remaining protein. The aqueous phase was then extracted once with chloroform/IAA (24:1) to remove residual phenol. Nucleic acids were then precipitated by the addition of two volumes of ethanol. For small-scale preparations this involved incubation for two minutes at room temperature followed by centrifugation in a microfuge for 5 minutes. For larger scale preparations incubation was for 20 minutes and the precipitate was collected by centrifugation for 15 minutes at 10,000g. The precipitates were washed with ice-cold 70% ethanol and the drained pellets dried under vacuum for 10-15 minutes. The pellet was then dissolved in TE buffer (10mM Tris-Cl pH 8.0; 1mM EDTA) containing 20mg ml^{-1} Ribonuclease A. The DNA thus obtained was suitable for transformation or restriction endonuclease analysis.

Where purer plasmid DNA was required for nucleotide sequencing the protocol above was modified to give higher quality DNA. The modifications followed were based on those described by Kraft *et al* (1988). Following the precipitation with potassium acetate in the alkaline lysis procedure the supernatant was treated with Ribonuclease A added to a final concentration of 50 mg ml^{-1} and incubated for 30-60 minutes at 37°C prior to phenol extraction. The subsequent precipitation of DNA involved incubation on dry ice for 30 minutes rather than the shorter room temperature precipitation of the standard alkaline lysis procedure. The washed and dried DNA was resuspended in 320 µl of TE buffer and was then re-precipitated to remove residual protein by the addition of 80µl of 4M NaCl and 400µl of 13% polyethylene glycol (PEG) 8000 with incubation on ice for 60 minutes. The

DNA was pelleted by centrifugation at 12,000g for 10 minutes, the pellet was washed with 70% ethanol and then dried before being dissolved in sterile distilled water.

2.7 Preparation of chromosomal DNA

Chromosomal DNA was prepared according to the method of Wilson (1990). Cells were grown overnight with vigorous shaking in 200ml of YT (8g tryptone, 5g yeast extract and 5g NaCl l^{-1}) in a 1 litre flask to ensure good aeration. 1.5ml aliquots of cells were pelleted in a microcentrifuge by spinning at full speed for 30 seconds at room temperature. The supernatant was removed without disturbing the pellet and the pellet was resuspended in 567µl of TE solution.(10mM Tris.Cl pH8.0: 1mM EDTA). To this was added 30µl of a 10% (v/v) sodium dodecyl sulphate solution and 3µl of Proteinase K (20mg/ml). The tubes were then mixed thoroughly by inversion and incubated at 37°C for 1 hr. After incubation, 100µl of 5M NaCl was added and the tubes were mixed well. This was followed by 80µl of CTAB/0.7M NaCl pre-heated to 65°C. The tubes were again mixed and incubated at 65°C for 15 mins. An equal volume of chloroform:isoamyl alcohol (24:1) was then added, the tubes were mixed and centrifuged at full speed in a microcentrifuge for 10 mins. The upper aqueous layer was taken off and mixed with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) and centrifuged at full speed in a microcentrifuge for 10 mins after thorough mixing. The upper layer was again removed and 0.6 volumes of isopropanol were added and the tubes were gently shaken for 15 mins, followed by centrifugation at full speed in a microcentrifuge for 10 mins. The resulting DNA pellet was then washed in 70% ethanol, recentrifuged, the supernatant removed and the pellet was dried in a vacuum desiccator before being dissolved in 10 µl of TE containing 20 mg ml^{-1} Ribonuclease A. The DNA obtained was suitable for restriction digestion and Southern blotting.

2.8 Restriction enzyme digestion

Digestions with restriction endonucleases were carried out according to the instructions of the manufacturer using the buffers provided. Enzymes were purchased from Bethesda Research Laboratories (BRL) or Pharmacia-LKB Ltd. Routine digests of plasmids included 500ng-1μg of DNA in a total volume of 15μl with 5-10 units of enzyme added. Digestion times ranged from 1-3 hours for plasmids to overnight for complete digests of chromosomal DNA.

2.9 Phosphatase treatment of DNA

In order to minimise the number of recircularised symmetrically cut vector molecules in ligation reactions the linearised plasmid DNA was treated with calf intestinal phosphatase (CIP) to remove terminal phosphate groups. One unit of phosphatase was added to the digested DNA. Reactions were incubated at 37°C for 30 minutes in the buffer recommended by the manufacturer. The CIP was removed by agarose gel electrophoresis followed by isolation of the DNA from the agarose. (See Method 2.12)

2.10 Treatment of digested plasmids with S1 nuclease

S1 nuclease was used to remove overhangs from non-compatible restriction enzyme sites so as to allow the blunt ended ligation of the two sites. Following the digestion of the plasmid DNA with the restriction enzymes, the DNA was purified by agarose gel electrophoresis. The required fragment was isolated from the gel and dissolved in 4μl of distilled H_2O. To this was added 6μl of S1 nuclease/buffer mix (Pharmacia) and the reaction was incubated at room temperature for 30 mins. The DNA was then again agarose gel purified before ligation.

2.11 Agarose gel electrophoresis

Restriction enzyme digests were analysed by submarine gel electrophoresis through agarose gels. Gels of 0.5-1.0% agarose prepared with TAE gel buffer (40mM Tris-acetate pH7.6, 1mM EDTA (TAE buffer) plus ethidium bromide at 0.5μg ml^{-1}) were used according to the sizes of fragments expected. Gels were run at constant voltage (5-10 V cm^{-1}) in TAE buffer. Samples were prepared by the addition of 0.15 volumes of type II loading buffer (0.25% bromophenol blue (w/v), 0.25% xylene cyanol (w/v), 30% glycerol (v/v)) (Sambrook *et al* 1989). Migration of sample was monitored by observation of the dye fronts. Gels were usually run until the bromophenol blue front had reached the end of the gel and they were visualised under U.V light. Gels were recorded using a UV image analysis system. Gels were calibrated using a 1kb ladder supplied by BRL

2.12 Isolation of DNA fragments from agarose gels.

Fragments generated from restriction digests were isolated using the Geneclean II system (Bio101 Inc.). The desired DNA fragments were identified by illumination with long-wavelength UV light and corresponding regions of the gel were excised with a sterile scalpel blade and placed into a 1.5ml Eppendorf tube. The weight of the gel slice was determined and 3x the volume of a 6M NaI solution was added. The tube was then placed in a 55°C water bath until the gel slice had melted. The solution was mixed and 5μl of Glassmilk™, a DNA binding silica matrix, was added, the tube was mixed again and incubated on ice for 5 mins with occasional mixing. The tube was then centrifuged in a bench-top microcentrifuge for 15secs and the supernatant was decanted. The pellet was then washed 3x with a NEW WASH™ solution diluted with ethanol and distilled H_20 as per manufacturers instructions, each wash being followed by a 15s spin. After the third spin,the supernatant was removed and the pellet was spun again for 10s to allow any residual solution to be removed and a volume of sterile

dH_2O was added equivalent to the original volume of Glassmilk used and the tube placed in a 60°C water bath for 3 mins. The tube was then spun for 30s and the supernatant was removed and placed into an Eppendorf tube.

2.13 Ligation of DNA.

Vector and insert DNA were routinely mixed in a molar ratio of 1:2 in a total volume of 10-20µl and ligated in a reaction containing 1 unit of T4 DNA ligase (BRL) using the supplied buffer. Reactions were incubated at room temperature for 2-16 hours, blunt end ligations were always incubated overnight in the presence of 5% PEG. Generally, half of the ligation reaction was used for transformations of competent cells.

2.14 Transformation of bacterial cells with plasmid DNA

Competent cells were freshly prepared, as required, by the method of Kushner (1978). The required strain was grown overnight in 5ml LB cultures and then diluted one-hundred fold into fresh LB. Cells were then grown at 37°C to an OD_{680} of approximately 0.4. Aliquots of cells (1.5ml) were harvested by centrifugation for 30 seconds and the drained pellets resuspended in 0.5 ml of sterile MOPS A solution (10mM MOPS pH7.0, 10mM RbCl). The cells were then pelleted with a further 20 second spin and resuspended in 0.5ml of MOPS B (100mM MOPS pH6.5, 10mM RbCl, 50mM $CaCl_2$) and left on ice for 60-90 minutes. Cells were collected by a 10 second spin and resuspended in 150µl of MOPS B with dimethyl sulphoxide included to 0.2% (v/v). DNA was added (10-200ng) and the mixture left on ice for 1 hour. The cells were heat-shocked at 55°C for 30 seconds and chilled on ice for 1 minute before 1ml of pre-warmed LB was added. The cells were then allowed to recover at 37°C for 1 hour before 100µl was spread on a pre-warmed LB-antibiotic plate. Plates were then incubated overnight at 37°C.

2.15 Transfer of DNA to nylon membranes

DNA was transferred from agarose gel to nylon membrane using a capillary transfer method (based on Southern, 1975). The gel was first soaked in 0.25M HCl for 7 minutes at room temperature in order to cause breakage of the DNA through depurination, this treatment improves the efficiency of transfer of large fragments. The DNA was then denatured by incubation of the gel with a solution of 1.5M NaCl; 0.5M NaOH for 25 minutes with gentle agitation. The denaturing solution was then removed and replaced with neutralising solution (3M NaCl; 0.5M Tris-Cl, pH 7.4), the gel was then incubated for a further 30 minutes. The gel was removed to a fresh container and rinsed with distilled water. The gel was then placed on a piece of Whatman 3MM paper so that there were no air bubbles between the paper and the gel. The sheet of paper rested on a glass bridge over a plastic box with its ends projecting into the box such that when the container was filled the paper would act as a wick. A piece of Hybond-N membrane which had been cut to the same size as the gel and wetted by immersion in 3xSSC (1xSSC is 0.15M NaCl: 0.015M sodium citrate) was then carefully placed onto the surface of the gel. Any air bubbles between the membrane and the gel were eliminated by the application of gentle pressure. Several layers of 3MM paper were placed on top of the membrane and a piece of Saranwrap film which had been cut to have an aperture the same size as the gel placed over the stack in order to reduce capillary movement of liquid bypassing the gel. A stack of paper towels approximately 10cm high was placed in contact with the gel/membrane sandwich and on top of this was put a glass plate and a 500g weight. The reservoir was then filled with 20xSSC to immerse the ends of the wick and the transfer allowed to proceed for 16 hours with fresh paper towels being added after one hour. After transfer the blotting apparatus was disassembled and the membrane washed in 3xSSC to remove any adhering agarose. The membrane was then air-dried before being wrapped in

Saranwrap film and exposed for 15 seconds to short wavelength U.V light to fix the DNA to the Hybond.

2.16 Radiolabelling of restriction fragments for use as hybridisation probes

Restriction fragments were radiolabelled using [α-^{32}P] dCTP with an oligolabelling kit from Pharmacia-LKB, using the method of Feinberg and Vogelstein (1983). This method involves denaturation of the DNA to be labelled and then mixing with hexadeoxyribonucleotides of random sequence. These random hexamers anneal to random sites on the DNA and so serve as primers for DNA synthesis by the Klenow Fragment of *E.coli* DNA polymerase I with radiolabelling being achieved by the inclusion of labelled nucleotide.

Restriction fragments to be labelled were recovered from agarose gels and 50ng in a volume of 3µl was denatured by heating at 95°C for 3 min and the DNA chilled immediately on ice. The denatured fragment was then added to a mixture comprising 10µl of Reagent mix (Pharmacia), 5µl (50µCi) of [α^{32}P] dCTP (Amersham) and 31µl of sterile water. One µl of the provided Klenow fragment was added and the reaction was allowed to proceed at 37°C for 3 hours. The effectiveness of labelling was evaluated by paper chromatography of 0.5 µl of the reaction mixture on DE81 ion exchange paper using 0.3M ammonium formate as the solvent. In this system labelled polynucleotide remains at the origin whereas unincorporated radioactive nucleotide migrates with the solvent front. After running, the chromatogram was wrapped in Saranwrap and autoradiographed using Fuji RX X-ray film. The film was developed after 10 minutes and showed that labelling had been highly efficient with essentially full incorporation of the label. The probe was denatured by heating at 95°C for 3 minutes before use.

2.17 Probing of Hybond-N filters

Filters were pre-hybridised and hybridised using a method based on that of Williams *et al.*, (1990). Pre-hybridisation was carried out for 1 hour at 65°C in sealed hybridisation chambers using 20ml of de-gassed pre-hybridisation solution (0.1g Marvel milk, 1.2g PEG8000, 2ml 10% SDS, 2ml 15x SSPE pH7.4 and 15ml distilled H_20. (1x SSPE is 0.15M NaCl, 10mM NaH_2PO_4 and 1mM EDTA pH7.4)). Hybridisation was carried out under similar conditions using pre-hybridisation solution to which the denatured probe had been added. Hybridisation was allowed to proceed overnight at 65°C with gentle shaking. The hybridisation solution was then removed and the filter subjected to three ten minute washes at 65°C with pre-warmed 3x SSC; 0.1% SDS followed by three ten minute washes with 0.5x SSC; 0.1% SDS. The filter was then blotted on 3MM paper to remove the excess wash solution and air-dried. The filter was then covered with Saranwrap before autoradiography using Fuji RX X-ray film. Cassettes were left at -70°C if the signal was not very strong or at room temperature otherwise.

2.18 Nucleotide sequencing

Plasmid DNA was sequenced using the chain-termination method of Sanger *et al* (1977) using a modified T7 DNA polymerase (Tabor and Richardson ,1987). Sequencing kits (Sequenase version 2.0) were purchased from United States Biochemicals and radio-isotope was purchased from Amersham International. The procedure may be divided into discrete stages. 1. Denaturation of double standed DNA template. 2. An oligonucleotide primer is annealed to the template. 3. A labelling reaction is performed using limiting concentrations of dNTPs including radio-labelled dATP. This continues to complete incorporation of label into DNA chains with lengths randomly distributed from several to several hundred nucleotides. 4. A termination reaction is performed at higher dNTP concentrations with a dideoxynucleotide

triphosphate included. DNA synthesis proceeds until all DNA chains are terminated with a dideoxynucleotide. 5. Reactions are stopped and the products separated using high-resolution denaturing polyacrylamide gel electrophoresis.

2.18.1 Denaturation of plasmid DNA

Approximately 5μg of plasmid DNA in a total volume of 20μl was denatured by the addition of 2μl of a freshly prepared solution of 2M NaOH; 2mM EDTA and incubation at room temperature for 10 minutes. The mixture was then neutralised by the sequential addition, on ice, of 8μl of 1M Tris-Cl pH 4.5 and 3μl of 3M sodium acetate. The DNA was precipitated by the addition of 75μl of ethanol (-20°C) and incubation on dry ice for 10 minutes before the pellet was collected by centrifugation for 10 minutes. The drained pellet was washed with 200μl of ice-cold 70% ethanol and then dried under vacuum for 10 minutes before being resuspended in 7μl of sterile water.

2.18.2 Sequencing reactions

Sequencing reactions were performed using 5μg of denatured plasmid DNA. For each reaction the DNA was dissolved in a total of 7μl of sterile water. The DNA was mixed with 2μl of reaction buffer and 1μl of the required primer (5ng.ml^{-1}) was added. The primer was annealed by heating the mixture at 65°C for 2 minutes and then allowing slow cooling to room temperature over 30 minutes. During this time 2.5μl aliquots of A, C, G and T termination mixes were placed in separate tubes which were pre-warmed to 37°C. The annealed DNA solution was collected at the bottom of the tube by brief centrifugation and to this was added 1μl of 0.1M dithiothreitol (DTT), 2μl of diluted labelling mix, 0.7μl of [α-^{35}S] dATP (approximately 3.6μCi) and 2μl of Sequenase version 2.0 which had been diluted 8-fold with ice cold dilution buffer immediately before addition. The contents of the tube were mixed and

incubated for 4 minutes at room temperature. At the end of this period 3.5µl of the labelling reaction mixture was transferred to each of the termination mix tubes and the termination reactions were allowed to proceed at 37°C for a further 5 minutes. Reactions were stopped by the addition of 5µl of Stop solution (95% formamide; 20mM EDTA; 0.05% bromophenol blue; 0.05% xylene cyanol) to each tube. Reactions were then stored at -20°C until required.

2.18.3 Electrophoresis and autoradiography

Sequencing gels were 6% acrylamide-7M urea made up in 0.09M Tris-borate;EDTA (2mM) pH 8.3 (TBE) buffer. Gels were pre-run at 1400-1500V with TBE in the upper and lower reservoirs for 30-45 minutes before samples were loaded. An aluminium plate to act as a heat-sink was attached to the gel plates to minimise distortion of the bands during electrophoresis caused by uneven heat-distribution. In order to increase the amount of sequence data obtained, two or three separate loadings of sample were run with 3µl being applied at each loading. The second sample was applied when the xylene cyanol front had migrated two-thirds of the way down the gel with the third sample (where required) being applied when the second xylene cyanol front had reached a similar position. Gels were then run until the final bromophenol blue dye front had reached the bottom of the gel. Following electrophoresis the gel was immersed in a 10% methanol /10% glacial acetic acid solution for 30 minutes before being transferred to Whatman 3MM paper. Gels were then dried down onto the paper under vacuum at 80°C for 90 minutes. Autoradiography was carried out overnight or longer at room temperature using Fuji RX X-ray film.

2.19 Polymerase chain reaction (PCR) amplification of dehalogenase genes

DNA fragments were amplified from chromosomal DNA prepared as described previously (Section 2.7). Primers were designed using conserved amino acid sequences from published L-2CP specific dehalogenases (Figure 5.7). 50ng template DNA was mixed with 0.5μM primer and 1.5mM dNTP's in the following PCR buffer: 67mM Tris.Cl pH8.8, 16.6mM ammonium sulphate, 6.7mM $MgCl_2$, 10mM 2-mercaptoethanol, 6.7μM EDTA pH8.0 and 170μg/ml Bovine Serum Albumin. The amplifiaction program was as follows: Denature at 94°C for 1.5 minutes, anneal at 30°C for 40 seconds and extend at 72°C for 45 second. This was repeated for 30 cycles. After completion the reaction mixture was electrophoresed on a 3% agarose gel as described in Section 2.11.

2.20 Polyacrylamide gel electrophoresis

Protein samples were routinely electrophoresed on polyacrylamide gels using a Mini-Protean II system from BioRad. Denaturing (SDS) gels were prepared with a 12% acrylamide; 0.1% SDS resolving gel prepared in 375mM Tris-Cl pH 8.8 and were polymerised by the addition of 0.1% ammonium persulphate and 0.05% TEMED. Stacking gels were prepared at 5% acrylamide in 125mM Tris-Cl pH 6.8 but were otherwise identical. Samples were prepared for denaturing gel electrophoresis by heating 10μg of protein for 2 minutes at 95°C with 0.3 volumes of sample buffer. The sample buffer consisted of 2% SDS; 0.1% bromophenol blue; 10% glycerol; 100mM DTT in 50mM Tris-Cl pH 6.8. Gels were generally run for 45 minutes at a constant voltage of 200 V in a buffer which was 25mM Tris, 250mM glycine (pH 8.3) which contained 0.1% SDS. The protein bands were stained with Coomassie blue R250 solution (comprising 0.5% (w/v) Coomassie blue R250 in 45% (v/v) methanol; 10% acetic acid) for 0.5-3 hours. Gels were then destained in a solution of 7.5% (v/v) acetic acid; 5% methanol for 2-16 hours. Gels were calibrated using SDS-7 markers from Sigma, the relative molecular weights of the

markers are as follows: bovine serum albumin (66 kDa); ovalbumin (45 kDa); glyceraldehyde-3-phosphate dehydrogenase (36 kDa); carbonic anhydrase (29 kDa); trypsinogen (24 kDa); trypsin inhibitor (20.1 kDa) and bovine lactoglobulin (14.1 kDa).

2.21 Non-Denaturing Polyacrylamide Gel Electrophoresis

In order for enzyme extracts to remain active when used for activity stains, non-denaturing conditions were required. Gels were prepared based on the method of Hardman (1980). Resolving gels containing 7.5% acrylamide in 325mM Tris.SO_4 pH 8.8 were polymerised by the addition of 0.3% ammonium persulphate and 0.1% TEMED. Stacking gels were formed from 4% acrylamide in 125mM Tris.SO_4 pH 6.8. These were polymerised as for the running gels. Gels were left covered overnight at 4°C before being used to allow the ammonium persulphate to decompose completely. Gels were run using a Mini-Protean II gel system from BioRad in 25mM Tris, 19mM glycine buffer (pH8.3) at a constant voltage of 200V and a temperature of 4°C. Samples were prepared as for cell extracts then an aliquot was mixed with 0.1 volumes of sample buffer (0.1% bromophenol blue; 10% glycerol; 100mM DTT in 50mM Tris.Acetate pH 6.8.). Samples were not heated before being loaded onto the gel. The gel was run until the dye front reached the bottom of the gel, usually 1.5-2 hours. Gels were then stained for dehalogenase activity.

2.22 Dehalogenase Activity Staining.

In order to visualise the dehalogenase activity, gels were stained by a method based on Weightman (1980). The gel was incubated in 50mM Cl^--free halogenated substrate for 30-40 mins at 30°C. Substrate was carefully removed and the gel was then placed in a 0.1M $AgNO_3$ solution and incubated in the dark until bands appeared on the gel due to the precipitation of AgCl. The gel was then washed with distilled water to remove the $AgNO_3$

and fixed by washing in 5% acetic acid for 10 minutes. The gel was again washed with distilled water and stored in the dark and photographed as required.

2.23 Amino-terminal amino acid sequencing

The amino terminii of purified dehalogenase proteins were sequenced by Dr. K. Lilley and Miss E. Cavanagh using an Applied Biosystems 470A gas-phase sequencer. A purified fraction of the enzyme was run on a 12% polyacrylamide-SDS mini-gel and then electroblotted onto a polyvinylidine difluoride (PVDF) membrane using 50mM glycine-50mM Tris pH 10.0 as the transfer buffer. The blot was stained with Coomassie blue R-250 and the protein band excised and loaded into the sequencer without polybrene.

2.24 Fast protein liquid chromatography (FPLC)

A Pharmacia FPLC system was used in protein purification procedures in accordance with the manufacturer's instructions. Purifications were performed at room temperature. The separation techniques employed were anion exchange with the Mono Q HR 5/5 column, hydrophobic interaction with a Phenyl Superose 5/5 column and gel filtration with a Superose 12 column. For molecular weight determination, two Superose 12 columns were used in series for greater resolution.

2.24.1 Anion Exchange Chromatography

Anion exchange chromatography was performed using a linearly increasing gradient of acetate ions. For the FPLC system two buffers are required, one containing no elution salt (Buffer A) and the other containing the final concentration of salt required (Buffer B). For the purification of the dehalogenases, Buffer A contained 20mM Tris.Acetate pH 8.0., 1mM EDTA, 5% Glycerol and Buffer B was Buffer A with the addition of 2M sodium

acetate. These buffers are then mixed via a gradient controller to give a linear gradient of 0-2M acetate. Samples for purification were prepared as for cell extracts and approximately 10mg of protein were applied to the MonoQ column per run. The column was run at a flow rate of 1ml min^{-1}. Fractions were collected and then assayed to determine which contained maximal dehalogenase activity. Active fractions were pooled for use in further purifications or for assaying and stored at 4°C.

2.24.2 Hydrophobic interaction chromatography

Hydrophobic interaction was carried out using a linearly decreasing ammonium sulphate gradient. Two buffer systems were used as for the anion exchange chromatography. Buffer A was 50mM NaH_2PO_4, 1mM EDTA, 2.5% glycerol, 1.7M $(NH_4)_2SO_4$ pH7.5. Buffer B was as Buffer A but without the ammonium sulphate. Buffers were mixed with a gradient mixer to give a linear gradient of 1.7M-0M $(NH_4)_2SO_4$.Samples were generally active MonoQ fractions mixed in a 1:1 ratio with 3.4M $(NH_4)_2SO_4$ to give equivalent salt concentration to the buffer.The column was run at a flow rate of 0.5ml min^{-1}. Fractions were collected and then assayed to determine which contained maximal dehalogenase activity. Active fractions were pooled for use in further purifications or for assaying and stored at 4°C.

2.24.3 Gel Filtration

Gel filtration was carried out using one or two columns depending on the degree of resolution required. Columns were pre-equilibrated overnight using a buffer containing 20 mM Tris.acetate, 0.1M sodium acetate pH 7.6. The sodium acetate is present to stop any interaction with the gel matrix. Generally, fractions from the MonoQ runs were applied to the gel filtration column with approximately 0.5 mg of protein being added per run. The column was run at a flow rate of 0.2 or 0.4ml min^{-1}, with the slower flow rate

being used for molecular weight determination. Protein elution was monitored as before and active fractions were determined by the halide release method. Active fractions were stored at 4°C. For molecular weight determination the columns were calibrated using molecular weight standards from Sigma, the relative molecular weights of these being: ß-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa) and cytochrome c (12.4 kDa).

2.25 Oligonucleotide synthesis

Oligonucleotides were synthesised by Mrs. D. Langton with an Applied Biosystems 380B DNA synthesiser using cyanoethyl phosphoamidite chemistry. The concentration of oligonucleotides was determined from the A_{260} of a diluted oligonucleotide solution using the conversion constant A_{260} of 1.0 is equivalent to a concentration of 33µg/ml.

2.26 Resolution of 2CP

Determination of the resolution of 2CP was carried out by measuring the release of bound Cl^- over a period of time in the presence of extracts or whole cells containing either one or two dehalogenases. A reaction mixture was prepared from 100mM Tris.acetate pH7.6 and 10mM 2CP. The reaction was started by the addition of whole washed cells or cell-free extract containing either one of the stereo-specific dehalogenases or both of the dehalogenases. Cells were grown to late exponential phase on 2CP and in the presence of IPTG. Cells were then centrifuged at 10,000g for 10 minutes at 4°C. The pellet was resuspended in 0.25 volumes of 100mM Tris.acetate pH 7.6 and centrifuged as before. The pellet was then resuspended in 0.1 volumes of 100mM Tris.acetate pH 7.6 and divided into two aliquots. One aliquot was used for the preparation of cell-free extracts as previously described, whilst the other was used as whole cells. Whole cells and cell-free extracts were

stored at 4°C until use. The reaction mixture was shaken in a 30°C water bath. At various time points aliquots were taken and divided, a sample being used for the immediate assay of Cl^- release whilst the remainder was then frozen in dry ice/IMS before being stored at -20°C. Once there was no further Cl^- release seen, an aliqout of cell-free extract or whole cells containing the dehalogenase of opposite stereo-specificity was added and Cl^- release was followed until there was no further Cl^- release was seen. Cl^- concentration was determined as described earlier (Section 2.5.2).

Chapter 3

Isolation and Analysis of Dehalogenase Genes

3.1 Isolation and analysis of dehalogenase genes

In order to isolate the genes encoding the dehalogenase enzymes it was decided to use the direct selection approach, that is to confer a new ability to the host cell, in this case it is the ability for the host to grow on a substrate that previously the host was unable to utilise. For this method of selection several criteria need to be met. First, the host needed to be able to take up into the cell the carbon source used for selection and second the host needs to be able to use the organic product of dehalogenation as a source of carbon and energy. 2CP was chosen as the substrate because it was converted to lactate by all three of the dehalogenase enzymes and so this increased the chances of detecting at least one of the genes, whereas the use of DCP as the substrate would have meant that only one of the dehalogenase genes could be detected. This approach had not been used previously to select dehalogenase genes. Previously, dehalogenase genes had been identified using probes designed from the N-terminal amino acid sequences of dehalogenase enzyme (Barth *et al.*, 1992). The product of 2CP dehalogenation is lactate, whilst the product of DCP dehalogenation is pyruvate, both of which are metabolisable by *E.coli*. In the absence of a cloned 2CP or DCP uptake system it was hoped that the halogenated compounds might be taken up by the lactate or propionate uptake systems present in *E.coli*. It was decided to use *E.coli* NM522 as host for ease of selection of recombinant transformants. Cloning of the dehalogenase genes using a PCR methodology was also attempted.

3.1.1 Construction of a genomic library

To isolate the dehalogenase genes a genomic library was to be constructed in *E.coli* NM522 and the transformants screened for their ability to grow on 2CP. *Rhizobium* genomic DNA was partially digested and size selection of the fragments in the 6-12 kb range was used as fragments of this size were

more likely to contain intact dehalogenase genes. Total genomic DNA was prepared from a culture inoculated with a single *Rhizobium* colony and 10 μg of this DNA was digested with 1U of the restriction enzymes *Eco*RI or *Hin*dIII for 1h at 37°C. The plasmid pUC19 was digested with the same restriction enzymes and then subjected to alkaline phosphatase treatment to minimise recircularization of the vector during subsequent ligations. After the time courses of the digestions were complete the DNA was electrophoresed through a 0.8% agarose gel and genomic DNA in the size region 6-12 kb was cut from the gel and purified using the Gene-clean protocol. The band of digested vector was also removed from the gel and extracted using the same procedure. Ligations were set up in the following vector:insert ratios (w:w); 1:1, 1:2, 1:5 and the ligations were left at 4°C for approximately 85 h. *E.coli* NM522 was transformed using the standard procedures and the recombinants plated onto LB-ampicillin plates and incubated overnight at 37°C. One aliquot of each transformation was plated onto LB-ampicillin plates that also contained IPTG and X-Gal in order to monitor insertion efficiency. After incubation approximately 150 colonies per plate were seen and the insertion efficiency was >75% in all but one ligation. This gave a total of 4000 recombinant colonies. This library was considered comprehensive enough based on the Clarke and Carbon equation (Clarke 1976) for it to be 99% certain of containing the genes of interest and so *E.coli* NM522 containing the gene library was screened for the ability of to grow on 2CP.

3.1.2 Screening for dehalogenase activity

All of the transformant colonies were replica plated onto minimal medium plates containing ampicillin and IPTG with 2CP as the sole source of carbon and energy. The antibiotic was present to act as a selection pressure to maintain the plasmid and also helped to prevent any non-plasmid

containing contaminants from growing on the plates. IPTG was present to act as an inducer for the *lac* promoter present on pUC19. It may be that this was the promoter that will be required for the expression of the cloned dehalogenase genes as opposed to a *Rhizobium* promoter that acts in *E.coli*. Replica plating was carried out with the use of sterile pieces of velvet that had the plate of transformants pressed against them followed by the selection plate. In this way it was convenient to screen large numbers of recombinants.

No colonies were seen after an incubation period of 10 days at 30°C. The library was constructed again and the screening procedure repeated, but there were again no positive clones. Possible reasons for the lack of positive clones are lack of expression or uptake of the substrate. Another reason was that the substrate was toxic towards the *E.coli* host, despite the fact that the *Rhizobium* was able to grow on medium containing the same 2CP. The toxicity of 2CP was investigated.

3.1.3 Toxicity of 2CP

Toxicity tests were carried out by streaking single colonies of *E.coli* NM522 onto minimal medium plates containing either 10mM glycerol with or without 10mM 2CP. *E.coli* NM522 was able to grow on glycerol only in the absence of 2CP, indicating that there was a toxic effect from the 2CP. The purified 2CP that was made for the dehalogenase assay (Section 2.5.1) was then tested for toxicity and was found to have no toxic effect against *E.coli* NM522 and so only purified 2CP was used subsequently for the screening and growth of recombinant colonies. The nature of the toxic compound was not determined but it was presumably an inorganic contaminant that was removed during the extraction. Preliminary experiments were carried out to ensure that *E.coli* NM522 had no dehalogenase activity and was unable to grow on 2CP. Cell-free extracts of *E.coli* NM522 grown on glycerol in the

presence of 2CP were made and tested for dehalogenase activity against 2CP and no activity was found. *E.coli* NM522 plated onto minimal medium plates containing 2CP as the sole carbon source showed weak background growth. This is due to the presence of lactate in the 2CP due to autodegradation.

3.1.4 Screening for dehalogenase activity using purified 2CP

The genomic library was constructed as before and screened in a similar fashion except that the 2CP used for the selection plates had been purified as described in Materials and Methods (Section 2.5.1). The selection plates were incubated at 30°C until colonies were seen to develop. After 96h a colony was seen to have grown strongly above the background growth on one of the plates and after 192h a further colony was seen on another of the plates which was also much stronger than the background growth. The background growth was due to the presence in the plates of lactate which is a product of 2CP auto degradation.

The two colonies were then picked with sterile toothpicks onto various media to test for retention of their characteristics. Both colonies grew overnight at 30°C on LB-ampicillin and grew after 72h on the minimal medium plates containing 2CP and ampicillin, showing that they had retained their initial characteristics. The putative clones were then mini-prepped to prepare some plasmid DNA for analysis and for retransformation of *E.coli* NM522. New transformants showed the same properties as the original library transformants indicating that the plasmid DNA was conferring the new growth ability to *E.coli* NM522. The two plasmids were designated pSC1 and pSC530.

The selection process was also carried out using DCP as the selecting carbon and energy source. Gene libraries were obtained by the same

methods as described previously and replica plated onto minimal medium plates containing extracted DCP, ampicillin and IPTG and incubated at 30°C. However no colonies grew on the selection medium. Construction and screening of a gene library was repeated but again no colonies able to grow on DCP were seen

The next stage of the analysis of the two clones was to characterise the cells carrying the two plasmids in terms of their growth capabilities and dehalogenase activities and for a detailed analysis of the cloned DNA that determined these characteristics.

3.2 Properties of *E.coli* NM522 carrying pSC1 or pSC530

3.2.1 Growth of cells carrying pSC1 or pSC530 on halogenated substrates

As was mentioned at the beginning of this chapter, for growth to occur the factors of uptake and metabolism need to be considered. Both clones isolated on 2CP were tested for growth on a variety of halogenated substrates to help clarify which of the dehalogenase genes had been cloned. Minimal medium plates were made containing either 2CP, D-2CP, L-2CP or DCP as sole carbon and energy source, ampicillin and IPTG. Onto these plates were streaked single colonies of cells carrying pSC1 or pSC530 and the plates were incubated at 30°C. Both of the clones allowed good growth on 2CP, D-2CP and L-2CP, but little or no growth on the plates containing DCP as the sole carbon source. Background growth on the DCP containing plates was due to pyruvate contamination. This indicated that both clones contained genes that coded for dehalogenases able to degrade both the D and L isomers of 2CP, but cells carrying these genes were unable to grow on DCP. This would initially imply that the cloned dehalogenase genes were for Dehalogenases I and III, and Dehalogenase II had not been cloned,

although it could not be ruled out that DCP utilisers had not been isolated because of problems with uptake of DCP into the cell rather than it not being degraded once in the cell.

In order to determine the specificities of the cloned dehalogenases, cell-free extracts were made and the dehalogenase activity towards 2CP and DCP was determined. Cells were grown until late log phase in the presence of 2CP as sole carbon source and with IPTG (0.3mM) present in the medium. The results of these assays can be seen in Table 3.1. The specific activity of cells harbouring pSC1 was approximately 66% that seen in the *Rhizobium* and there was no activity towards DCP. This concurs with the results seen from the growth tests, where there was no growth on DCP. From these results it would suggest that the plasmid pSC1 codes for the Dehalogenases I and III, as these have activity towards D- and L-2CP but are not active towards DCP. The results of the assays with extracts from the cells containing pSC530 show that whilst this extract was active towards 2CP, albeit with approximately 11% of the Rhizobial activity, there was also detectable activity towards DCP. This implies that the DNA present in pSC530 codes for Dehalogenase II. The observation that *E.coli* NM522::pSC530 was unable to grow on DCP as a sole carbon and energy source, even though there was dehalogenase activity within the cell means that there may be a problem with the uptake of DCP into the cell, i.e. the uptake system for DCP had not been cloned along with the dehalogenase gene and the cell was unable to use any of its endogenous uptake systems to mobilise DCP into the cell. From these results it can be seen that the two clones isolated would appear to represent a *Rhizobium* DNA fragment encoding for Dehalogenases I and III and a separate DNA fragment coding for Dehalogenase II.

Table 3.1

Comparison of dehalogenase specific activities in cell-free extracts from *Rhizobium* , *E.coli* NM522::pSC1 and *E.coli* NM522::pSC530. Cells were grown on 10mM 2CP with the addition of IPTG where required. Values are a mean of at least five determinations.

	Specific activity (μmol Cl- released $min^{-1}mg$ $protein^{-1}$)	
Strain	D/L-2CP	DCP
Rhizobium	0.45	0.23
E.coli NM522::pSC1	0.32	ND
E.coli NM522::pSC530	0.05	0.025

ND= Not Detectable.

3.2.3 Growth rates of cells carrying either pSC1 or pSC530

Growth rates of cells carrying the two clones were determined. These rates can be seen in Figure 3.1 and Figure 3.2 and from this it can been seen that the rate of growth of *E.coli* NM522::pSC1 on 2CP was significantly quicker than that of *E.coli* NM522::pSC530 on 2CP, with doubling times being 16h and 93h respectively. The rate of *E.coli* NM522::pSC1 was approximately 2.5-fold slower when compared to the reported rate for the *Rhizobium* of 6h (Allison 1981) and also slower than the growth rate of *E.coli* NM522 on lactate, indicating that it was the process of uptake/dehalogenation that was rate limiting rather than the metabolism of the lactate formed from 2CP. There was no reason for the *E.coli* NM522 to have a specific uptake system for 2CP and uptake of the halogenated substrate into the cell was probably via a propionate or lactate uptake system that was induced by 2CP, although at perhaps lower levels than if induced by the optimal inducer, so the rate of uptake into the cell would probably be decreased due to the bulkier nature of 2CP compared to lactate or propionate. This combined with the lower dehalogenase specific activity seen in cells carrying pSC1 compared to the *Rhizobium* could lead to the slower doubling times seen for cells carrying the plasmid. This would apply also to the *E.coli* NM522::pSC530, which combined with the even lower dehalogenase specific activity in the cell would give the reduced growth rate seen. There was a 6-fold difference in the dehalogenase specific activities of extracts from cells carrying pSC1 and pSC530 and this difference was reflected in their growth rates, with cells carrying pSC530 growing approximately 6-fold slower than those carrying pSC1. This implies that there was an effect of dehalogenase specific activity on growth rate. Experiments which clarify the situation are described later (Section 4.1.2).

Figure 3.1

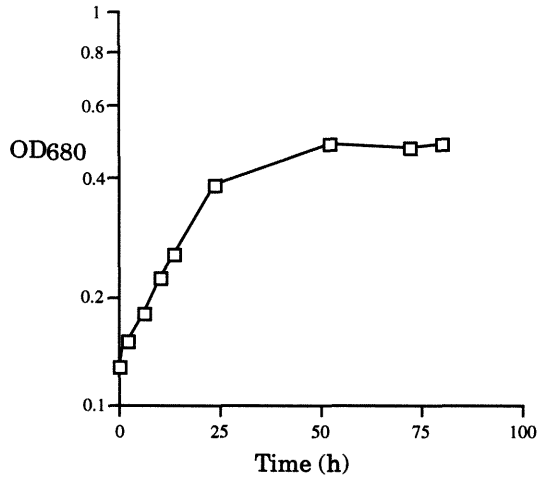


Figure 3.1

Growth of cells containing pSC1 with 10mM 2CP as the sole source of carbon and energy and in the presence of IPTG (0.3mM) plotted logarithmically.

Figure 3.2

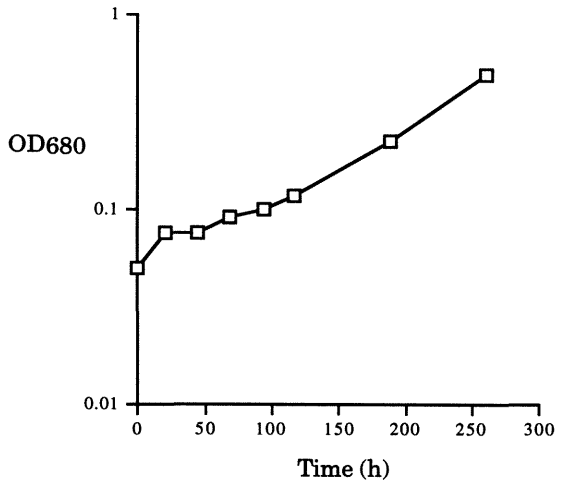


Figure 3.2

Growth of cells containing pSC530 on 10mM 2CP as the sole source of carbon and energy and in the presence of IPTG (0.3mM) plotted logarithmically.

3.2.4 Induction of dehalogenase activity

In order to determine whether the clones contained regulatory genes that were recognised and expressed in *E.coli* as well as the structural genes, it was decided to see if the presence of 2CP was able to induce dehalogenase activity or if the genes were expressed via the *lac* promoter of pUC19 rather than a cloned and expressible regulatory system. Cells containing either pSC1 or pSC530 were grown in the presence of ampicillin with either 2CP or IPTG as an inducer, and 10mM lactate as a carbon source. Table 3.2 shows that the addition of 2CP as an inducer had little effect on the dehalogenase activity of cells containing pSC1, only on the addition of IPTG does one see a significant increase in the levels of activity, and the addition of both compounds shows no higher levels of induction than that with IPTG alone. This implies that the insert carried by pSC1 does not contain regulatory genes that are recognised by *E.coli* RNA polymerase and required for the induction of the dehalogenase genes by 2CP, a known inducer of dehalogenase activity in the *Rhizobium* (Allison 1981) and that the dehalogenase genes are under control of the *lac* promoter of pUC19. This finding could be confirmed by reversing the orientation of the insert, which would abolish dehalogenase expression if the genes were under the control of the *lac*promoter. The results of the induction of cells carrying pSC530 were similar to that of pSC1, with the presence of IPTG being required for the induction of the dehalogenase activity, so again the regulatory genes recognisable in *E.coli* have not been cloned on this fragment.

3.2.5 Activity staining of cell-free extracts

A series of activity stains were carried out to compare the electrophoretic mobilities of the expressed dehalogenase proteins with those of the

Table 3.2

A comparison of the effects of inducers upon the dehalogenase specific activities of cell-free extracts from *E.coli* NM522 containing either pSC1 or pSC530. All cells were grown with 10mM lactate as the carbon source, ampicillin, plus stated inducer. Specific activities given are μmol Cl^- released $min^{-1}mg\ protein^{-1}$. Values are a mean of three determinations.

	Inducer			
Plasmid	None	2CP	IPTG	2CP+IPTG
pSC1	0.06	0.11	0.33	0.32
pSC530	ND	ND	0.05	0.05

ND=Not Detectable

Rhizobium proteins. Native PAGE was carried out and gels were incubated with either 2CP or DCP before being stained with $AgNO_3$. The extracts from pSC1 containing cells showed bands with activity against 2CP which had the same electrophoretic mobility as that of the upper band of the *Rhizobium* extracts which was known to correspond to Dehalogenases I and III. There was no lower band in the region where Dehalogenase II was known to migrate (Allison 1981) and so this was seen as further evidence that pSC1 carried the genes that coded for Dehalogenases I and III. Extracts from cells containing pSC530 gave no positive activity stains. The reason for this was probably that as the level of activity in the extracts was low and there was possibly a further loss of activity during the activity staining procedure, then there was insufficient activity for the bands to be visualised.

3.3 Molecular analysis of pSC1 and pSC530

3.3.1 Analysis of insert DNA in pSC1 and pSC530

Cultures inoculated with single colonies containing either pSC1 or pSC530 were grown overnight and plasmid DNA was prepared and this DNA was then subjected to restriction enzyme analysis. The DNA was digested with the enzyme *Eco*RI, the enzyme used for the cloning and then electrophoresed. The restriction pattern can be seen in Figure 3.3. Both lanes containing plasmid DNA show the common pUC19 band at 2.7kbp, but pSC1 contains a 6.5kbp fragment whereas pSC530 contains a 3.0kbp fragment. This would suggest that the two clones obtained by the selection method are not the same DNA, but are two individual regions of *Rhizobium* DNA both of which allow *E.coli* NM522 to grow on 2CP. The two fragments must represent distinct regions of DNA as both contain no internal *Eco*RI sites and so pSC530 cannot be a smaller subclone of pSC1.

Figure 3.3

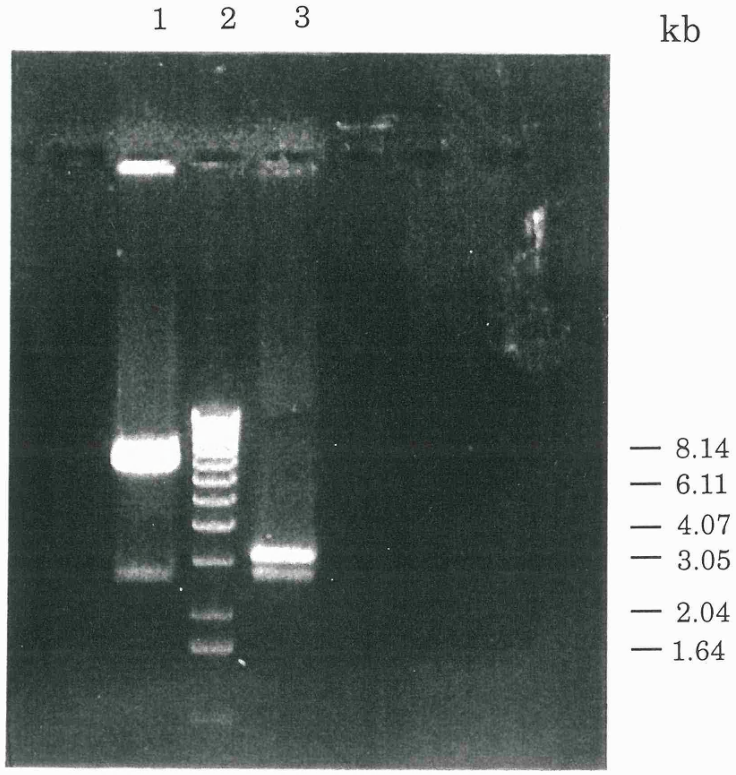


Figure 3.3

Restriction enzyme digest of plasmids pSC1 and pSC530. Plasmids were completely digested with *Eco*RI and electrophoresed on a 0.8% agarose gel. Lane 1: pSC1, Lane 2: 1kb ladder molecular weight markers, Lane 3: pSC530.

3.3.2 Restriction mapping of pSC1 and pSC530

Further characterisation of the two clones was undertaken by mapping of the inserts with a variety of restriction endonucleases, all of which were known to cut the multiple site polylinker of pUC19 only once. Through a series of single and double restriction enzyme digests it would be possible to determine the location of restriction sites within the cloned DNA. The resulting map would provide the locations of restriction sites that could be used for subcloning purposes. The restriction map that resulted from the digests is shown in Figure 3.4. From the map it can seen that the two inserts are quite distinct, they share no common restriction sites and they therefore form two regions of the *Rhizobium* genome that encode dehalogenase enzymes. What cannot be told from the restriction map was whether the two regions are contiguous in the genome or are spatially separate. In order to investigate this question and to determine if there was any homology between the two fragments and therefore the genes which they encode, it was decided to carry out Southern blot analysis.

3.3.3 Southern blot analysis of *Rhizobium* genomic DNA with pSC1 and pSC530

Two probes were prepared for use in the Southern blots by removing the inserts from the plasmids pSC1 and pSC530 and labelling the complete insert DNA. *Rhizobium* genomic DNA (5μg) was digested separately overnight with *Eco*RI and *Hin*dIII and run out on a 0.8% agarose gel before being transferred to a Hybond-N nylon membrane. An *Eco*RI digest of each of the plasmid DNAs was also included on the gel to act as a positive control and to see if there was any cross-hybridisation between the dehalogenase genes. After gel-purification in order to remove any vector DNA the probes were prepared by labelling the fragments with [α-^{32}P]

dCTP using the random priming hexanucleotide method of Feinberg and Vogelstein (1983).

The membranes were allowed to hybridise to the probe overnight at 65°C. The membranes were washed at high stringency and were then autoradiographed at room temperature or at -70°C depending on the strength of the signal. The results of these Southern blots can be seen in Figures 3.5 and 3.6, along with a photograph of the corresponding agarose gel prior to DNA transfer.

From the autoradiograph using the pSC1 probe it can be seen that a single hybridisation band is seen in each genomic DNA track, with the band being at 6.5 kbp in the *Eco*RI track and greater than 10 kbp in the *Hin*dIII track. This was as expected since there are no *Hin*dIII sites within the pSC1 insert in which case a fragment larger than the insert should be seen. There was no hybridisation between the pSC1 probe and pSC530 DNA. The autoradiograph of the membrane probed with pSC530 shows the expected band at 3.0 kbp in the *Eco*RI track and shows two bands in the *Hin*dIII track, one at approximately 3.5 kbp and the other at 4.5 kbp. Two hybridising *Hin*dIII bands were expected due to the internal *Hin*dIII site in pSC530, one of which should be approximately twice the intensity of the other. The reason for the difference in the intensity of the bands was that digestion of pSC530 with *Hin*dIII gives rise to two fragments, of 1kbp and 2kbp. The hybridising fragment containing the 2kbp region of the pSC530 insert will give twice the signal of the hybridising fragment containing the 1kbp region. These autoradiographs show that the two clones do not have sufficient homology to each other to allow hybridisation to occur under the conditions used and also that the two regions of DNA are not contiguous in the *Rhizobium* genome because there was no common *Hin*dIII band seen. If

Figure 3.4

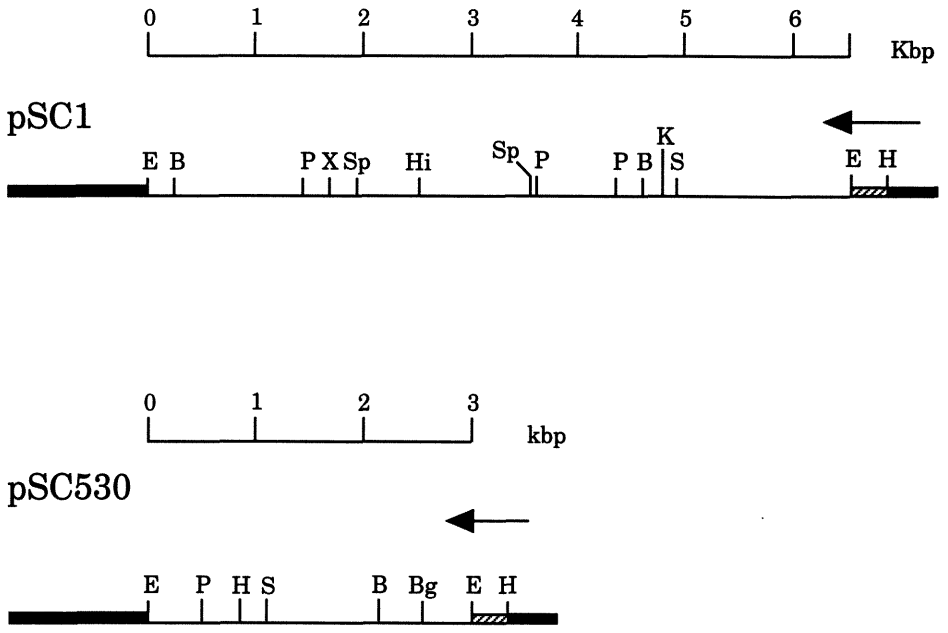


Figure 3.4

Restriction maps of the plasmids pSC1 and pSC530. Key to enzymes:E=*Eco*RI, B=*Bam*HI, H=*Hin*dIII, P=*Pst*I, X=*Xho*I, Sp=*Sph*I, K=*Kpn*I, S=*Sal*I and Hi=*Hin*cII. Bold line denotes the plasmid pUC19, hatched region indicates polylinker (not to scale) and fine line indicates insert. Arrow denotes direction of transcription of the *lac* promoter.

the two regions were contiguous, they would share a common band within *Hin*dIII digested genomic DNA because of the arrangments of *Hin*dIII sites in the fragments. As there was an internal *Hin*dIII site within pSC530 but not within pSC1, a *Hin*dIII digest of the genomic DNA would give rise to a band containing all of the pSC1 DNA and a proportion of the pSC530 DNA if the regions were contiguous, so this band would hybridise to both probes. As this was not seen, the two fragments must be from spatially distinct regions of the genome. It was possible that one or both of the fragments may be located on a plasmid, as had been seen for a number of dehalogenase genes (Kawasaki *et al* 1992 and Hardman *et al* 1986), and due to the nature of the genomic DNA preparation it was possible to precipitate large plasmids along with chromosomal DNA. Attempts to obtain plasmid DNA from the *Rhizobium* were unsuccessful and so it was assumed that the genes are chromosomally located.

3.4 PCR amplification of dehalogenase genes

A possible method of cloning the L-2CP specific dehalogenase gene was to use PCR to amplify an internal region of the gene from chromosomal DNA and use this region as a probe to obtain the full length gene. Comparisons of sequenced L-2CP dehalogenases showed three possible regions from which PCR primers could be designed (Figure 5.7) and primers were designed from two of these regions. The primers were, 5'-3':

SC1:GCGATCCTGTCAACGG			SC2:GGATTAGACGAAACAAA		
A	T	C	CG	T	T
T	A	T		C	C
C		G		G	G

These primers were used for PCR reactions as described in Materials and Methods (Section 2.19). No consistent bands could be amplified from the *Rhizobium* chromosomal DNA and so this method was unable to yield a DNA frgament to be used as a probe for identifying the dehalogenase genes.

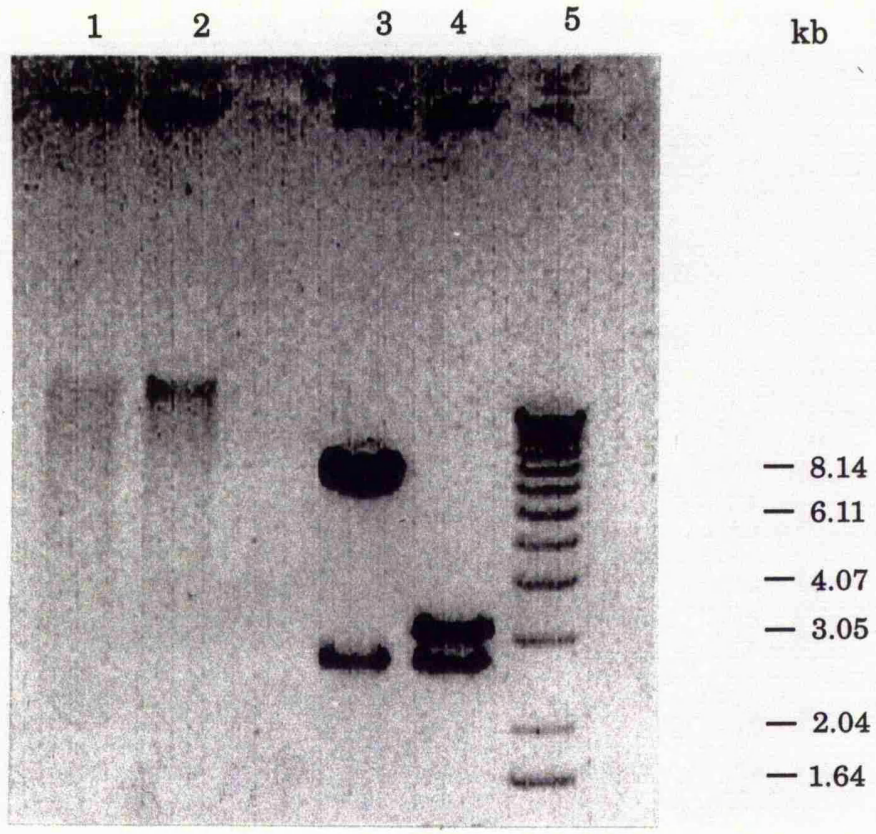
Figure 3.5

Southern blot analysis of *Rhizobium* chromosomal DNA digested with *Eco*RI or *Hin*dIII. 5μg of DNA were digested completely before being run out on a 0.8% agarose gel and transferred to Hybond-N nylon membrane. The membrane was then probed with radiolabelled 6.5 kbp *Eco*RI fragment from pSC1. *Eco*RI digested pSC1 and pSC530 are included as controls. The membrane was treated and the autoradiograph exposed as described in Chapter 2 (Section 2.17). Lane 1: *Eco*RI digested chromosomal DNA, Lane 2: *Hin*dIII digested chromosomal DNA, Lane 3: pSC1, Lane 4: pSC530 and Lane 5 1kb ladder molecular weight markers.

a) Agarose gel of DNA prior to blotting for Southern blot analysis.

b) Autoradiograph of hybridised membrane

a)



b)

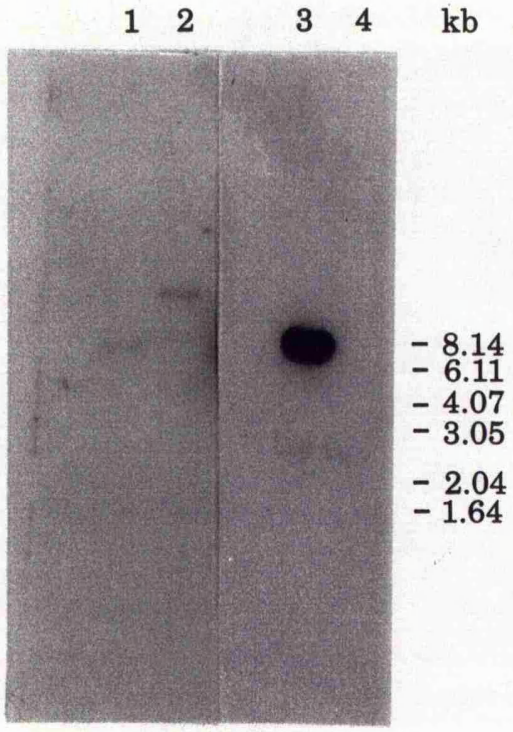


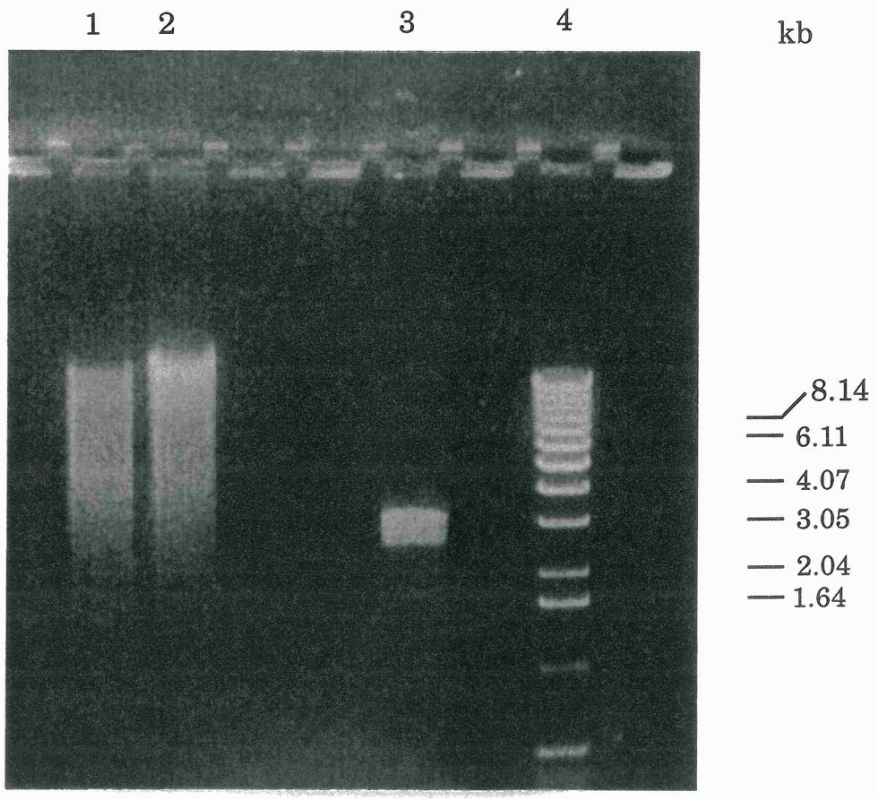
Figure 3.6

Southern blot analysis of *Rhizobium* chromosomal DNA digested with *Eco*RI or *Hin*dIII. 10μg of DNA were digested completely before being run out on a 0.8% agarose gel and transferred to Hybond-N nylon membrane. The membrane was then probed with radiolabelled 3.0 kbp *Eco*RI fragment from pS530. *Eco*RI digested pSC530 is included as a control. The membrane was treated and the autoradiograph exposed as described in Chapter 2 (Section 2.17). Lane 1: *Eco*RI digested chromosomal DNA, Lane 2: *Hin*dIII digested chromosomal DNA, Lane 3: pSC530 and Lane 4: 1kb ladder molecular weight markers.

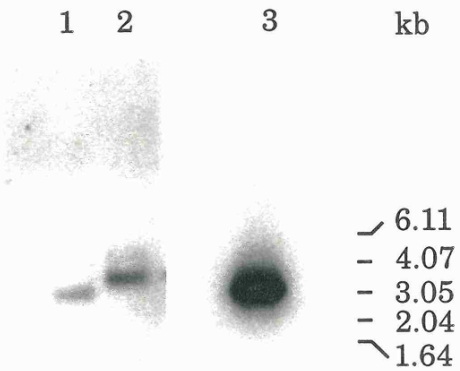
a) Agarose gel prior to blotting for Southern blot analysis

b) Autoradiograph of hybridised membrane

a)



b)



The reason for the failure of this method became apparent when the *Rhizobium* dehalogenase nucleotide sequence was obtained (Section 5.1.2).

3.5 Summary

The results presented in this chapter show that by using direct selection methods, it was possible to identify the cloned *Rhizobium* dehalogenase genes. However it was not possible to identify directly genes that enabled *E.coli* NM522 to grow on DCP. Activity towards this compound had been measured in cells containing pSC530 but growth was not observed. Apparently there are other proteins that may be required for growth on DCP and the most likely candidate for the missing enzyme was a DCP uptake system. Experiments have shown that the two clones isolated carry genes for different dehalogenase enzymes as determined by their substrate specificities, with the dehalogenases expressed from pSC1 being able to act only on 2CP, whilst the dehalogenase expressed from pSC530 is able to act on 2CP and DCP.

The model presented by Leigh (1986) shows the proposed arrangement of the dehalogenase genes as determined by studies of the mutants that were obtained. This proposed that the regulatory gene(s) was distinct from the structural genes. The clones obtained do not possess a regulatory gene that was expressed in *E.coli*, even in the case of pSC1 where there was a large region of DNA cloned and so these clones have been less informative in terms of substantiating the model for the regulation of the genes as proposed by Leigh than would be liked. Leigh proposed that the genes were positively regulated and this is borne out in the cloning of the dehalogenase genes, as the lack of a regulatory gene product leads to no induction of the genes, as was seen for the genes on pSC1 and pSC530

Restriction enzyme analysis of the two clones had shown that they are distinct regions of DNA as they contain no common DNA and further analysis by Southern blotting showed that the two fragments of DNA are not from a contiguous region of the *Rhizobium* genome but are spatially separate. They also show little similarity to each other as determined by the lack of hybridisation between the two fragments. The Southern blotting experiments also show that the clones were obtained from the same organism and neither arose from a contaminant that was obtained during the selection procedure. At this stage it was believed that pSC1 codes for Dehalogenases I and III, whilst pSC530 codes for Dehalogenase II, although it cannot be ruled out at this stage that either one or both of the clones may code for a previously undescribed dehalogenase enzyme.

Further work on characterising the enzymes produced by pSC1 and pSC530 was undertaken and this will be presented in the following chapters, along with the analysis of the two DNA fragments and their encoded genes.

Chapter 4

Further Analysis of pSC1 and pSC530

4.1 Further Analysis of pSC1 and pSC530

The previous chapter dealt with the isolation and primary characterisation of two regions of *Rhizobium* genomic DNA that coded for dehalogenase enzymes. This chapter will describe the further analysis of these clones and the dehalogenases that they encode. They will be dealt with in two separate sections, each dealing with an individual clone.

4.1.1 Analysis of pSC1

In order to analyse further pSC1 it was decided to make a series of subclones of the original construct and to see how this effected the dehalogenase activity of the clone.

4.1.2 Subcloning of pSC1

From the data obtained during the restriction mapping of pSC1 it could be seen where there were useful sites from which subclones could be obtained. The first subclone that was constructed was the deletion of approximately 2.0 kbp to form a 4.8 kbp *Eco*RI-*Sal*I fragment from map position 0 to map position 4.8 (Figure 3.4). pSC1 DNA was digested with the appropriate restriction enzymes and the desired fragment was purified using agarose gel electrophoresis and isolated from the gel using the Geneclean procedure. This DNA fragment was then ligated into *Eco*RI-*Sal*I digested pUC19 and transformed into *E.coli* NM522. Recombinant colonies were selected on the basis of antibiotic resistance and a white coloration using the blue/white insertional inactivation selection method and a number of the colonies were grown overnight in order to prepare plasmid DNA from them. Restriction analysis of the mini-prepped DNA showed that it contained the expected insert and so the plasmid was designated pSC248. Experiments to determine the dehalogenase activities of the subclone were then carried out. The new plasmid retained the charcteristics of the parent plasmid, that is it

allowed growth on both the D and L isomers of 2CP. Cell-free extracts of *E.coli* NM522::pSC248 grown to late exponential phase with 2CP as the sole carbon and energy source and in the presence of IPTG were made and assayed against 2CP in order to determine the level of dehalogenase activity and it was found that the specific activity was 0.61 μmol min^{-1}mg $protein^{-1}$. This is twice the activity seen for the parent plasmid, and greater than that seen in the *Rhizobium*. If, however, the growth rates on 2CP of cells carrying pSC1 or pSC248 are compared, then there is little difference, the values for the doubling times being 16h and 15h respectively. This shows that the increase in enzyme activity has little effect on the rate of growth and so there must be a factor other than the dehalogenase activity that is now limiting the rate of growth. The obvious candidate for this is the uptake system, which, as mentioned in the previous chapter, could be the propionate or lactate uptake system. If the rate at which the 2CP is transported into the cell is less than the rate at which the dehalogenases are able to metabolise it to lactate the growth rate would become transport dependant. As was described in the previous chapter (Section 3.2.3) the difference between the growth rates of the cells carrying either pSC1 or pSC530 was proportional to the difference in their dehalogenase specific activity. It has now been shown that as the specific activity increases, there is no further decrease in the growth rate of the cells, so the dehalogenase activity is no longer the rate limiting step for growth. The specific activity at which the switch from dehalogenase activity to uptake being rate limiting occurs somewhere between the two values for pSC1 and pSC248, i.e. 0.3 μmol min^{-1}mg $protein^{-1}$ and 0.6 μmol min^{-1}mg $protein^{-1}$.

The next subclone to be constructed was a 2.0 kbp *Pst*I fragment from map position 1.4 to map position 3.4 (Figure 3.4). The pSC245 DNA was digested completely with *Pst*I and then electrophoresed on a 0.8% agarose gel in order to purify the required band. After extracting the band from the

gel it was ligated into *Pst*I digested pUC18 and pUC19. A number of recombinant transformants were obtained using the blue/white selection procedure and DNA was prepared from these colonies and used for restriction mapping experiments in order to determine the orientation of the fragment using an asymmetric *Xho*I restriction site. A construct was found that contained the fragment in the correct orientation in pUC18 and this plasmid was designated pSC2. Experiments were then carried out to determine the dehalogenase activities expressed by pSC2. Cells containing pSC2 were plated out onto selective medium containing separately either D- or L- 2CP and incubated at 30°C. Once again this plasmid enabled *E.coli* to grow on both of the isomers of 2CP and so it seemed likely that it contained the genes coding for the Dehalogenases I and III. The next stage was to assay cell-free extracts for activity against 2CP. Cells were grown with 2CP as the sole carbon and energy source and IPTG as inducer until late exponential phase, harvested and cell-free extracts were made and the specific activity of the extracts determined. The activity towards 2CP was 1.02 μmol min^{-1}mg $protein^{-1}$ This is approximately 3x the activity seen in the parent clone pSC1, but determination of the growth rate of *E.coli* NM522::pSC2, which had a doubling time of 15h, showed that it was not significantly quicker than that of *E.coli* NM522::pSC1, indicating that the hypothesis that the rate limiting factor changed from dehalogenase activity to uptake may be correct. SDS-PAGE analysis was carried out as the level of activity of pSC2 indicated that the protein was probably produced in quite high amounts and so it might be possible to determine the size of the proteins produced from pSC2 using SDS-PAGE. The gel showed two bands each produced to a level of approximately 10% of total cell protein, with an approximate molecular weight of 29 kDa and 31 kDa, respectively. In order to determine if these two bands were associated with the dehalogenase enzymes, a construct was made which reversed the orientation of pSC2 and

so abolished the expression of the dehalogenase genes as it is known from work in Chapter 3 that the genes are expressed from the *lac* promoter of pUC. The plasmid pSC219 was constructed by digesting pSC2 with *Eco*RI and *Hin*dIII and after gel-purifying the corresponding band by agarose gel electrophoresis the fragment was ligated into *Eco*RI/*Hin*dIII digested pUC19 as this has the effect of reversing the orientation of the insert DNA with respect to the *lac* promoter of pUC18. After transformation into *E.coli* NM522 cells containing pSC219 were grown in minimal medium containing glycerol as the carbon source for growth, ampicillin to maintain the presence of the plasmid and 2CP and IPTG were present to induce any dehalogenase activity. Cell-free extracts were prepared and assayed for activity. No dehalogenase activity was detected and SDS-PAGE analysis showed that the two proteins previously seen to be expressed in high levels were no longer present. This indicated that both these proteins were expressed from the plasmid pSC2 and absence of the proteins lead to loss of dehalogenase activity so these proteins were likely to be subunits of one or more dehalogenase enzymes.

4.2 Determination of the number of dehalogenase enzymes present on pSC2

The next question that needed to be approached was whether the two proteins expressed from pSC2 were subunits of two dehalogenase enzymes with the conformation of α_n and $ß_n$ or of a single dehalogenase enzyme with an $\alpha_n\beta_n$ conformation. The former possibilty seemed more likely because results reported in Chapter 3 showed that pSC1 codes for activities that would be expected for dehalogenases I and III, but not dehalogenase II. In order to clarify this, two approaches were taken. The first method was to subclone further the plasmid pSC2 to see if disruption of expression of one of the genes leads to loss of one or both dehalogenase activities. At the

same time it would be possible to determine which of the overexpressed proteins was encoded by a discrete region of DNA. The second method involves purifying and characterising the dehalogenases to demonstrate how many proteins are associated with each dehalogenase activity and so the number of dehalogenases can be defined. Both of these approaches were carried out in parallel. The results for the subcloning approach will be presented first followed by those for the protein purification approach.

4.2.1 Subcloning of pSC2

In order to produce subclones that would only express one of the cloned proteins, restriction sites which cut within the putative structural genes were required. As the insert of pSC2 was 2.0 kbp in size and this amount of DNA is approximately the minimum required to code for the two approximately 30 kDa proteins that were known to be expressed from the plasmid, then sites occurring in the first and last third of the insert should disrupt only one of the two genes. The first subclone made involved the use of the *Xho*I site at map position 1.7 (Figure 3.4). The pSC2 plasmid DNA was digested completely with *Xho*I and *Hin*dIII and the DNA was purified using agarose gel electrophoresis. The desired DNA band was isolated from the gel and the DNA was then exposed to S1 nuclease which removed the single stranded incompatible overhangs generated by the two restriction enzymes so the blunt ends of the DNA can then be ligated together. The reaction was heat inactivated and the DNA was electrophoresed on an agarose gel in order to purify the required band. The DNA was then ligated overnight and transformed into *E.coli* NM522. Several recombinant colonies were selected and DNA was prepared for analysis. Plasmid DNA with the desired restriction pattern was identified and the plasmid was designated pSC3. Cells carrying pSC3 were plated onto minimal medium plates containing either 2CP, D-2CP or L-2CP as the sole carbon and energy

source and IPTG. Growth was seen on these plates containing 2CP and D-2CP after incubation at 30°C so it was decided to determine the dehalogenase activity of cells carrying pSC3. Cells containing pSC3 were grown to late exponential phase with 2CP as the sole carbon and energy source and IPTG. The cells were harvested and extracts prepared. These extracts were then assayed using 2CP and had a specific activity of 1.13μmol min^{-1}mg $protein^{-1}$. Assays were then carried out using the D and L isomers individually as the substrates for the dehalogenase. Dehalogenation was only found to occur when D-2CP was the substrate, there was no Cl^- release seen when L-2CP was present. Activity towards a single stereoisomer was confirmed by the results of the growth tests. SDS-PAGE analysis of the cell-free extracts showed that there was only one of the two previously highly expressed bands present and this was the lower 29 kDa band. From these results it can be seen that the plasmid pSC3 expressed a single dehalogenase with activity towards only D-2CP. This evidence showed that the 29 kDa protein expressed by pSC3 is the subunit of Dehalogenase III, the dehalogenase that acts only on D-2CP. It also suggested that there were two discrete dehalogenase enzymes encoded by the genes on the pSC2 plasmid, rather than a single non-stereospecific $\alpha_n\beta_n$ dehalogenase enzyme.

To confirm the suggestion that there were two discrete dehalogenases it would be necessary to produce a subclone that only expressed the L-2CP dehalogenase activity. In order to do this, the *Hin*cII restriction site at map position 2.5 (Figure 3.4), which combined with the *Hin*cII site in the polylinker of pUC18 would allow the removal of the 1 kbp of insert DNA known to contain the gene for Dehalogenase III and result in the expression of the remaining dehalogenase gene. pSC2 DNA was completely digested with *Hin*cII and the required DNA band was purified by agarose gel electrophoresis and the DNA isolated from the gel. This DNA was then

ligated overnight and transformed into *E.coli* NM522. The transformants were selected by antibiotic resistance and then plasmid DNA was prepared from a number of colonies and subjected to restriction digests to determine if the correct plasmid had been constructed. The correct restriction pattern was obtained and the plasmid was designated pSC4. To ascertain the nature of the dehalogenase activity encoded by pSC4, cells were grown to late exponential phase with 2CP as sole carbon and energy source and IPTG present to induce the *lac* promoter of pUC18. Cell-free extracts were prepared and the dehalogenase activity towards 2CP determined. The extracts were highly active towards 2CP, with a specific activity of 2.86 μmol $min^{-1}mg$ $protein^{-1}$ and when tested with the individual D and L isomers of 2CP, only activity towards the L isomer was seen, as was expected. This was again confirmed by growth tests that showed that cells carrying pSC4 were able to grow on L-2CP but not D-2CP. These results show that the subcloning had resulted in the construction of a plasmid that encoded solely the L-specific 2CP dehalogenase activity, that is Dehalogenase I. SDS-PAGE analysis of the cell-free extracts from cells containing pSC4 was carried out and only the 31 kDa protein was expressed, indicating that it was this protein that formed the subunit of Dehalogenase I, the enzyme that acts stereospecifically on L-2CP.

4.2.2 Summary of the subcloning of pSC1

The result of the first set of subcloning experiments was the construction of a plasmid, pSC2, which contained the genes for two proteins which were highly expressed during induction by IPTG and were associated with dehalogenase activity. The next set of subcloning experiments allowed pSC2 to be subcloned further to give two plasmids each of which contained the structural gene for one of the two proteins encoded by pSC2, i.e. pSC3 expressed only the D-2CP specific Dehalogenase III now referred to as

Haloacid dehalogenase D (HadD), whilst pSC4 expressed only the L-2CP specific Dehalogenase I, now referred to as Haloacid dehalogenase L (HadL). The subcloning also confirmed that the two proteins expressed from the plasmid pSC2 during induction by IPTG were the individual subunits of two distinct dehalogenase enzymes and not the subunits of the same dehalogenase. A summary of the subcloning and the activities of the subclones can be seen in Figure 4.1 and Table 4.1. Figure 4.2 shows the proteins expressed from the plasmids pSC2, pSC3 and pSC4.

4.2.3 Purification of the dehalogenases expressed from pSC2

The *Rhizobium* dehalogenases had never been previously been purified to homogeneity even though partial purification had been accomplished by means of ammonium sulphate precipitation and DEAE-cellulose ion-exchange chromatography (Allison 1981, Leigh 1986). Early in the project, purification of the *Rhizobium* dehalogenases was attempted. If the dehalogenases could be purified, their N-terminal sequences could be used to design probes for identifying clones of the dehalogenase genes. Initial attempts were made to follow the methodology of Leigh (1986), but this was unsuccessful and so a protocol using FPLC was devised. The conditions under which the dehalogenase activity bound to a MonoQ 5/5 column were determined, but separation of the dehalogenases from other cellular proteins was poor. A second step was introduced with the loading of the active MonoQ fractions onto a Phenyl Superose 5/5 column. A decreasing gradient of 1.7M-0M ammonium sulphate was run, with the dehalogenase activity being bound to the column until the end of the gradient. When the fractions were assayed, a significant percentage of the activity (70%) was

Figure 4.1

Summary of the subcloning of pSC1. Key to enzymes: E=*Eco*RI, B=*Bam*HI, P=*Pst*I, X=*Xho*I, Sp=*Sph*I, K=*Kpn*I, S=*Sal*I and Hi=*Hin*cII. Bold line denotes the plasmid pUC19 (pSC248) or pUC18 (pSC2/3/4), fine line indicates insert. Hatched region indicates polylinker (not to scale). Arrow shows direction of transcription of *lac* promoter.

Figure 4.1

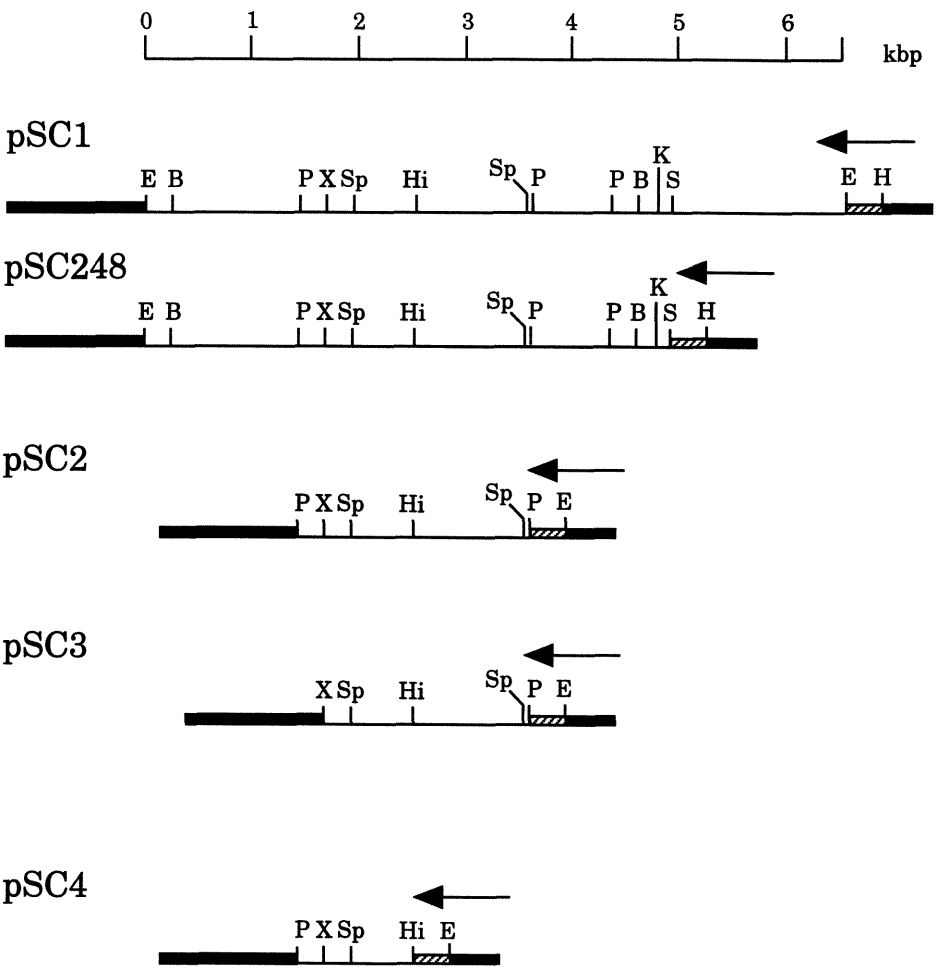

0
1
2
3
4
5
6
kbp
pSC1
E B
P X Sp
Hi
Sp
P
K
P B
S
E H
pSC248
E B
P X Sp
Hi
Sp
P
K
P B
S
H
pSC2
P X Sp
Hi
Sp
P E
pSC3
X Sp
Hi
Sp
P E
pSC4
P X Sp
Hi E

Table 4.1

Dehalogenase activities of subclones of the plasmid pSC1. All cells were grown with 10 mM 2CP as the sole carbon and energy source and with IPTG (0.3mM) present to induce the *lac* promoter. Values are a mean of at least five determinations.

	Specific Activity (μmol Cl^- release $min^{-1}mg\ protein^{-1}$)
E.coli NM522::pSC2	1.02
E.coli NM522::pSC3	1.13
E.coli NM522::pSC4	2.86

Figure 4.2

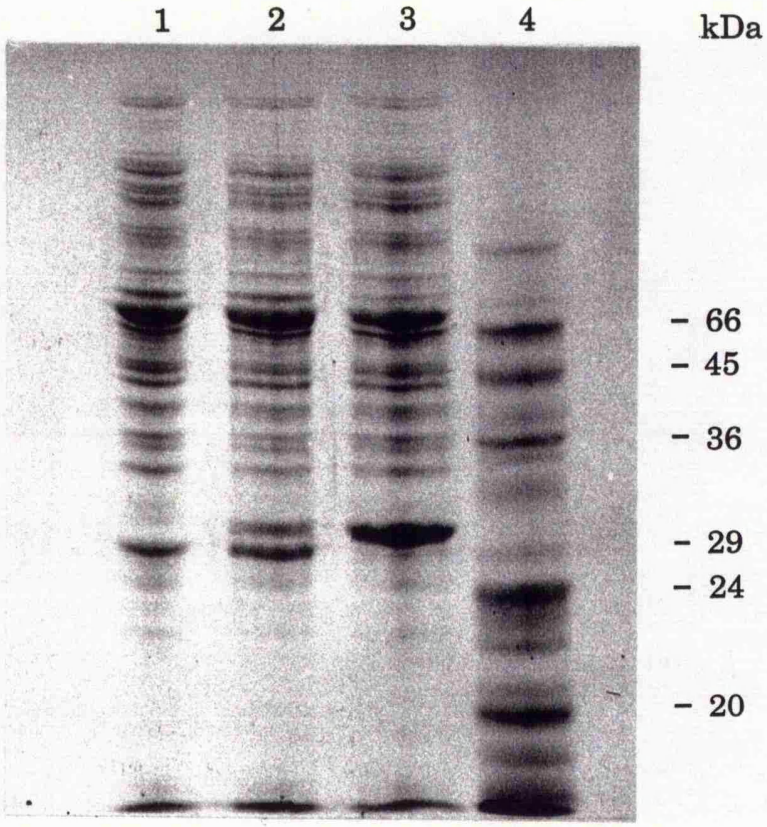


Figure 4.2.

SDS-PAGE gel of proteins expressed from pSC2,pSC3 and pSC4. Lane 1 : pSC3, Lane 2: pSC2, Lane 3: pSC4 and Lane 4: Molecular Weight Markers.

lost. As the presence of ammonium sulphate was not detrimental to dehalogenase activity (Allison 1981), the properties of the Phenyl Superose column must cause the loss of activity. Gel-filtration was not thought to be a feasible third step due to the low level of activity and the poor level of purification achieved up to that point. At this stage the purification was left to concentrate on the direct selection, but the FPLC methodology was employed in the purification of the cloned dehalogenases.

As the levels of expression from pSC2 were quite high, it was decided to purify both dehalogenases to homogeneity using FPLC techniques that had been determined for the *Rhizobium* dehalogenase. Purification of the dehalogenases would allow the N-terminal amino acids to be sequenced, allowing the correct open reading frame to be assigned when the genes were sequenced.

Cells containing the plasmid pSC2 were grown in 200ml of minimal medium with 2CP as the carbon source and IPTG was added to a final concentration of 0.3mM to ensure the full induction of the *lac* promoter on the plasmid. The cells were grown at 30°C with vigorous shaking until they had reached the late exponential phase of growth. The cells were then harvested by centrifugation and the cell pellet was washed in 0.1 volumes of 100mM Tris.acetate, 1mM EDTA, 10% glycerol buffer pH7.5 and then centrifuged as before. The pellet was then resuspended in 0.04 volumes.of the same buffer The cell-free extract was prepared by sonicating the cell suspension followed by centrifugation to remove any unbroken cells and cellular debris. The resulting supernatant was assayed for dehalogenase activity and kept at 4°C until used for the purification.

The supernatant containing 10mg of protein in 4ml was loaded onto a MonoQ HR 5/5 anion exchange column in 20mM Tris.acetate, 1mM EDTA, 5% glycerol buffer pH8.0 and eluted using a linear gradient of 0-2M sodium

acetate over a volume of 15ml, at a flow rate of 1ml min^{-1}. The eluted proteins were detected by their absorbance at 280nm and the elution profile was recorded. The eluted proteins were collected in 1ml fractions and were stored on ice before being assayed for dehalogenase activity. In order to allow the rapid assaying of a large number of fractions the initial assay was carried out qualitatively in a microtitre plate. To 100µl of 100mM Tris.acetate buffer pH7.6 was added 1µl of 1M Cl^--free 2CP and 10µl of fraction. This was done for all of the fractions eluted from the column. The plate was then incubated at 30°C for 10 minutes and then 10µl of 0.25M ferrous ammonium sulphate followed by 10µl of mercuric thiocyanate-saturated ethanol was added and the reaction mixture was thoroughly stirred. A colour change indicated the presence of free chloride ions. Fractions which showed dehalogenase activity could then be quantitatively assayed by the usual method. This micro-assay method also quickly showed how many regions there were of dehalogenase activity in the elution profile and how closely associated they were.

During initial protein purification runs all of the dehalogenase activity was bound to the column. When the salt concentration was increased a single peak of dehalogenase activity eluted from the column at a salt concentration of 520mM sodium acetate. This showed that the two dehalogenases had a very similar surface charge as they eluted at the same ionic conditions. This was unexpected as previously it had been shown that it was possible to separate Dehalogenases I and III by ion-exchange chromatography (Leigh 1986), albeit using a different ion exchange medium and elution buffer. Changing the FPLC conditions by varying the nature of the gradient over which the proteins were eluted and/or changing the pH of the elution buffer, had no effect on the separation of the dehalogenases and activity was always eluted as a single peak. The elution profile of the dehalogenases from the MonoQ 5/5 column is shown in Figure 4.3. Analysis by SDS-PAGE

of the eluted proteins from the fractions containing dehalogenase activity and those neighbouring fractions that were inactive showed that the presence of activity was associated with the presence of the two proteins that were highly expressed from pSC2. This did not, however, enable the question of the number of dehalogenases encoded on pSC2 to be answered.

After the MonoQ column each of the dehalogenases accounted for approximately 35% and 35% respectively of the total protein, and although they did not need to be 100% pure before they were blotted and sequenced, the fewer contaminating proteins present, the less chance of having ambiguous sequence. The next purification step was gel-filtration chromatography. This was carried out using a Superose 12 column with an elution buffer of 100mM Tris.acetate, 100mM sodium acetate pH7.6 and a flow rate of 0.4ml min^{-1}. The presence of the extra Na acetate was to prevent any interaction with the gel-filtration matrix. As well as giving further purification, use of this column would also show whether the native dehalogenases were of a similar size, or if they are of greatly differing sizes that would facilitate their separation. The elution was monitored as before and the fractions tested by the micro-assay. Once again a single region of dehalogenase activity was seen, indicating that the dehalogenases were of a similar native molecular weight.

After the gel-filtration step, the fractions containing dehalogenase activity were separated by SDS-PAGE and this allowed the extent of the purification achieved to be monitored. The SDS-polyacrylamide gel showed that some degree of further purification had been achieved and so it was decided to determine the N-terminal amino acid sequence of the dehalogenases from the gel-filtration fractions. Two wells were loaded on

Figure 4.3

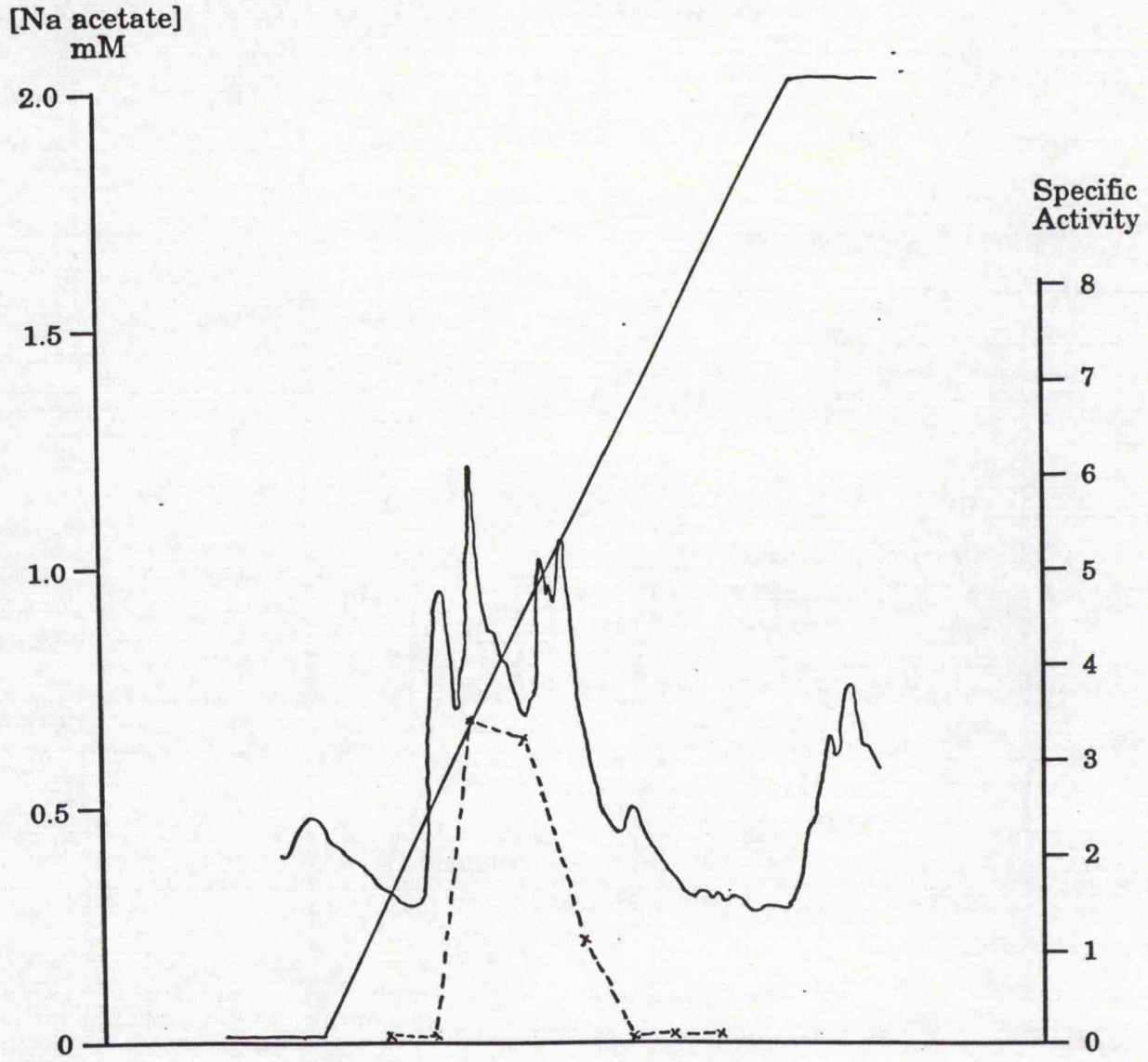


Figure 4.3

Elution profile of extracts from cells containing pSC2. Extract was loaded onto a MonoQ 5/5 column and eluted with a linear 0-2M sodium acetate gradient in buffer as described in Chapter 2 (Section 2.23.1). The elution profile of dehalogenase specific activity (μmol Cl^- min^{-1} mg $protein^{-1}$) is superimposed on the protein trace.

an SDS-polyacrylamide gel with 10μg of protein and electrophoresis was carried out. The gel was then supplied to the protein sequencing facility at the University of Leicester. The sequence analysis was carried out by Dr. K. Lilley and Miss. E. Cavanaugh as described in Chapter 2. Sequences were obtained for the two proteins that were overexpressed from pSC2, with the first eight residues being obtained for the 29 kDa protein, with one unassigned residue at position seven and the first thirteen residues were obtained for the 31kDa protein with residues four and five being unassigned. The unassigned residues may be either cysteine, serine, histidine or arginine (E. Cavanaugh, pers. comm.). These sequences are shown in Figure 4.4 and Figure 4.5. When the sequences were compared with the N-terminal sequence of the α-fragment of ß-galactosidase there was no similarity indicating that neither of the proteins were expressed as fusion proteins formed from the α-peptide and the cloned protein. The sequences also showed no similarity to each other or to other published dehalogenase N-terminal amino acid sequence.(Schneider *et al* 1991, Barth *et al* 1992 and Jones *et al* 1992).

The construction of plasmids pSC3 and pSC4 allowed the confirmation that the two proteins expressed from pSC2 were the two dehalogenases HadD and HadL and so it was decided to purify the HadD and HadL separately using the two constructs that allowed them to be expressed individually.

Cells carrying either pSC3 or pSC4 were grown on minimal medium with 2CP as the sole carbon source and IPTG (0.3mM) present to ensure full induction of the *lac* promoter. The cell-free extracts were then used for the purification of the dehalogenases using the purification scheme already designed. Each extract was applied to the MonoQ 5/5 column under the same conditions and the fractions were collected and assayed as described earlier. Both of the dehalogenases eluted at the same salt concentration,

Figure 4.4

N-terminal sequence of Dehalogenase I (HadL) expressed from the plasmid pSC2

	Residue	**Amount**(pmoles)
1.	Ser	2.45
2.	Leu	13.47
3.	Lys	5.55
4.	X	?
5.	X	?
6.	Ile	8.20
7.	Lys	3.68
8.	Ala	8.35
9.	Leu	7.86
10.	Thr	6.65
11.	Phe	6.38
12.	Asp	6.00
13.	Thr	4.66

X indicates an undetermined residue

Figure 4.5

N-terminal sequence of Dehalogenase III (HadD) expressed from the plasmid pSC2

	Residue	**Amount**(pmoles)
1.	Met	5.19
2.	Ile	7.61
3.	Asp	3.94
4.	Leu	8.21
5.	Pro	6.33
6.	Arg	3.44
7.	X	?
8.	Pro	6.96

X indicates an undetermined residue.

(520mM sodium acetate) as was seen previously when they co-eluted. Visual analysis of the active fractions by SDS-PAGE showed that the dehalogenases each constituted approximately 80% of the total protein in each fraction and it was decided that this was pure enough to allow an accurate size determination by gel-filtration. The most active fraction for each of HadD and HadL was applied separately to two Superose 12 columns linked in series. The combination of two columns allowed a more accurate determination of the molecular weight due to the greater volume. The samples were run in the buffers previously stated with a flow rate of 0.2ml min^{-1}. After each run of the dehalogenase to be sized, the columns were calibrated with known molecular weight standards. The molecular weight calculated for HadD was approximately 58 kDa, whilst the molecular weight calculated for HadL was approximately 60 kDa. These results show that in their native form, the dehalogenases HadD and HadL have a dimeric conformation, with subunit sizes of 29 kDa and 31 kDa, as determined by SDS-PAGE, respectively. As it was known that HadD and HadL both co-eluted when passed through a Superose 12 column, it is reasonable to assume that they both have similar native molecular weights, and this was seen when the two proteins were passed through the linked Superose 12 columns. The elution profiles for the MonoQ column are shown for both HadD and HadL in Figures 4.6 and 4.7. An SDS-polyacrylamide gel showing the purification of HadD and HadL is shown in Figure 4.8.

Figure 4.6

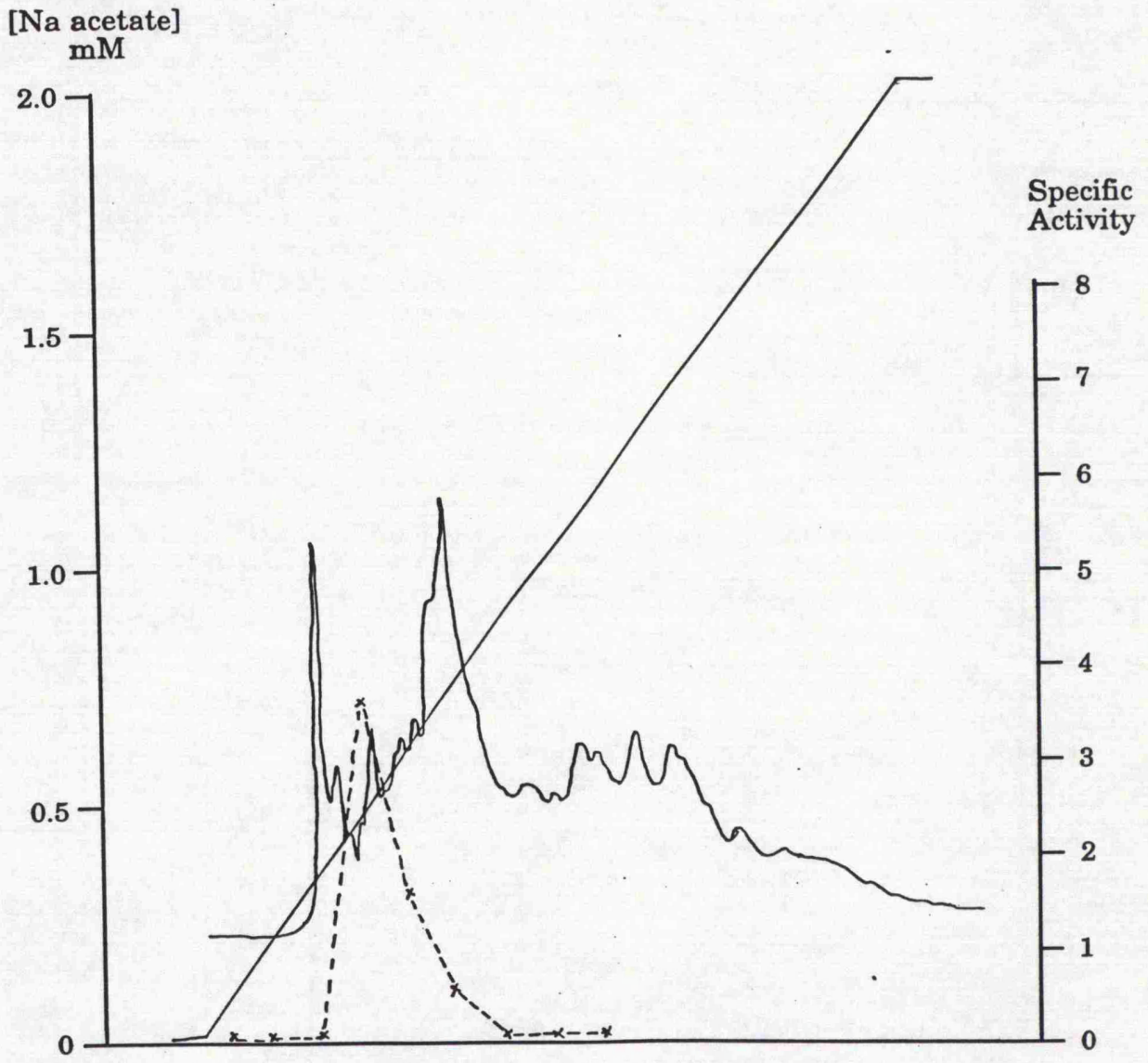


Figure 4.6

Elution profile of extracts from cells containing pSC3. Extract was loaded onto a MonoQ 5/5 column and eluted with a linear 0-2M sodium acetate gradient in buffer as described in Chapter 2 (Section 2.23.1). The elution profile of HadD dehalogenase activity is superimposed on the protein trace.

Figure 4.7

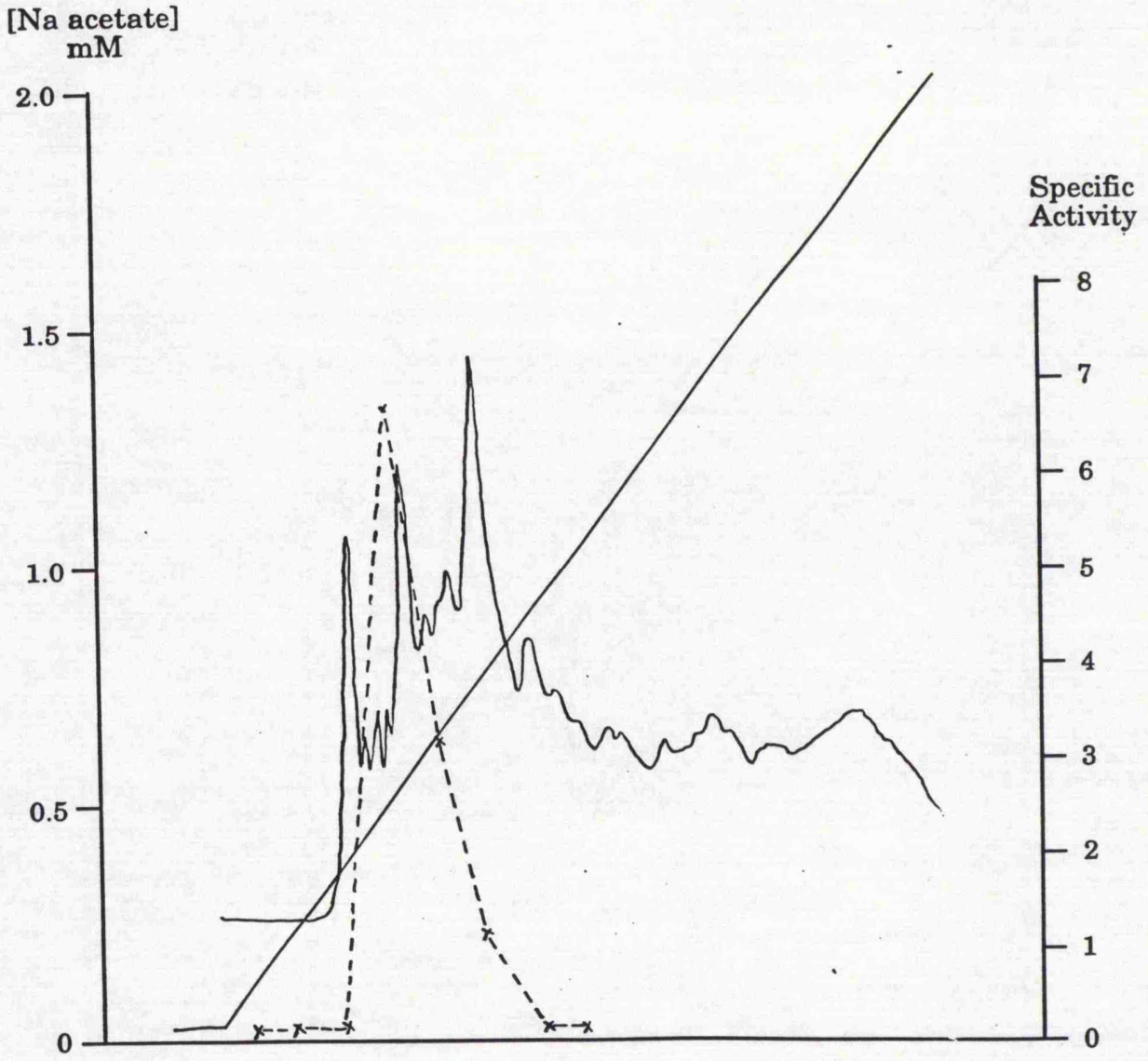


Figure 4.7

Elution profile of extracts from cells containing pSC4. Extract was loaded onto a MonoQ 5/5 column and eluted with a linear 0-2M sodium acetate gradient in buffer as described in Chapter 2 (Section 2.23.1). The elution profile of HadL dehalogenase activity is superimposed on the protein trace.

Figure 4.8

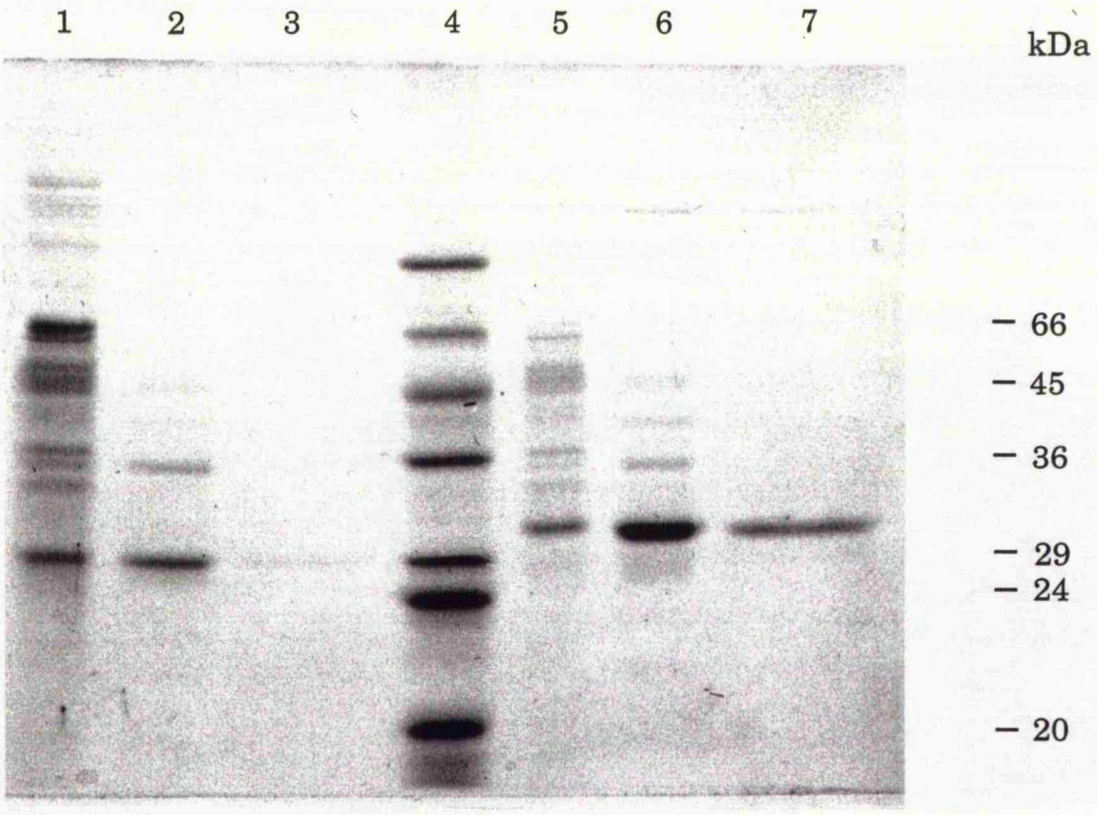


Figure 4.8

SDS-PAGE gel of the purification of HadD and HadL. Fractions contained maximum dehalogenase specific activity. Lane 1: cell-free extract of cells containing pSC3, Lane 2: MonoQ fraction, Lane 3: Superose 12 fraction, Lane 4: Molecular weight markers, Lane 5: cell-free extract of cells containing pSC4, Lane 6: MonoQ fraction and Lane 7: Superose 12 fraction.

4.2.4 Purification of HadL from the *Rhizobium* sp.

In order to establish that the cloned dehalogenase genes were from *Rhizobium* sp. and confirm the Southern blotting analysis evidence presented in Chapter 3, it was decided to purify one of the dehalogenases from the *Rhizobium* and then have the N-terminal amino acid sequence determined. This would then be compared to the amino acid sequence that had been determined for the cloned dehalogenases. Although purification of a Rhizobial dehalogenase was not possible before, now that the dehalogenases had been identified from the cloned genes, identification of the equivalent dehalogenases from the *Rhizobium* would be easier. It was decided to use the protocol devised for the purification of the cloned dehalogenases. A problem in purifying the dehalogenase was the low level of expression in the *Rhizobium*. This was over come by growing larger volumes of cells, for purification of the dehalogenase it was usual to grow at least 3l of cells, compared to the routine growth of 200ml of cells for the purification of the over-expressed proteins.

The *Rhizobium* was grown on minimal medium with 2CP present as the sole carbon source and the inducer for the dehalogenase genes. Extracts were prepared as for the cloned dehalogenases and applied to the MonoQ 5/5 column under the same buffer conditions. The dehalogenase activity was seen to elute at the same ionic strength, and both D and L-2CP dehalogenase activities were present. The most active fractions were analysed by SDS-PAGE and although the fraction containing the maximum dehalogenase activity contained a considerable number of proteins, it was decided to use a gel-filtration procedure next rather than a Phenyl Superose hydrophobic interaction column which was known from previous experiments to cause inactivation of the dehalogenase enzymes. The most active fraction from the MonoQ run was loaded onto two Superose 12

columns linked in series and run under the same buffer conditions used for the purification of the cloned dehalogenases, but with a slower flow rate of 0.2ml $min^{-1}$to achieve a greater degree of resolution. The most active fraction was determined by assaying the fractions using the micro-assay and then this fraction and the neighbouring less active fractions were analysed by SDS-PAGE. It was possible to see a clear relationship between the dehalogenase activity and a protein band in the region of 31 kDa (HadL). There was also a band in the correct region for HadD, but this was much weaker. The difference in amounts seen for the *Rhizobium* dehalogenases but not seen for the cloned dehalogenases could be due to differing levels of expression of the dehalogenases in the *Rhizobium* (Leigh 1986). The ratios of dehalogenase specific activity between HadL and HadD present in *Rhizobium* mutants were reported to vary from between 2.5:1 to 3.5:1 depending on the inducer (Leigh 1986). After purification the ratio betwen HadL and HadD was approximately 4:1, similar to that seen in *Rhizobium* mutants and indicating that the two dehalogenases had co-purified in a similar manner to the cloned dehalogenases. It was decided to have the 31 kDa protein sequenced and this was carried out as for the cloned dehalogenase proteins.

The N-terminal amino acid sequence was obtained for the first eleven residues of the *Rhizobium* dehalogenase and is presented in Figure 4.9. As can be seen, it is identical to the N-terminal amino acid sequence obtained for the cloned HadL enzyme. This is confirmation that the cloned HadL dehalogenase is from the *Rhizobium* and as both HadD and HadL are found on a contiguous region of DNA then by inference the HadD dehalogenase must also be from the *Rhizobium*. HadD and HadL therefore represent Dehalogenases III and I, respectively.

Figure 4.9

N-terminal amino acid sequence for the Dehalogenase I (HadL) enzyme purified from the *Rhizobium* sp.

	Residue	**Amount**(pmoles)
1.	Ser	2.36
2.	Leu	5.14
3.	Lys	3.42
4.	Lys	4.19
5.	Arg	5.44
6.	Ile	4.06
7.	Lys	3.96
8.	Ala	6.41
9.	Leu	5.72
10.	Thr	2.43
11.	Phe	3.87

4.3 Resolution of 2CP using HadD and HadL dehalogenases

As it was now possible to produce extracts that contained either the HadD or HadL dehalogenase, it was decided to use a combination of these extracts to check the stereospecificity of the dehalogenases and also to see if it was possible to resolve a racemic mixture so that only one of the stereoisomers was removed, a possible industrial use for stereospecific enzymes. The efficiency of both *in vivo* and *in vitro* systems would be tested, as an industrial environment would require the use of whole cells to avoid the cost of harvesting large volumes of cells and making cell-free extracts.

Cultures of cells carrying either pSC2, pSC3 or pSC4 were grown to late exponential phase with 2CP as the sole carbon source and IPTG present to fully induce dehalogenase expression. The cells were harvested, washed and divided into two aliquots. One aliquot was used to make cell-free extracts for use in the *in vitro* experiments whilst the other aliquot was used for the *in vivo* experiments. The cell-free extracts were prepared as described previously and stored at 4°C until use, as were the washed cells.

To reaction vessels containing 100 mM Tris.acetate pH7.6 was added 10mM purified 2CP. Then either whole cells or cell-free extract was added to start the reaction, which was carried out in a shaking water bath at 30°C. The volume of whole cells was such that equivalent numbers of cells were added to each reaction and the volume of cell-free extract was such that the same units of activity were added. Time point aliquots were taken and Cl^- release monitored until a plateau was seen for the reactions containing the either HadD or HadL. At this point, the dehalogenase (either whole cells or extracts) of the opposite stereospecificity was added to the reactions and Cl^- release monitored until it was seen to cease again. The results of these experiments can be seen in Figures 4.10 and 4.11. As can be seen for the *in vitro* experiments (Figure 4.11), in the cases where there was only HadD or

HadL present, the dehalogenase was able to release approximately 50% of the available Cl^-, as determined by the total amount of Cl^- released when extracts prepared from cells containing pSC2 where both dehalogenases are present were used. This indicated that the racemic mixture of 2CP is a 1:1 ratio of D- and L- isomers. Upon the addition of the second dehalogenase, there was the release of the remaining 50% of the available Cl^-. As the reactions are unable to release all of the Cl^- without the presence of both HadD and HadL indicated that the dehalogenases were stereospecific, and that racemic 2CP was composed of approximately equal proportions of D and L-2CP. These *in vitro* experiments showed the feasibility of using this system to resolve a racemic mixture to leave the desired stereoisomer. Investigations using an *in vivo* system for the resolution were also carried out as this is more analogous to an industrial situation (Taylor 1984, Taylor 1988) As can be seen in Figure 4.10 whole cells behaved similarly to cell-free extracts, with the reaction ceasing with 50% of the Cl^- liberated, and going to completion with the addition of the alternative dehalogenase containing cells. The most noticeable difference between the two systems is the time scale over which the reactions take place, with the *in vitro* system having released all of the bound Cl^- within 2 hours of the addition of both HadD and HadL, whereas the use of the *in vivo* system resulted in the complete dehalogenation taking approximately 10 hours, a 5-fold decrease in the rate. This is presumably due to the necessity for the substrate to be transported into the cell before the dehalogenation reaction can take place. The Cl^- ion released will diffuse rapidly out of the cell, and therefore the time for release of the ion from the 2CP molecule to its presence in the medium of the reaction mixture is negligible. The decreased rate of resolution must therefore be a direct consequence of the rate of uptake into the cell.

Figure 4.10

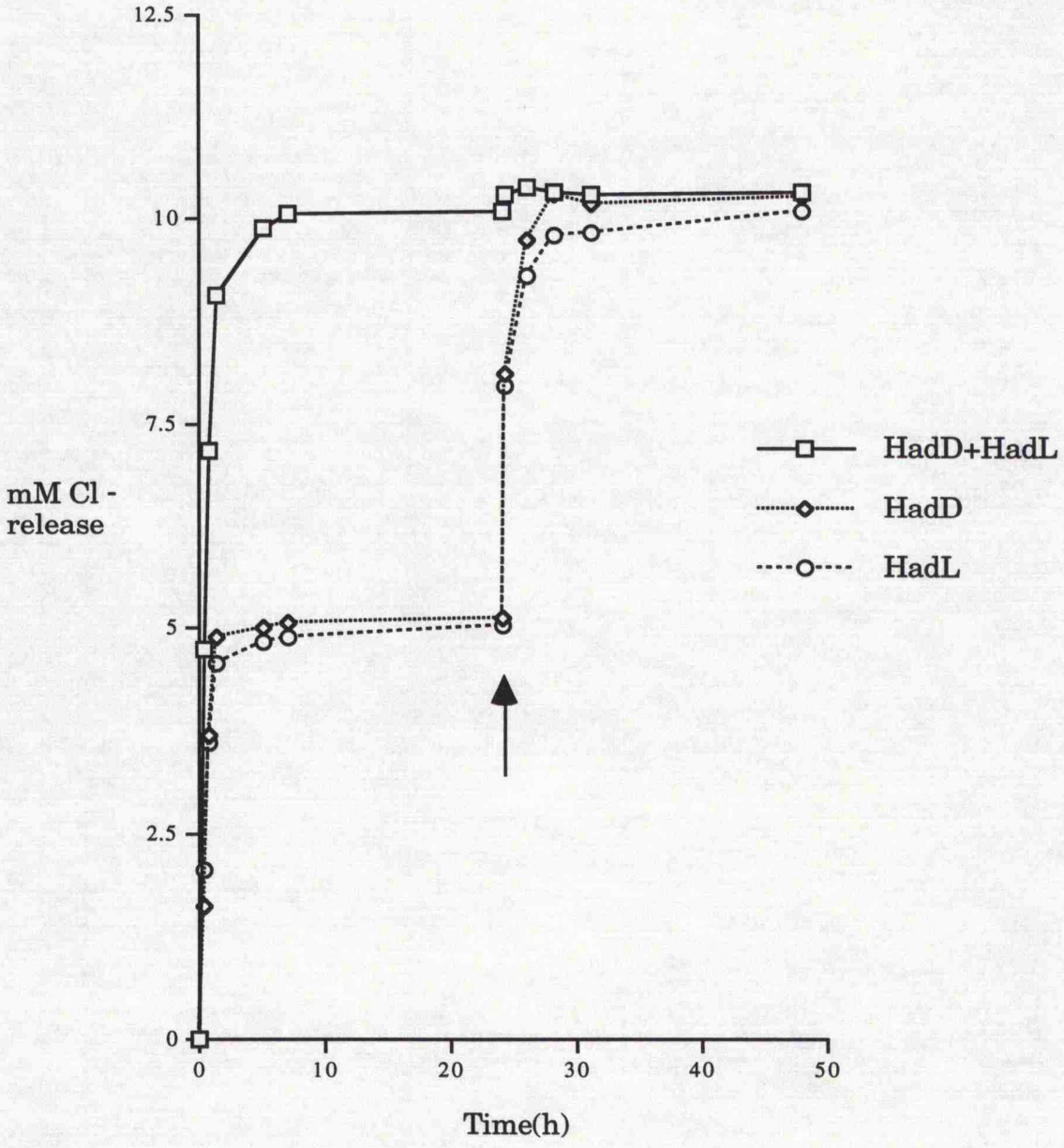


Figure 4.10

Resolution of racemic 2CP by whole cells carrying either pSC2, pSC3 or pSC4. Key indicates the dehalogenase present at the start of the experiment, arrow indicates the addition of dehalogenase of opposite stereospecificity.

Figure 4.11

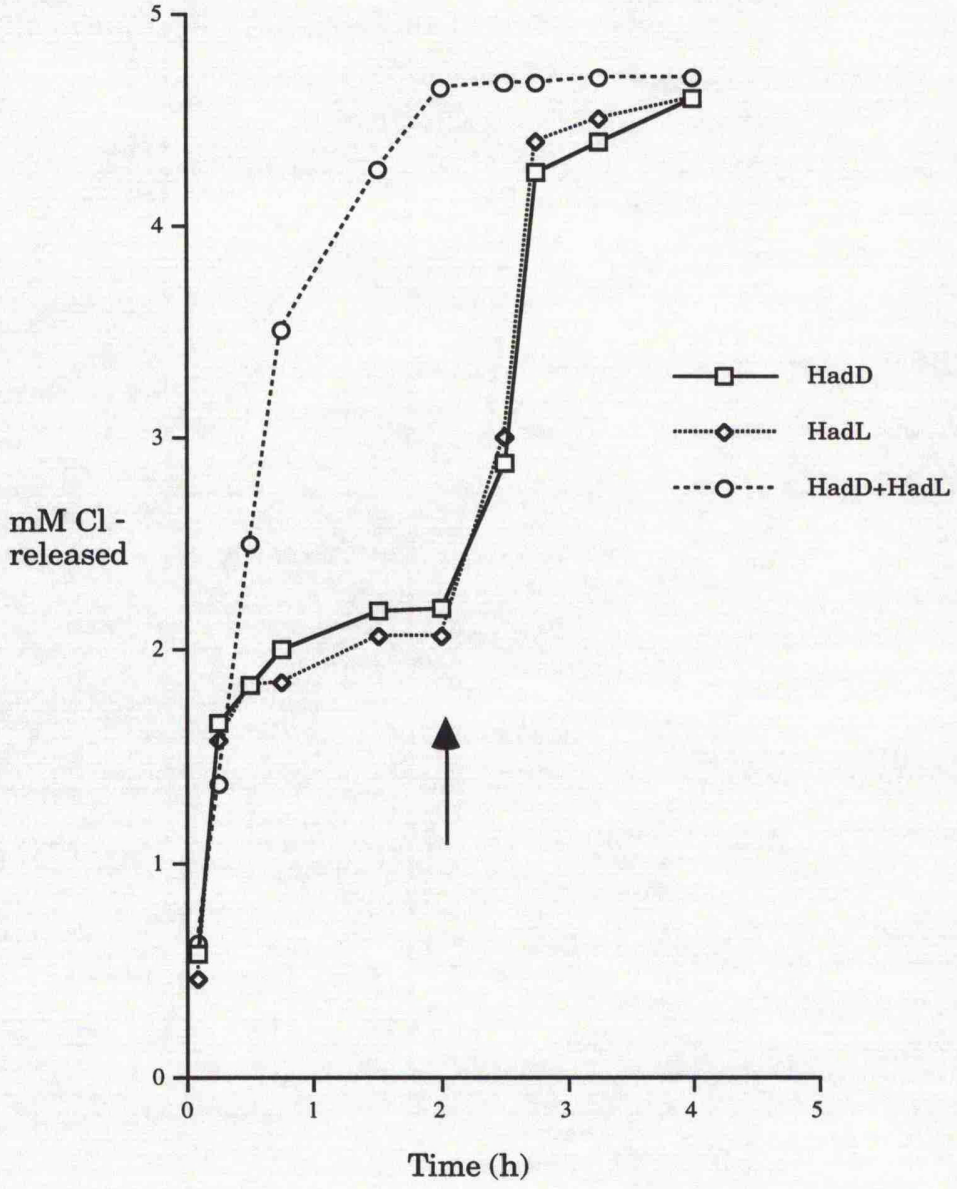


Figure 4.11

Resolution of racemic 2CP by cell-free extracts of cells carrying either pSC2, pSC3 or pSC4. Key indicates the dehalogenase present at the start of the experiment, arrow indicates the addition of dehalogenase of opposite stereospecificity.

4.4 Analysis of pSC530

As described in Chapter 3, cell-free extracts from cells containing plasmid pSC530 showed dehalogenation activity towards 2CP and DCP, indicating that the dehalogenase present may be Dehalogenase II or a previously undescribed dehalogenase. In order to determine more precisely the number and nature of the dehalogenases present, a series of subclones were constructed.

4.4.1 Subcloning of the plasmid pSC530

From the data obtained during the restriction enzyme mapping of the plasmid it was known that there were several restriction enzyme sites that could be used to produce subclones. The first subclones to be produced were the *Hin*dIII-*Eco*RI fragment, map position 0.8 to 3.0 and the *Eco*RI-*Bam*HI fragment at map positions 0 to 2.1. pSC530 plasmid DNA was digested to completion with the appropriate restriction enzymes, the digested DNA was electrophoresed and the desired bands were isolated from the gel. The DNA was then ligated into the appropriately digested polylinker of pUC18 for the *Hin*dIII-*Eco*RI fragment and pUC19 for the *Eco*RI-*Bam*HI fragment, so as to maintain the transcription of the *lac* promoter through the insert in the same direction as that seen in pSC530. *E.coli* NM522 was transformed with the ligated DNA and plated out onto LB-amp plates. Several recombinant colonies were selected and these were streaked out onto minimal medium plates containing 2CP and IPTG. Plasmid DNA was also prepared from these colonies and analysed.

The *Eco*RI-*Bam*HI subclone gave rise to transformants that were unable to grow on 2CP as the sole carbon and energy source and cells grown on lactate and induced with IPTG showed no dehalogenase activity in cell-free extracts. From this it can be assumed that the *Bam*HI site within pSC530

lies internally in the structural gene of the dehalogenase and so digestion at this site results in the loss of the expression of the gene and therefore the loss of dehalogenase activity. Also the loss of all dehalogenase activity through the disruption at only one restriction site indicates that a single gene may be responsible for the dehalogenase activity. The *Hin*dIII-*Eco*RI subclone gave rise to colonies that were able to grow on 2CP as the sole carbon source. The plasmid DNA that was obtained from these colonies was analysed and shown to have the predicted restriction map and this plasmid was designated pSC520. In order to determine the level of expression of the dehalogenase activity, cell-free extracts of cells containing pSC520 were prepared. Cells were grown until late exponential phase with 2CP present as the sole carbon and energy source and IPTG present to induce the *lac* promoter. Cells were harvested and the extracts prepared and assayed for their activity towards 2CP. The specific activity determined for the cell-free extracts was 0.065 μmol min^{-1}mg $protein^{-1}$. This is similar to the specific activity seen for the parent plasmid pSC530, so it appears that the structural gene(s) lies within the *Hin*dIII-*Eco*RI fragment, with the *Bam*HI site situated within the gene (See Figure 3.4). In order to try to increase the level of expression a subclone was made that aimed to decrease the distance between the *lac* promoter and the start of the dehalogenase gene. To do this the *Bgl*II site at map position 2.5 was utilised. pSC520 DNA was completely digested with *Bgl*II and *Sma*I and agarose gel purified. The desired DNA fragment was then isolated from the gel, the overhang of the *Bgl*II site removed with S1 nuclease and the DNA ligated overnight. *E.coli*NM522 was then transformed with the ligation mix and the recombinant cells checked for their ability to grow on 2CP. No growth was seen with 2CP as the sole carbon and energy source and cell-free extracts made from cells carrying the plasmid showed no dehalogenase activity. These results indicate that cutting at the *Bgl*II site abolishes expression of

the dehalogenase gene and so the site is still within the structural gene. This result also shows that in pSC520 the start of the dehalogenase gene must be within 500 bp of the *lac* promoter, as this was the amount of DNA deleted when the previous subclone was constructed. Under these circumstances one would normally expect to see quite a high level of expression, but this is not the case when the cell-free extracts are analysed by SDS-PAGE. It may be that the RBS upstream of the dehalogenase gene is poor or does not function well in *E.coli*, or that there is a problem with folding of the protein leading to a low level of activity. Subclones of pSC530 are shown in Figure 4.12.

4.4.2 Characterisation of the dehalogenase activity of pSC520

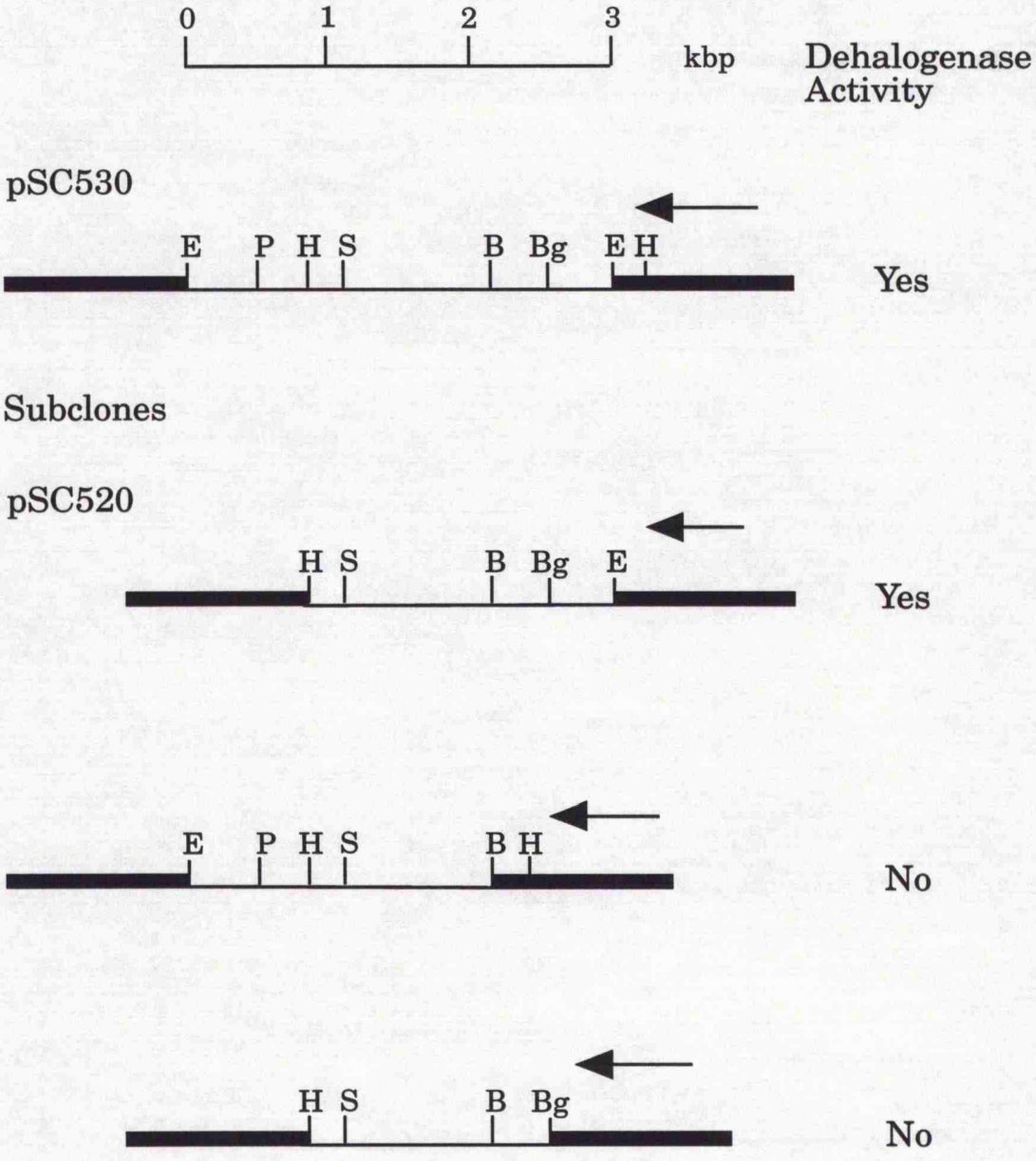
To determine the number and nature of the dehalogenases encoded by the plasmid pSC520 a protein purification protocol based on that used for the purification of HadD and HadL was used. Cell-free extracts were prepared from cells grown on 2CP in the presence of IPTG and these extracts were applied to the MonoQ anion-exchange column under the same buffer conditions as those used for the previous purifications. However, the resulting active fractions contained many contaminating proteins so it was decided to try the gel-filtration column as a further purification step. However, after this step there was little increase in the purity and it was impossible to determine which protein band was the one associated with the dehalogenase activity.

As there were some difficulties in obtaining a high-expressing subclone it was decided to characterise the dehalogenase activity in terms of its substrate profile and inhibition. By comparing this profile to that known for Dehalogenase II it would be possible to say if the cloned dehalogenase enzyme was Dehalogenase II or another dehalogenase not previously

Figure 4.12

Restriction map of pSC530 and subclones derived from it. Dehalogenase activity of each subclone is shown, where there was no activity, the subclone was not named. Vector DNA is indicated by the thick line, with the polylinker shown when present. Arrow denotes direction of transcription of the *lac* promoter. Key to enzymes: E=*Eco*RI, H=*Hin*dIII, B=*Bam*HI, Bg=*Bgl*II, S=*Sal*I and P=*Pst*I.

Figure 4.12


0
1
2
3
kbp
Dehalogenase Activity
pSC530
E P H S B Bg E H
Yes
Subclones
pSC520
H S B Bg E
Yes
E P H S B H
No
H S B Bg
No

characterised. In order to construct the substrate profile, the following halogenated substrates were used; 2CP, DCP, MCA, DCA, monobromoacetate (BrA) and monoiodoacetate (IA). Assays were carried out in 100mM Tris.acetate pH7.6, with the halogenated substrate at a concentration of 10mM. Released halide was determined by the method of Bergman and Sanik (1957). The results of these assays are shown in Table 4.2. As can be seen from this table, the profile obtained for the dehalogenase activity with the cell-free extracts from cells containing pSC520 is the same in order of substrate preference to that previously reported for Dehalogenase II, (Allison 1981, Leigh 1986), i.e. the levels of activity seen are BrA>IA>MCA>2CP>DCP>DCA. The relative activities between the substrates are also very similar to those previously reported, with only slight variations between the values. As a further method of characterising the dehalogenase activity it was decided to carry out an inhibition experiment. Cell-free extract was preincubated in the presence of 1mM N-ethylmaleimide (NEM), a thiol binding agent, for 5 minutes and then used for an assay containing 2CP as the halogenated substrate. The assay was conducted as normal and the specific activity was calculated. The results are shown in Table 4.3. There is an inhibition of 93%, which compares well with the reported 97% inhibition for the Dehalogenase II enzyme (Allison 1981). NEM is known to have little inhibitory effect on Dehalogenases I and III from previous studies (Allison 1981, Leigh 1986). From the substrate profile and the inhibition results it was concluded that the dehalogenase activity found in the cell-free extracts of cells containing the plasmid pSC520 was due to the presence of the cloned gene for the *Rhizobium* Dehalogenase II and the expression of that gene in the *E.coli* cells.

Table 4.2

Substrate profile for the dehalogenase activity of cell-free extracts prepared from *E.coli* cells containing the plasmid pSC520. Specific activity is given in μmol halide released min^{-1} $mg\ protein^{-1}$

Halogenated Substrate	Specific Activity	Relative Activity (%)	Relative Activity (%)[1]
2CP	0.096	100	100[2]
DCP	0.58	60	53[2]
MCA	0.18	187	172[2]
DCA	0.018	19	19[2]
BrA	1.056	1100	1123[3]
IA	0.24	250	257[3]

[1]Relative activity denotes activity for dehalogenase II from the work of [2]Leigh (1986) and[3]Allison (1981). BrA=Bromoacetate, IA=Iodoacetate.

Table 4.3

Inhibition of dehalogenase activity of crude cell-free extracts from cells containing pSC520. Equivalent amounts of dehalogenase was treated as according to the Materials and Methods (Section 2.5.5) and then assayed as normal with 2CP as the substrate. Specific activity is given in μmol Cl^- release min^{-1}mg $protein^{-1}$. The value quoted for dehalogenase II is for the purified enzyme.

Assay Conditions	dehalogenase specific activity	dehalogenase II specific activity(Allison 1981)
No pre-incubation	0.075	0.356
Pre-incubated with 1mM NEM	0.005	0.011
Inhibition (%)	93	97

4.5 Summary

The results presented in this chapter describe the further analysis of the clones that had been obtained. The plasmid pSC1 was subcloned to give a series of plasmids, each of which was then analysed further. The plasmid pSC2 was a 2.0 kbp internal *Pst*I fragment of pSC1 and was shown to have the same growth properties as pSC1. Cell-free extracts showed the same dehalogenase activity, although the specific activity was elevated due to the high percentage of the total protein being cloned dehalogenase protein, approximately 10% of the total cell protein for each of the dehalogenases. The dehalogenase enzymes were purified using the FPLC system and then the N-terminal amino acid sequences were determined. These sequences showed no homology to each other or to previously published dehalogenase N-terminal amino acid sequence. The plasmid pSC2 was then subcloned to give the two plasmids pSC3 and pSC4 each of which has been shown to encode the gene for a single dehalogenase enzyme. The plasmid pSC3 encoded the gene for the D-2CP specific dehalogenase HadD, which was purified and shown to have a native molecular weight of 58 kDa with a subunit size of approximately 29 kDa, whilst the plasmid pSC4 coded for the gene for the L-2CP specific dehalogenase HadL, which was purified and shown to have a native molecular weight of 60 kDa with a subunit size of approximately 31 kDa, indicating that both of the proteins had a dimeric conformation. As HadD and HadL co-eluted from the ion-exchange column then they must have a similar surface charge and so probably iso-electric point. The HadL dehalogenase was then purified from the *Rhizobium* from which the gene was cloned originally. The purified Rhizobial dehalogenase was then subjected to N-terminal amino acid sequencing and the outcome was identical to that obtained for the cloned HadL enzyme, indicating that the clone was obtained from *Rhizobium* genomic DNA, confirming the Southern blotting results presented in Chapter 3.

It was then demonstrated that the HadD and HadL dehalogenases could be used individually to resolve a racemic mixture of 2CP and leave only the desired stereoisomer. This was demonstrated for both *in vivo* and *in vitro* conditions, with the major difference between the two conditions being the time scale for completion of the reaction, with the *in vitro* conditions taking longer due to the necessity for transport of the substrate into the cell. Both enzymes were shown to be completely stereospecific in their reaction, confirming the results published previously (Allison 1981, Leigh 1986).

The analysis of pSC530 showed that only one subclone (pSC520) that was able to express dehalogenase activity could be obtained. Any disruption of the region of insert closest to the *lac* promoter leading to the loss of activity. Purification of the dehalogenase was not successful due to the low level of expression and so characterisation was carried out in terms of substrate profiling and enzyme inhibition to try to establish the nature of the dehalogenase encoded on pSC520. It was seen that the substrate profile in terms of reactivity was the same as that reported for Dehalogenase II and comparison of the ratios of activity of the various substrates with those known for Dehalogenase II showed they were alike. The enzyme inhibition experiment showed that the dehalogenase activity was very susceptible to inhibition by NEM, as has been previously reported for dehalogenase II. From this evidence it seems likely that the dehalogenase encoded on the plasmid pSC520 was Dehalogenase II.

Analysis of pSC1, pSC530 and their respective subclones has shown that all three of the dehalogenases present in the *Rhizobium* sp have been cloned and their genes expressed in *E.coli*.

Chapter 5

Nucleotide sequencing of the *hadD* and *hadL* genes

5.1 Nucleotide sequencing of the *hadD* and *hadL* genes

One of the original aims of the project was to use the nucleotide sequence of the stereospecific dehalogenases, HadD and HadL, to try to determine if small scale changes in conserved regions were responsible for the stereospecificity of the enzymes. The nucleotide sequence could also be used to show the evolutionary relationship of the dehalogenases. In order to investigate these questions it was necessary to have fulfilled two aims. Firstly, the genes to be sequenced needed to be cloned and secondly one must be able to recognise the sequence of the genes of interest within the cloned DNA. A clone had been obtained which was known to contain the complete structural genes for HadD and HadL, (pSC2), and the N-terminal amino acid sequence had been obtained for both of the dehalogenases, allowing the identification of their open reading frames.

Nucleotide sequencing was carried out on the pSC2 plasmid. It was known that the genes were transcribed in the direction of the *lac* promoter of pUC and that the first gene to be transcribed was the *hadD* gene. Use of the Reverse Primer, which binds to the pUC polylinker and directs sequencing in the same direction as the transcription of the *lac* promoter, would allow the start of the *hadD* structural gene to be located. Plasmid sequencing was carried out and a region of DNA was identified where there was an open reading frame identical to that predicted by the N-terminal amino acid sequence. This started 143 nucleotides from the cloning site, at map position 3.3 (Figure 3.4). Sequencing of the gene was then continued using custom synthesised oligonucleotide primers. These primers were based upon sequence obtained from previous sequencing runs and by manipulation of the sequence it was possible to design primers that allowed the sequencing of both strands and in both directions. Sequencing of the insert using the Universal Primer was also carried out, giving sequence on

the opposite strand and in the opposite direction starting at map position 1.5. The use of the same sequencing strategy involving oligonucleotide primers was undertaken for sequence obtained from the Universal Primer. In this way it was possible to sequence all of the DNA from map positions 1.5 to 3.5 on both strands and so data from one strand would be corroborated with data from the other.

5.1.1 Nucleotide sequence of the *hadD* gene

The complete nucleotide sequence of the *hadD* gene is shown in Figure 5.1. There was only one open reading frame (ORF) of any significance. The correct reading frame of the *hadD* gene consisted of 798 bp which encoded for a 266 amino acid protein with a predicted subunit molecular weight of 29,386 Da, which corresponded well to the predicted molecular weight of the subunit of approximately 29 kDa for the purified protein. There was a good ribosome binding site (RBS) (Shine and Dalgarno, 1975) centred at -14 bp upstream of the initiation codon. Although this upstream distance was greater than that usually seen for *E.coli* genes, it corresponds to the distance seen for a *Rhizobium* protein that has been expressed in *E.coli* (De Maagd *et al*, 1992) and so this may be a feature of *Rhizobium* genes. The percentage G+C content for the *hadD* gene was 56%, which was slightly less than the reported 59-64 % (Kreig 1984) predicted for a *Rhizobium* gene. The codon usage was also comparable to that reported for *Rhizobium* sp (Brown 1991). The codon usage is shown in Table 5.1. The predicted amino acid composition of the HadD protein is shown in Table 5.2.

5.1.2 Nucleotide sequence of the *hadL* gene.

The complete nucleotide sequence of the *hadL* gene is shown in Figure 5.2. The correct reading frame of the *hadL* gene started 177 bp downstream of the

Figure 5.1

```
ATG ATA GAT CTT CCG AGG CAT CCC CCA TCG ATG CTG CCT GTC ATT
Met ile asp leu pro arg his pro pro ser met leu pro val ile

CGA ACT GTC CCG GAG CAT GCC GCC ACA GGA GAG TTA AAG AGA CGC
arg thr val pro glu his ala ala thr gly glu leu lys arg arg

TAC GAT GCG GTA AAG TCG GCG TTC GAC GTT CCG TGG ATG GGG GTC
tyr asp ala val lys ser ala phe asp val pro trp met gly val

GTC GCA TGG CTC ATA CCA ATA TCG CGC TTC TTC GAT GCT CTG TGG
val ala trp leu ile pro ile ser arg phe phe asp ala leu trp

GAG GGG CTT GAG CGA GTC GCA GGT ACA CGG GCA TTC CAA GAT GCT
glu gly leu glu arg val ala gly thr arg ala phe gln asp ala

TGC CGC GCA ATG CGA GCG GCA ACG GAA GCA GGC GTA GAA CGA AGC
cys arg ala met arg ala ala thr glu ala gly val glu arg ser

CTC GGG ATC TCG CCG TTG CGC ACC GCC TAC AAG ACC TCG GTT ACG
leu gly ile ser pro leu arg thr ala tyr lys thr ser val thr

ATC CGC GCG AGA TCG GGG AGA TCC GAA CGA TCA TCG AGG TCT TCT
ile arg ala arg ser gly arg ser glu arg ser ser arg ser ser

CGC ATG GCA ACT ATC ATA CAT CTG CTG GCA ACC GTC AGC CGC TAT
arg met ala thr ile ile his leu leu ala thr val ser arg tyr

CTC TTG GAG GAC TTC AGA GGA ACA CAA GCA ACT ATC ATA CAT CTG
leu leu glu asp phe arg gly thr gln ala thr ile ile his leu

CTG GCA AAC CGT CAG CCG CTA TCT CTC TCG GGA GGG GAC TTC AGA
leu ala asn arg gln pro leu ser leu ser gly gly asp phe arg

GGA ACC ACA AGT TTT CGA GAC ATC CCC CCG GCT TCC CCG CAC ATC
gly thr thr ser phe arg asp ile pro pro ala ser pro his ile
```

```
TTT CAT CAA CCG ATC TTT GAT GGA GCC GCA CCA TCT GAC GAA CAC
phe his gln pro ile phe asp gly ala ala pro ser asp glu his

ACG CGC GGA ATC TTT GCA GAT ATT CAG GCA CAC TGG CTG CCG ATC
thr arg gly ile phe ala asp ile gln ala his trp leu pro ile

CTA ATA CTG ACT ATC GAG CGC TCA CGT GGC CAA GCT ACT TCA CCT
leu ile leu thr ile glu arg ser arg gly gln ala thr ser pro

GGC GTG GGC CAG GTT AGA CAG TCC ATC GAT CGC CGA CCC TGT CAC
gly val gly gln val arg gln ser ile asp arg arg pro cys his

AGC CTA ACT GCA CTG AGT CAA TGT CAA GCT CGC TGT TGC CTG CGA
ser leu thr ala leu ser gln cys gln ala arg cys cys leu arg

TAT CTA CTG TCC GAA TCC AGT CTA CTG CCA TAG
tyr leu leu ser glu ser ser leu leu pro AMB
```

Figure 5.1

Nucleotide sequence of *hadD* and the corresponding predicted amino acid sequence.

Table 5.1

Codon usage of the *Rhizobium hadD* gene. Each codon is shown with the amino acid encoded and the number of occurances of the codon in the gene.

TTT phe 4	**TCT** ser 4	**TAT** tyr 2	**TGT** cys 3
TTC phe 6	**TCC** ser 5	**TAC** tyr 2	**TGC** cys 2
TTA leu 1	**TCA** ser 3	**TAA** och 0	**TGA** opa 0
TTG leu 2	**TCG** ser 8	**TAG** amb 1	**TGG** trp 4

CTT leu 2	**CCT** pro 2	**CAT** his 5	**CGT** arg 2
CTC leu 4	**CCC** pro 3	**CAC** his 4	**CGC** arg 11
CTA leu 5	**CCA** pro 4	**CAA** gln 6	**CGA** arg 8
CTG leu 12	**CCG** pro 9	**CAG** gln 4	**CGG** arg 1

ATT ile 2	**ACT** thr 6	**AAT** asn 0	**AGT** ser 3
ATC ile 11	**ACC** thr 4	**AAC** asn 1	**AGC** ser 3
ATA ile 6	**ACA** thr 4	**AAA** lys 0	**AGA** arg 6
ATG met 5	**ACG** thr 3	**AAG** lys 3	**AGG** arg 2

GTT val 3	**GCT** ala 5	**GAT** asp 7	**GGT** gly 1
GTC val 6	**GCC** ala 4	**GAC** asp 5	**GGC** gly 4
GTA val 2	**GCA** ala 14	**GAA** glu 5	**GGA** gly 6
GTG val 1	**GCG** ala 4	**GAG** glu 6	**GGG** gly 5

Table 5.2

Predicited amino acid composition of HadD from the nucleotide sequence of the *hadD* gene.

AMINO ACID	RESIDUES PER SUBUNIT
ALA	27
ARG	30
ASN	1
ASP	12
CYS	5
GLN	10
GLU	11
GLY	16
HIS	9
ILE	19
LEU	26
LYS	3
MET	5
PHE	10
PRO	18
SER	26
THR	17
TRP	4
TYR	4
VAL	12
CALCULATED MOL. WT.	**29386 Da**

stop codon of the *hadD* gene. This was much greater than the 23 bp distance between the stop codon of the *Pseudomonas putida* AJ1 *hadD* gene and the start of its *hadL* gene. However, as reported in Chapter 4, in the *Rhizobium* there was a difference seen in the level of expression between the HadL and HadD dehalogenases and so the longer intergene distance may carry a region effecting the levels of expression of the *hadL* gene, although this region seems not to function in *E.coli* as there was a 1:1 ratio of expression seen for the HadD and HadL proteins. The *hadL* gene consisted of 840 nucleotides including the initiating ATG codon. This gene coded for a 279 amino acid protein, (the initiating amino acid is absent from the mature protein), with the HadL protein having a predicted subunit molecular weight of 30,868 Da, which closely matched the estimated subunit molecular weight of 31kDa for the purified protein. There was a RBS centred at -16 upstream from the initiation codon, again a greater distance than that normally seen for *E.coli* but in close agreement to that seen for the *hadD* gene. The percentage G+C content of the *hadL* gene was 55% which was very similar to that of the *hadD* gene and the codon usage showed similarity to that of the *hadD* gene and to that reported for *Rhizobium* sp shown in Table 5.3 (Brown 1991). The predicted amino acid composition is shown in Table 5.4.

As was stated in Chapter 3, an attempt was made to amplify a portion of the *hadL* gene using PCR with primers designed from L-2CP specific dehalogenases. From the sequence now available for *hadL* it is possible to see that the regions from which the primers were designed were not present in the gene, so the PCR method would not have been succesful with the primers used in this study. Neither the *hadD* or the *hadL* gene had upstream sequences that showed similarity to the -10/-35 type promoter.

Figure 5.2

```
ATG AGT TTA AAA AAA AGA ATC AAG GCA CTA ACG TTT GAC ACG GGC
Met ser leu lys lys arg ile lys ala leu thr phe asp thr gly

GGC ACG GTT GCT CGA TTG GCA GTA CCG GCT TCC GAG ATG CTT TTG
gly thr val ala arg leu ala val pro ala ser glu met leu leu

AGA CGG CAG GCC GCC GGC ACG GAA TCA ATC GAG ACT GGG CGG TAC
arg arg gln ala ala gly thr glu ser ile glu thr gly arg tyr

TGG CCA ATG AAC TGC CGC CGT AGG TCA ATG CAG GCA ATG CTG AAC
trp pro met asn cys arg arg arg ser met gln ala met leu asn

CTC GGT CGG GAG CCC CCC CGT CAT ACA ACT TTG ATG GTG CGT CAC
leu gly arg glu pro pro arg his thr thr leu met val arg his

CAA TTC TCG CTT GAT GCG ATC TTA GCC GAG GAG GGA CTG GAT GTT
gln phe ser leu asp ala ile leu ala glu glu gly leu asp val

TTC GAC GAT GAG GAT CGT GCA CAT TGC TGG GAT GCA CCT CAC AGT
phe asp asp glu asp arg ala his cys trp asp ala pro his ser

TTC GAT CCT GGC GAT GTC CGG GAT GGT CTG GCA AGA CTA CGA GAC
phe asp pro gly asp val arg asp gly leu ala arg leu arg asp

CGA TAT ATC GCT GTG TCG TTC ACT TTC GTC TCG CAT CGG CTC ATC
arg tyr ile ala val ser phe thr phe val ser his arg leu ile

ATA GAC ACG ACG TCA GTC GTA ACC GGC CTG ATG TGG ATG CGA TCC
ile asp thr thr ser val val thr gly leu met trp met arg ser

TGT CTT GTG AGG GAA TGG GTG TCT ACA AGC CAT TGC CAG CAG ATA
cys leu val arg glu trp val ser thr ser his cys gln gln ile

TGC GAA AGC GGC GGC TAT GCT TCA CGT AAA GCC CTG AGG AAT GCC
cys glu ser gly gly tyr ala ser arg lys ala leu arg asn ala
```

```
TTA TGG TCG CAT GCC ATC GTT TCG ATC TTG ATG CAG CGC GAA ACC
leu trp ser his ala ile val ser ile leu met gln arg glu thr

TCG GCT TCA GGA CAG CCA CTA ATA AAT AGG CCG GAT GAG TGG GGC
ser ala ser gly gln pro leu ile asn arg pro asp glu trp gly

AAG GCT ATT GGT CCT CAG AAA CCT CCG CCG GGC AGT GAG CCT TAC
lys ala ile gly pro gln lys pro pro pro gly ser glu pro tyr

GAC ATT GAG CTG AAT AGC TTT CTC GAG CTT GCG GCA TTT CTG GAA
asp ile glu leu asn ser phe leu glu leu ala ala phe leu glu

AGC GAA AGC TCG CTA AAA GCA AAC GCC ATC AGC GGT GCG GTG CGA
ser glu ser ser leu lys ala asn ala ile ser gly ala val arg

GGC AGC GCC GCT AGA CCA ACA GCT CCT GGC CGA TAC GCG CTC GTT
gly ser ala ala arg pro thr ala pro gly arg tyr ala leu val

GCG CTG TAC GCC TCG GCG GAA GGC AAG TGA
ala leu tyr ala ser ala glu gly lys OPA
```

Figure 5.2

Nucleotide sequence of *hadL* and the corresponding predicted amino acid sequence. Initiating methionine is removed to give mature enzyme.

Table 5.3

Codon usage of the *Rhizobium hadL* gene. Each codon is shown with the amino acid encoded and the number of occurances of the codon in the gene.

TTT phe 3	**TCT** ser 1	**TAT** tyr 2	**TGT** cys 1
TTC phe 5	**TCC** ser 2	**TAC** tyr 4	**TGC** cys 4
TTA leu 3	**TCA** ser 5	**TAA** och 0	**TGA** opa 1
TTG leu 4	**TCG** ser 8	**TAG** amb 0	**TGG** trp 6

CTT leu 4	**CCT** pro 6	**CAT** his 5	**CGT** arg 5
CTC leu 4	**CCC** pro 2	**CAC** his 2	**CGC** arg 2
CTA leu 4	**CCA** pro 3	**CAA** gln 1	**CGA** arg 6
CTG leu 8	**CCG** pro 4	**CAG** gln 7	**CGG** arg 5

ATT ile 2	**ACT** thr 3	**AAT** asn 3	**AGT** ser 3
ATC ile 8	**ACC** thr 2	**AAC** asn 3	**AGC** ser 7
ATA ile 3	**ACA** thr 3	**AAA** lys 5	**AGA** arg 4
ATG met 9	**ACG** thr 6	**AAG** lys 3	**AGG** arg 4

GTT val 4	**GCT** ala 8	**GAT** asp 9	**GGT** gly 4
GTC val 3	**GCC** ala 9	**GAC** asp 5	**GGC** gly 12
GTA val 2	**GCA** ala 8	**GAA** glu 7	**GGA** gly 2
GTG val 5	**GCG** ala 6	**GAG** glu 10	**GGG** gly 1

Table 5.4

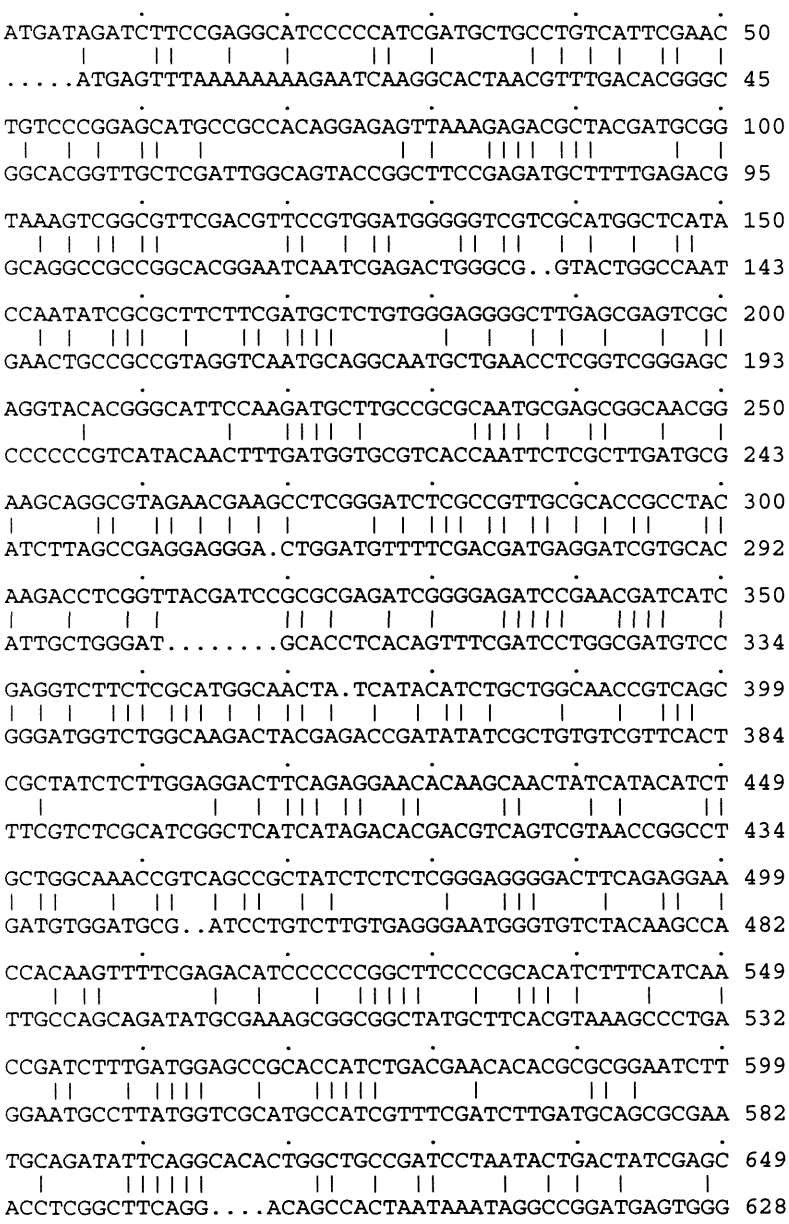
Predicited amino acid composition of HadL from the nucleotide sequence of the *hadL* gene.

AMINO ACID	RESIDUES PER SUBUNIT
ALA	31
ARG	26
ASN	6
ASP	14
CYS	5
GLN	8
GLU	17
GLY	19
HIS	7
ILE	13
LEU	27
LYS	8
MET	9
PHE	8
PRO	15
SER	26
THR	14
TRP	6
TYR	6
VAL	14
CALCULATED MOL. WT.	**30868 Da**

5.1.3 Comparison of the nucleotide sequence of *hadD* and *hadL* with published dehalogenase sequences

Using the GAP algorithm it was possible to compare the nucleotide sequences of *hadD /hadL* and the protein sequences of HadD/HadL with known sequences for other dehalogenase genes and proteins. Comparison of the two genes to each other showed that there was 40% identity at the nucleotide level between the *hadD* and *hadL* genes, but at the protein level there was only 18% identity. These comparisons are shown in Figure 5.3 and Figure 5.4. This is intriguing, as the similarity at the DNA level seems significant, but at the amino acid level the identity was spread throughout the sequence, with no large regions of similarity in evidence. This situation compares to that seen for the *P.putida* AJ1 *hadD* and *hadL* genes and their corresponding predicted protein sequences (Barth, 1992). Comparison of the two dehalogenase genes from this organism showed that at the nucleotide level there was again 41% identity, but when the predicted protein sequences were aligned, the identity was 20%, analagous to the situation seen in the *Rhizobium*. This was interpreted to mean that the HadD and HadL dehalogenases from the *Pseudomonas* strain were not related (Jones *et al* 1992). When the Rhizobial *hadD* gene and the predicited amino acid sequence was compared to the *hadD* gene and predicted amino acid sequence from *Pseudomonas putida* AJ1 there was 41% identity at the nucleotide level and a 23% identity between the two amino acid sequences, which was comparable to the level seen between the dehalogenase genes from the *Rhizobium*. There was only one region where there was stretch of conserved amino acids, with 70% identity over the 13 amino acids. This was located in the N-terminal region of the dehalogenases. The alignment of the D-2CP specific dehalogenase genes from the *Rhizobium* and the

Figure 5.3


ATGATAGATCTTCCGAGGCATCCCCCATCGATGCTGCCTGTCATTCGAAC 50
.....ATGAGTTTAAAAAAAAGAATCAAGGCACTAACGTTTGACACGGGC 45
TGTCCCGGAGCATGCCGCCACAGGAGAGTTAAAGAGACGCTACGATGCGG 100
GGCACGGTTGCTCGATTGGCAGTACCGGCTTCCGAGATGCTTTTGAGACG 95
TAAAGTCGGCGTTCGACGTTCCGTGGATGGGGGTCGTCGCATGGCTCATA 150
GCAGGCCGCCGGCACGGAATCAATCGAGACTGGGCG..GTACTGGCCAAT 143
CCAATATCGCGCTTCTTCGATGCTCTGTGGGAGGGGCTTGAGCGAGTCGC 200
GAACTGCCGCCGTAGGTCAATGCAGGCAATGCTGAACCTCGGTCGGGAGC 193
AGGTACACGGGCATTCCAAGATGCTTGCCGCGCAATGCGAGCGGCAACGG 250
CCCCCCGTCATACAACTTTGATGGTGCGTCACCAATTCTCGCTTGATGCG 243
AAGCAGGCGTAGAACGAAGCCTCGGGATCTCGCCGTTGCGCACCGCCTAC 300
ATCTTAGCCGAGGAGGGA.CTGGATGTTTTCGACGATGAGGATCGTGCAC 292
AAGACCTCGGTTACGATCCGCGCGAGATCGGGGAGATCCGAACGATCATC 350
ATTGCTGGGAT.........GCACCTCACAGTTTCGATCCTGGCGATGTCC 334
GAGGTCTTCTCGCATGGCAACTA.TCATACATCTGCTGGCAACCGTCAGC 399
GGGATGGTCTGGCAAGACTACGAGACCGATATATCGCTGTGTCGTTCACT 384
CGCTATCTCTTGGAGGACTTCAGAGGAACACAAGCAACTATCATACATCT 449
TTCGTCTCGCATCGGCTCATCATAGACACGACGTCAGTCGTAACCGGCCT 434
GCTGGCAAACCGTCAGCCGCTATCTCTCTCGGGAGGGGACTTCAGAGGAA 499
GATGTGGATGCG..ATCCTGTCTTGTGAGGGAATGGGTGTCTACAAGCCA 482
CCACAAGTTTTCGAGACATCCCCCCGGCTTCCCCGCACATCTTTCATCAA 549
TTGCCAGCAGATATGCGAAAGCGGCGGCTATGCTTCACGTAAAGCCCTGA 532
CCGATCTTTGATGGAGCCGCACCATCTGACGAACACACGCGCGGAATCTT 599
GGAATGCCTTATGGTCGCATGCCATCGTTTCGATCTTGATGCAGCGCGAA 582
TGCAGATATTCAGGCACACTGGCTGCCGATCCTAATACTGACTATCGAGC 649
ACCTCGGCTTCAGG....ACAGCCACTAATAAATAGGCCGGATGAGTGGG 628

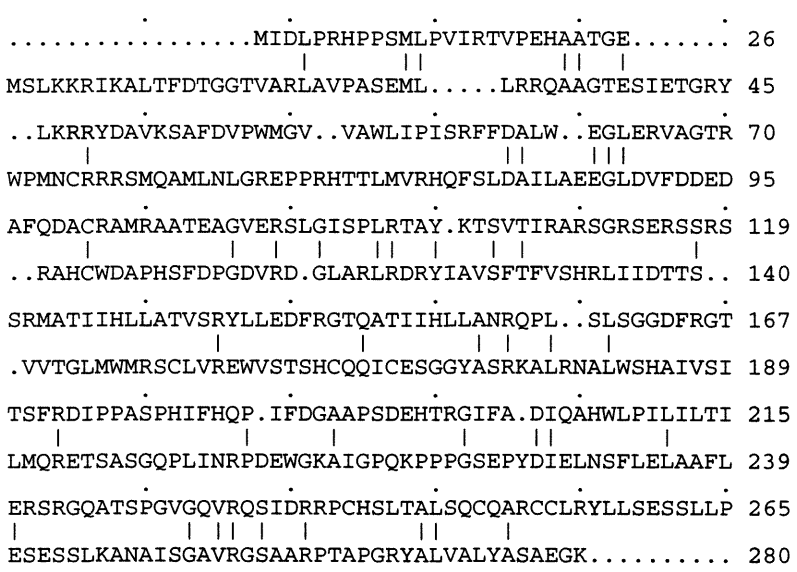
```
GCTCACGTGGCCAAGCTACTTCACCTGGCGTGGGCCAGGTTAGACAGTCC 699
||      |        ||      ||||      ||||   |       |   |
GCAAGGCTATTGGTCCTCAGAAACCTCCGCCGGGCAGTGAGCCTTACGAC 678

ATCGATCGCCGACCCTGTCACAGCCTAACTGCA.CTGAGTCAATGTCAAG 748
||  ||  |        ||  ||  |   ||   | |||  |  |  ||    |||
ATTGAGCTGAATAGCTTTCTCGAGCTTGCGGCATTTCTGGAAAGCGAAAG 728

CTCGCT.GTTGCCTGCGATATCTACTGTCCGAATCCAGTCTACTGCCATA 797
||||||    ||   ||  |||  | ||  ||   | ||  |  |  |  ||
CTCGCTAAAAGCAAACGCCATCAGCGGTGCGGTGCGAGGCAGCGCCGCTA 778

G.................................................. 798
|
GACCAACAGCTCCTGGCCGATACGCGCTCGTTGCGCTGTACGCCTCGGCG 828
```

Identity: 39.69%

Figure 5.3

Comparison of the nucleotide sequences of *hadD* and *hadL*. *hadD* sequence is the upper sequence. Vertical lines represent identical nucleotides.

Figure 5.4



Identity: 17.95%

Figure 5.4

Comparison of the predicted protein sequences for the HadD and HadL dehalogenases from the *Rhizobium*. The HadD sequence is uppermost. Vertical bars represent identical amino acids.

Pseudomonas and their predicted protein products is shown in Figure 5.5 and Figure 5.6. Whether this region is conserved amongst D-2CP specific dehalogenases can not be answered at present as there have been only two of these dehalogenases reported and so will only be answered by the discovery of further D-2CP specific dehalogenases and the sequencing of their genes. Comparison of the Rhizobial HadL protein with the amino acid sequence of the HadL protein from *Pseudomonas putida* AJ1 (Jones, 1992) showed that there was again a 40% identity at the nucleotide level and a 20% identity at the amino acid level between the two L-2CP specific dehalogenases. This was in contrast to the findings of Jones (1992), where it was seen that the amino acid sequence of HadL was significantly similar to that of two L-2CP dehalogenases of *Pseudomonas* sp. strain CBS3 (Schneider, 1991) and also showned good similarity to a L-2CP specific Hdl IVa dehalogenase from *Ps. cepacia* MBA4 (Murdiyatmo *et al* 1992) There were three regions of particularly high similarity, with the overall similarity being 51% to DehcII, 38% to DehcI and 36% to Hdl IVa. These regions of similarity are shown in Figure 5.7. In these regions where there were high levels of similarity, the *Rhizobium* HadL showed little, if any similarity. The alignment of the Rhizobial *hadL* and the *Pseudomonas hadL* genes and their predicted protein products is shown in Figure 5.8 and Figure 5.9. This may indicate that there are several subgroups within the dehalogenase enzyme family and the *Rhizobium* HadL represents a different group to that of the other L-2CP specific dehalogenase enzymes.

With the complete nucleotide sequences of the *Rhizobium hadD* and *hadL* genes it was possible to carry out searches to determine if there was any similarity between these sequences and sequences residing in the databases. Using the FASTA software available on the IRIX system at the

Figure 5.5

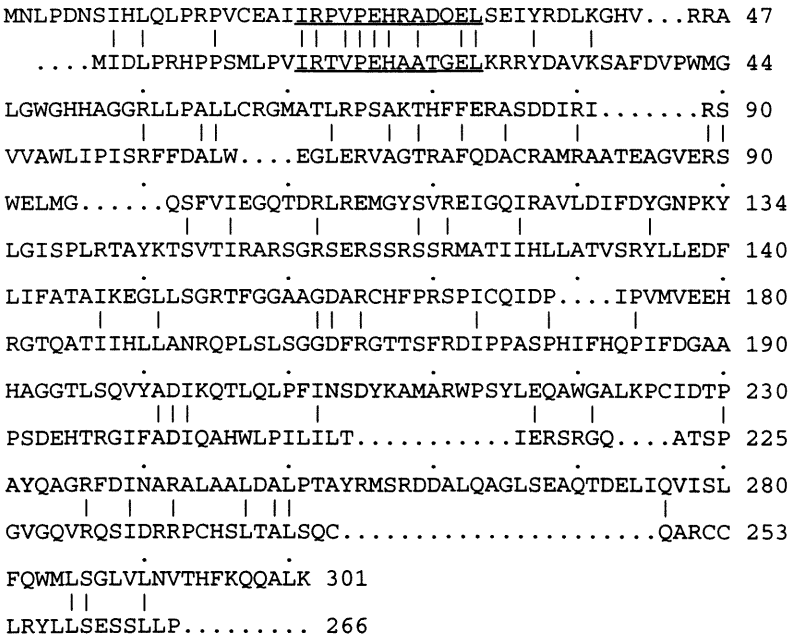

.....ATGATAGATCTTCCGAGGCATCCCCCATCGAT 32
ATGAACCTGCCTGACAATTCCATCCATCTTCAACTGCCGCGGCCCGTCTG 50
GCTGCCTGTCATTCGAACTGTCCCGGAGCATGCCGCCACAGGAGAGTTAA 82
CGAGGCAATCATCCGCCCGGTTCCGGAACACCGTGCGGATCAGGAACTGA 100
AGAGACGCTACGATGCGGT.AAAGTCGGCGTTCGACGTTCCGTGGATGGG 131
GCGAAATCTACCGGGATTTGAAGGGCCACGTTCGGCGTGCCCTGGGTTGG 150
GGTCGTCGCATGGCTCATACCAATATCGCGCTTCTTCGATGCTCTG..TG 179
GGTCATCACGCAGGCGGTCGCCTACTACCGGCCCTTCTTTGCAGAGGCAT 200
GGAGGGGCTTGAGCGAGTCGCAGGTACACGGGCATTCCAAGATGCTTGCC 229
GGCGACGCTTCGGCCGTCGGCGAAAACGCATTTTTTCGAGCGTGCCAGTG 250
GCGCAATGCGAGCGGCAACGGAAGCAGGCGTAGAACGAAGCCTCGGGATC 279
ATGACATACGGATCCGGTCCTGGGAGCTCATGGGACAGTCCTTTGTCATC 300
.TCGCCGTTGCGCACCGCCTACAAGACCTCGGTTACGATCCGCGCGAGAT 328
GAGGGCCAGACAGACCGGCTACGGGAGATGGGTTATTCGGTGCGTGAAAT 350
CGGGGAGATCCGAACGATCATCGAGGTCTTCTCGCATGGCAACTATCATA 378
CGGGCAGATCCGGGCAGTGCTGGACATCTTCGATTACGGCAA.TCCGAAA 399
CATCTGCTGGCAACCGTCAGCCGCTATCTCTTGGAGGACTTCAGAGGAAC 428
TATCTGATTTTCGCCACTGCCATCAAGGAAGGCCTGCTGAGCGGCCGCAC 449
ACAAGCAACTATCATACATCTGCTGGCAAACCGTCAGCCGCTATCTCTCT 478
GTTCGGCGGCGCGGCAGGCGATGCGCGATGTCATTTCCCGCGATCGCCGA 499
CGGGAGGGGACTTCAGAGGAACCACAAGTTTTCGAGACATCCCCCCGGCT 528
TCTGCCAGATCGACCCGATTCCGGTAATGGTTGAGGAACACCACGCGGGC 549
TCCCCGCACATCTTTCATCAACCGATCTTTGATGGAGCCGCACCATCTGA 578
GGCACCCTC.TCGCAGGTTTATGCCGATATCAAGCAGACCCTGCAATTGC 598

Identity: 41.25%

Figure 5.5

Comparison of the nucleotide sequences of the *Rhizobium hadD* gene and the *Pseudomonas putida* AJ1 *hadD* gene. The *Rhizobium* gene is uppermost. Vertical bars represent identical nucleotides.

Figure 5.6



Identity: 23.17%

Figure 5.6

Comparison of predicted amino acid sequences of the *Rhizobium* HadD and the *Pseudomonas putida* AJ1 HadD. The *Pseudomonas* HadD is uppermost. Vertical bars represent identical residues. Underlined region represents largest area of similarity.

Figure 5.7

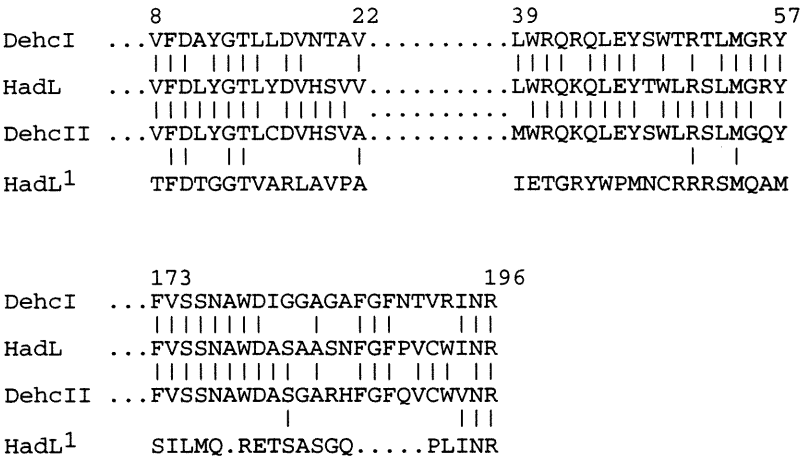
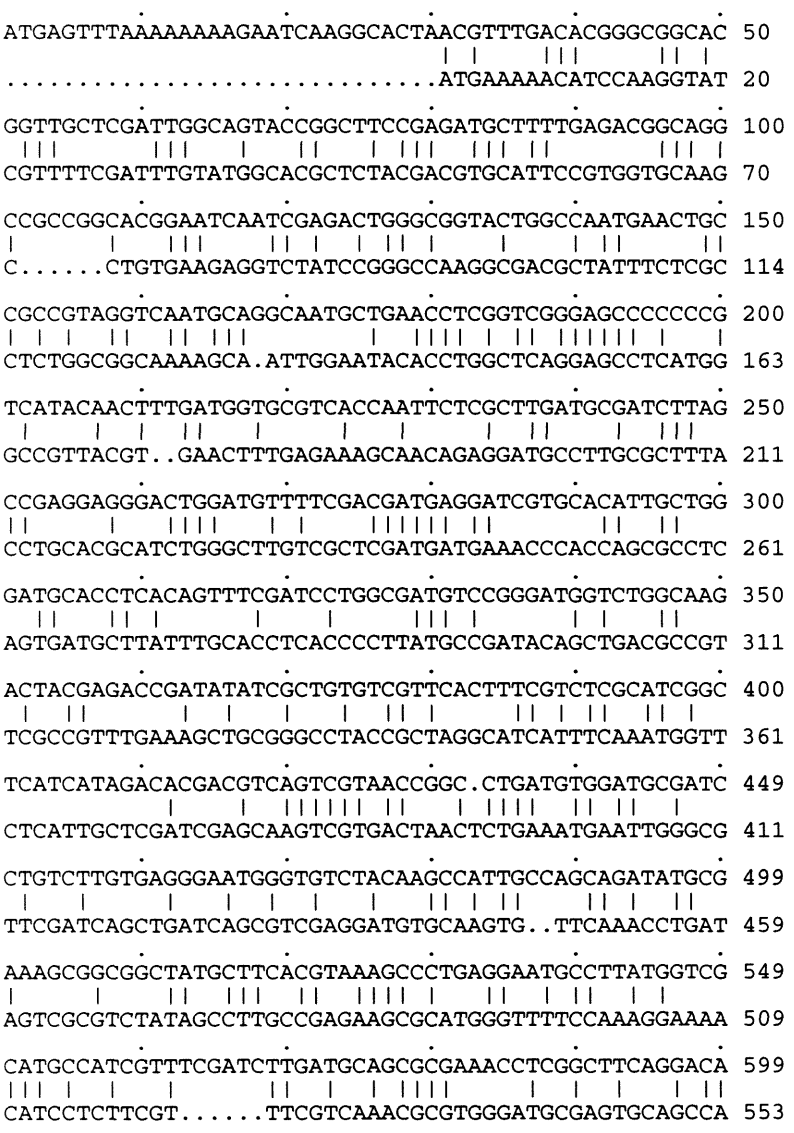


Figure 5.7

Amino acid sequence showing regions of homolgy between HadL from *Pseudomonas putida* AJ1, DehcI and DehcII dehalogenases from *Pseudomonas* sp. strain CBS3 and [1]HadL from the *Rhizobium*. Numbers above the amino acids indicate residue number, vertical bars represent identical residues.

Figure 5.8

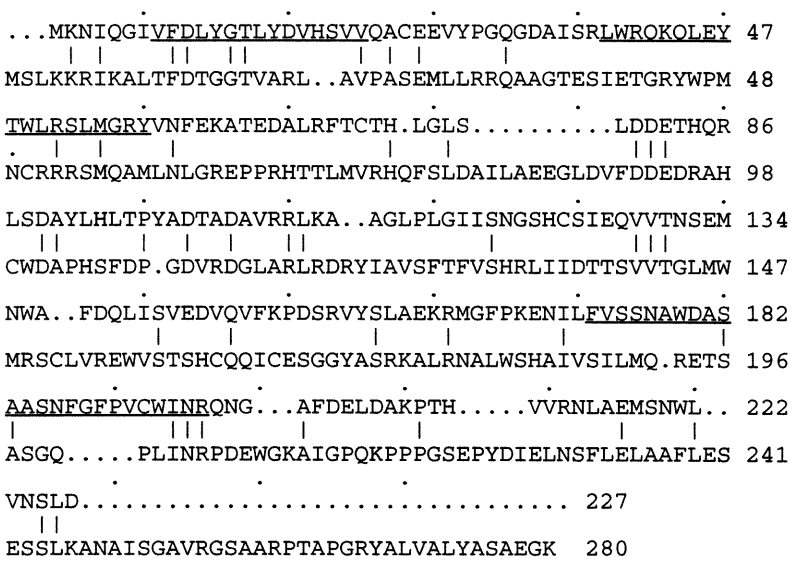

ATGAGTTTAAAAAAAAGAATCAAGGCACTAACGTTTGACACGGGCGGCAC 50
ATGAAAAACATCCAAGGTAT 20
GGTTGCTCGATTGGCAGTACCGGCTTCCGAGATGCTTTTGAGACGGCAGG 100
CGTTTTCGATTTGTATGGCACGCTCTACGACGTGCATTCCGTGGTGCAAG 70
CCGCCGGCACGGAATCAATCGAGACTGGGCGGTACTGGCCAATGAACTGC 150
C......CTGTGAAGAGGTCTATCCGGGCCAAGGCGACGCTATTTCTCGC 114
CGCCGTAGGTCAATGCAGGCAATGCTGAACCTCGGTCGGGAGCCCCCCCG 200
CTCTGGCGGCAAAAGCA.ATTGGAATACACCTGGCTCAGGAGCCTCATGG 163
TCATACAACTTTGATGGTGCGTCACCAATTCTCGCTTGATGCGATCTTAG 250
GCCGTTACGT..GAACTTTGAGAAAGCAACAGAGGATGCCTTGCGCTTTA 211
CCGAGGAGGGACTGGATGTTTTCGACGATGAGGATCGTGCACATTGCTGG 300
CCTGCACGCATCTGGGCTTGTCGCTCGATGATGAAACCCACCAGCGCCTC 261
GATGCACCTCACAGTTTCGATCCTGGCGATGTCCGGGATGGTCTGGCAAG 350
AGTGATGCTTATTTGCACCTCACCCCTTATGCCGATACAGCTGACGCCGT 311
ACTACGAGACCGATATATCGCTGTGTCGTTCACTTTCGTCTCGCATCGGC 400
TCGCCGTTTGAAAGCTGCGGGCCTACCGCTAGGCATCATTTCAAATGGTT 361
TCATCATAGACACGACGTCAGTCGTAACCGGC.CTGATGTGGATGCGATC 449
CTCATTGCTCGATCGAGCAAGTCGTGACTAACTCTGAAATGAATTGGGCG 411
CTGTCTTGTGAGGGAATGGGTGTCTACAAGCCATTGCCAGCAGATATGCG 499
TTCGATCAGCTGATCAGCGTCGAGGATGTGCAAGTG..TTCAAACCTGAT 459
AAAGCGGCGGCTATGCTTCACGTAAAGCCCTGAGGAATGCCTTATGGTCG 549
AGTCGCGTCTATAGCCTTGCCGAGAAGCGCATGGGTTTTCCAAAGGAAAA 509
CATGCCATCGTTTCGATCTTGATGCAGCGCGAAACCTCGGCTTCAGGACA 599
CATCCTCTTCGT......TTCGTCAAACGCGTGGGATGCGAGTGCAGCCA 553

Identity: 40.41%

Figure 5.8

Comparison of the nucleotide sequences of the *Rhizobium hadL* gene and the *Pseudomonas putida* AJ1 *hadL* gene. The *Rhizobium* gene is uppermost. Vertical bars represent identical nucleotides.

Figure 5.9



Identity: 20.64%

Figure 5.9

Comparison of the predicted protein sequence for the *Rhizobium*HadL and the *Pseudomonas putida* AJ1 HadL. The *Pseudomonas* protein is uppermost. Vertical bars represent identical residues. Underlined regions represent regions of similarity seen for L-2CP specific dehalogenases (See Figure 5.7).

University of Leicester both nucleotide and protein sequence databases, (EMBL release no. 36 and SWISSPROT release no. 26) respectively, were searched. No significant similarities were found between the *Rhizobium* dehalogenases and the databases.

5.2 Summary

The nucleotide sequence for the *hadD* and *hadL* genes was obtained in full and the two genes and their predicted protein products compared. The genes showed some similarity at the DNA level but there was less similarity at the protein level, indicating that the two genes may not be related, and the different stereospecific activities of the two proteins did not arise by a simple mutation of one gene to form the other. This was analagous to the situation seen for the equivalent dehalogenases from *Pseudomonas putida* AJ1 where the genes were reported as being unrelated (Jones *et al* 1992). Comparison of the two genes with genes encoding similar dehalogenase enzymes show that there was also little similarity to other dehalogenases. Comparison of the two HadD dehalogenases so far reported showed only one region where there was significant similarity and this was located in the N-terminal region of the protein. Comparison of the Rhizobial HadL dehalogenase showed no regions of significant similarity, compared to the three highly similar regions known to exist between L-2CP specific dehalogenases, which have now been seen in the comparison of five L-2CP specific dehalogenases.

This information determined from the nucleotide sequence would suggest that the Rhizobial *hadD* and *hadL* genes were completely distinct, being related neither to each other nor to other known dehalogenase genes.

Chapter 6

Discussion and Future Work

6.1 Discussion and future work

The original aims of this project were to clone the genes responsible for the stereospecific dehalogenases expressed by the *Rhizobium* and to study these genes at the molecular level. By doing this any differences between the genes that are perhaps responsible for the stereospecific activities of the dehalogenases may be highlighted. By comparing the genes and their protein products to each other and to published dehalogenase genes and enzymes, the evolutionary relatedness could be established. This may help explain the presence of multiple dehalogenases in the same organism.

6.2 Growth rates of cells carrying dehalogenase clones

The growth rate of host cells on 2CP carrying any dehalogenase seems to be limited by the dehalogenase specific activity until a certain value is reached, then the rate limiting step is switched to the uptake of 2CP. This was clearly demonstrated during the subcloning of the plasmid pSC1. When subclones were produced that had 3x greater dehalogenase activities, there was no change in the growth rate of the cells, implying that a factor other than dehalogenase activity limited growth. As the doubling time of *E.coli* grown on lactate was less than that seen for growth on 2CP then it seemed unlikely that the metabolism of the product of 2CP dehalogenation, lactate, was now rate limiting and so the only remaining factor was uptake. The plasmid pSC2 had only enough DNA to encode the dehalogenase structural genes, the growth rate seen for these cells had to be that for a system with no cloned uptake mechanism. Therefore an uptake system for 2CP was not believed to be cloned, on the basis of the small difference in the growth rate between pSC1 containing cells and pSC2 containing cells. If the genes required for uptake had been cloned and expressed on pSC1,

there would be enhanced uptake and therefore quicker growth. The growth of cells carrying pSC1 was slightly slower than cells carrying pSC2.

The actual mechanism for HAA uptake has not been described, however the effect of HAA uptake on the cell and its modulation have been investigated. The toxic effects of MCA and DCA were investigated by Weightman and co-workers (1985) and some of the MCA^r and DCA^r mutant strains that they isolated were attributed to a reduction in the uptake rate of HAA's into the cell. It was postulated that accumulation of HAA's intracellularly led to toxic effects and therefore reduction of the rate of uptake helped reduce the intracellular concentration. The role of permeases in the protection of the cell was confirmed by the work of Slater *et al.*, (1985). A mutant strain of *P.putida* PP3 that was resistant to MCA and DCA, PP42, was different for the parent strain only in the absence of one of two uptake systems, as measured by uptake of radiolabelled MCA and this loss allowed resistance to HAA's. This suggests that it may not be possible to clone the HAA uptake system and express it in *E.coli*, as such strains may have hypersensitivity to HAA toxicity and so would not have been isolated using the direct selection method. It would be possible to investigate the uptake of 2CP into *E.coli* and determine if uptake was via the lactate or propionate systems, or if another system was responsible. This could be achieved in several ways, depending on the availabilty of radiolabelled compounds. Cells could be grown on a carbon source such as succinate, in the absence and presence of 2CP and uptake of radiolabelled propionate or lactate could be determined. If there was an increase in the level of uptake when cells are grown in the presence of 2CP, this may show that the 2CP was able to induce a particular uptake system and so that system was likely to be involved with uptake of 2CP. An alternative method would be to compete uptake of radiolabelled propionate or lactate with non-radiolabelled 2CP to

see if there was any effect. If radiolabelled 2CP was available, uptake of 2CP could be competed with unlabelled propionate or lactate. Use of radiolabelled DCP in uptake experiments would confirm the findings that *E.coli* cells carrying a DCP dehalogenase clone are unable to grow on DCP because of the absence of an uptake system. It was possible that the *Rhizobium* permease gene was located near to the dehalogenase structural genes, in a similar fashion to the *dehI* associated permease in *P.putida* PP3 (Slater *et al.*, 1985). Use of pSC1 and pSC530 to indentify flanking regions of genomic DNA that may contain the permease gene can be envisaged.

6.3 Regulation of dehalogenase expression

Neither of the clones showed any induction of dehalogenase expression in the presence of the known inducer 2CP. There were two possible reasons for this, either the regulatory region or gene had not been cloned or the regulatory gene had been cloned but did not function in *E.coli*. There was approximately 3.0kbp of DNA upstream of the dehalogenase structural genes on pSC1, so it may be expected that if the regulator gene was located near to the structural genes, then it would be present in this region. If this was the case then the reason for the lack of regulation must be that the regulatory gene was not expressed in *E.coli*. The only example of a dehalogenase regulator gene that has been studied is the $dehR_I$ gene from *Pseudomonas putida* PP3 (Thomas *et al.*, 1992b). This gene was constitutively expressed in *E.coli* and acts as a positive regulator via interactions with the substrates. The regulatory system for the *Rhizobium* dehalogenases is possibly similar, as in the absence of the regulator protein in cells carrying the plasmids there was little dehalogenase activity, so the regulator was required for efficient transcription if the structural gene promoters were used. This could only be confirmed by obtaining regulated clones and seeing how the presence of a regulator protein effected

dehalogenase expression and activity. A possible way to determine if the regulator gene had been cloned and was not being expressed would be to transform the plasmid into either a host that was more closely related to the *Rhizobium*, such as a pseudomonad, or into a mutant strain of *Rhizobium*, that was lacking in dehalogenase activity, such as the mutant strain constructed by Leigh (Leigh 1986). This would involve the cloning of the insert into vectors that would function in organisms other than *E.coli*. Vectors such as pTB244 for *Pseudomonas* sp. or pBIN for the *Rhizobium* would be appropriate. It may then be possible to examine the regulation of dehalogenase expression in more detail. It may also be possible to use the pSC1 insert to purify the dehalogenase regulatory protein from the *Rhizobium*. Upstream of the dehalogenase structural genes would be the sequence that the regulator protein would recognise. Affinity columns could be made that each contain different regions of the upstream DNA from the dehalogenase genes Use of this DNA in affinity chromatography would identify a protein that specifically bound to that region of DNA. Sequence from this protein could be used to make a probe to identify the dehalogenase regulator gene. This process would also allow the identification of the regulatory sequence, as only those that contained the regulatory sequence would bind the regulator protein, therefore the location of the regulatory sequence could be identified and the sequence then characterised.

6.4 Dehalogenase gene organisation

Restriction mapping of the clones showed that they were from distinct regions of the genomic DNA. There was a possibility that the two clones were a contiguous region of genomic DNA that had been split as a result of the cloning. Southern analysis was used to determine the relationship of the clones to each other. This showed that they were not contiguous in the genome and also that they shared little similarity at the DNA level as they

showed no hybridisation under the conditions used. The Southern analysis was not able to show how closely associated the two clones are on the genome. This would be possible by the use of such techniques as chromosome walking. Using the pSC1 insert, a contiguous region of DNA could be identified. This could then be probed with the pSC530 insert to determine if the desired DNA region is present. If not, a new probe could be made from the contiguous region and used to identify a further region of DNA that could be probed for the presence of the pSC530 insert. This could be continued until the region of genomic DNA containing the pSC530 insert had been identified and so the distance between the pSC1 and PSC530 inserts could be established. The Southern analysis also suggests that from the number of hybridising bands seen, the *Rhizobium* contains only one copy of each of the genes.

6.5 Differences in expression levels of HadD and HadL

During the purification of the Rhizobial HadL, a difference in the relative amounts of HadD and HadL could be seen. This reflectec either a preferential purification of the HadL dehalogenase over the HadD enzyme or a difference in levels of expression in the *Rhizobium*. As there was co-purification of the two dehalogenases from cells carrying pSC2, it would seem that the purification method should not be biased towards one dehalogenase. This would imply that the two dehalogenases are expressed at different levels in the *Rhizobium*. Other evidence for the difference in expression was a 2.5x to 3.5x difference in the specific activities of HadD and HadL when measured in *Rhizobium* mutants (Leigh 1986). However, this difference could have been due to factors other than levels of expression, such as k_{cat} or protein stability, which were not addressed by Leigh (1986) and so cannot be discounted.

The model proposed by Leigh (1986) showed only one promoter for both the *hadD* and *hadL* genes. If this was correct, for there to be differential expression other factors need to be considered, such as ribosome binding site affinity or mRNA structure and stability. If investigation into the regulation of pSC1 showed that the clone was regulatable in a different host, it would be possible to study the expression of HadD and HadL from their own promoter and so ascertain if the differences in specific activity were due to differing levels of expression. Combined with this could be the determination of the kinetic properties of HadD and HadL, which would show if it was these that caused the differences in specific activity.

6.6 Dehalogenase activity of cells carrying pSC530

It was not possible to purify the dehalogenase activity from cells containing pSC530 or the subclone pSC520 due to the relatively low levels of expression and the poor separation of the dehalogenase from other cellular proteins. Characterisation of the dehalogenase activity by means of substrate profiling seemed to indicate that there was a very high probability that the dehalogenase cloned on pSC530 was the Dehalogenase II from the *Rhizobium*, as the relative activity profile against a range of substrates was the same as the Rhizobial Dehalogenase II. In terms of future work, it would appear that plasmid pSC530 and the encoded dehalogenase offer the most obvious and interesting continuation of the research. Southern blotting has shown, although the data was not presented here, that the insert of the plasmid pSC520 can be used to identify a *Hin*dIII genomic DNA fragment that was larger than the original pSC530 insert. In comparison with the restriction map of pSC530, the extra DNA must lie upstream of the dehalogenase gene. If this DNA was important for the expression of the dehalogenase then cloning the *Hin*dIII fragment identified in the Southern blot may lead to an insert with better expression of the

dehalogenase. Increased expression of the dehalogenase will allow easier identification of the dehalogenase protein during purification. A scenario similar to that for the HadD and HadL could be envisaged where the N-terminal amino acid sequence would be determined, thus allowing an open-reading frame to be assigned when the structural gene is sequenced. Sequencing of this gene would give the first sequence of a non-stereospecific 2CP dehalogenase and it would therefore be interesting to compare the sequence to those of stereospecific dehalogenases to see if there was any degree of relatedness. With the exception of the *hadL* gene from the *Rhizobium* reported here, the L-isomer specific dehalogenases all seem to have regions of high similarity and if any of these regions are seen in the sequence of the non-specific dehalogenase this may indicate a common link.

Once levels of expression had been achieved that allow the purification of large amounts of the dehalogenases, as was the case for HadD and HadL, research into the mechanisms of the dehalogenase reaction could be envisaged. The production of large amounts of protein facilitates the crystallisation of the enzyme, a process carried out for the haloalkane dehalogenase from *Xanthobacter autotrophicus* (Franken *et al.*, 1991). This would help identify the catalytic residues if the crystal could be formed in the presence of a non-metabolisable substrate analogue. This approach in conjunction with the use of random and site-directed mutagenesis would lead to the identification of the residues important for catalysis. The mutagenesis approach was used successfully by Asmara *et al.*, (1993) where they identified two residues of the HdlIVa dehalogenase as the key for catalysis. The residues were present in the conserved region not seen in either the HadD or HadL dehalogenases from the *Rhizobium*, therefore the key residues for these proteins must differ from those for the HdlIVa dehalogenase. It would be of interest to define the catalytic residues of

HadD and HadL and the non-specific dehalogenase, which would probably have a catalytic cysteine residue as the dehalogenase activity is sensitive to sulphydryl-blocking reagents.

6.6 Nucleotide sequence of *hadD* and *hadL*

The nucleotide sequencing showed that the *hadD* and *hadL* genes from the *Rhizobium* were 40% identical to each other over the length of the gene, but there are no significant stretches where the nucleotide sequence of the genes are conserved. When the predicted amino acid sequences were compared, there was a very low percentage identity (18%) between the two dehalogenases. This means that if the two genes were related, the mutations have occurred not in the third 'wobble' base of the codon, but in the first two bases, so changing the amino acid sequence. It would seem unlikely that one dehalogenase was able to mutate to such an extent as to change over 80% of the amino acid sequence and still retain functionality differing essentially in the stereospecific nature of the reaction. It seems more likely that over a period of time the *Rhizobium* had obtained via a form of genetic transfer genes for both dehalogenases and arranged them on the chromosome in such a way that they are co-ordinately regulated. The sequence of the non-specific dehalogenase would be of interest here. If this sequence showed significant similarity to either the HadD or HadL dehalogenases, the non-specific dehalogenase may have been the ancestral enzyme for a stereospecific dehalogenase. If this was the case then the *Rhizobium* would only have obtained one of the stereospecific dehalogenases by genetic transfer. The organism may have originally evolved a non-stereospecific dehalogenase from which a stereospecific dehalogenase arose, under the control of the same regulator. The organism then obtained a further dehalogenase and via genetic rearrangement this too was put under the control of the same regulator. It is interesting to note

that the HadL dehalogenase from the *Rhizobium* does not share the conserved sequences seen for the five other published L-specific dehalogenases. This would imply that whilst the other L-specific dehalogenases all arose from a common ancestral enzyme, the evolution of the Rhizobial HadL must have been from a different source, possibly the non-specific dehalogenase or a generalised hydrolase. Hardman (1991) suggests that the evolution of dehalogenases may be a result of the modification of hydrolases under the selection pressure applied in the laboratory. This may be possible when a single dehalogenase occurs, but for multiple dehalogenase systems with their complex regulation and from evidence such as the dehalogenase gene sequences presented in this thesis, a more complex and long-term evolutionary process would appear to be involved. The evolution of the strain *P.putida* PP3 reported by Senior *et al.*, (1976) took place over a much longer time period than is usual for the enrichment selection of organisms and was not the result of the change of a hydrolase to a dehalogenase but the increasing expression of existing dehalogenase genes. This form of evolution, rather than the modification of existing cellular hydrolases under laboratory pressure, would seem to be more able to account for the diversity of dehalogenases, with evolution being a long-term process rather than one able to take place over a matter of weeks in a laboratory.

6.7 Resolution of racemic mixtures

The use of enzymes as biocatalysts is being seen as a necessary progression in the fields of biotransformation and the production of chemicals, especially chiral intermediates. For this reason, the demonstration presented here of the use of dehalogenases to resolve racemic mixtures shows an important use for these enzymes. The dehalogenases were able to resolve a racemic mixture completely, which was akin to a process described in a number of

patents (Taylor, 1984, 1985). There were several limitations to the system described in this work that would need further research to remove. Firstly, the dehalogenases were expressed under the control of the *lac* promoter which requires the presence of IPTG for induction. In order to maintain the cost of any process as low as possible, removing the need for an expensive inducer would be desirable. The dehalogenase genes could be engineered so that expression was controlled by a promoter that was induced by temperature for example. The other major drawback seen from the experiments is the time span over which whole cells resolve the mixture when compared to the time span when cell-free extracts are used. This time difference was likely to be due to the slow uptake of the substrate into the cell. This could be overcome by permeabilising the cells with a compound such as toluene, but for an industrial application this would mean repeating this process for each batch of cells. The better alternative would be to clone the *Rhizobium* uptake system, or to utilise the uptake system of another organism. As it was possible that the uptake system from the *Rhizobium* is the same for both 2CP and DCP, then the simplest method of cloning the uptake system would be to use a clone that is known to act on DCP, i.e. pSC530 and then to make an additional genomic library in cells containing this plasmid. The library would need to be made using a plasmid with a different origin of replication to pUC, such as pACYC, so there would be no problems with incompatibility groups. Any colonies that were able to grow on DCP must now carry an uptake system for DCP and very possibly 2CP. This clone could then be engineered so it was co-expressed with either HadD or HadL and so an enhancement of the uptake of 2CP would be seen *in vivo*, thus making the process more efficient. Before this approach could be taken, toxicity of DCP to *E.coli* would need to be established, and also the problems of hypersensitivity, as mentioned earlier, need to be considered. Alternatively, it is known that there is a HAA permease on the

DEH transposable genetic element (Slater *et al.*, 1985) so this uptake system could be combined with the dehalogenases from the *Rhizobium* to form a catabolic plasmid capable of both uptake and metabolism. Another possible alternative to the use of the cloned genes would be to use a modified *Rhizobium*. It would be possible to use a combination of the pSC530 insert and either the cloned *hadD* and *hadL* genes to knockout the corresponding genes in the *Rhizobium*. This would then give an organism that was only capable of metabolising one of the two stereoisomers, so resolving the mixture, but there would be none of the problems of uptake or regulation of expression that occur when the dehalogenases are expressed in *E.coli*. This is the similar to the approach proposed in the ICI patents (Taylor, 1984 and Taylor 1988) whereby a mutation would be used to create a strain that is unable to dehalogenate the L-isomer of D/L-2CP.

Table 6.1

Organism	Mechanism	Gene name	Location of gene	Deduced M_r (kDa)	Native M_r (kDa)
Pseudomonas sp. CBS3	B1	*dehCI*	chromosome	25.4	41
	B1	*dehCII*	chromosome	25.6	64
P.putida AJ1/23	B1	*hadL*	chromosome	25.7	79
	B4	*hadD*	chromosome	33.6	135
P.putida PP3	B3	*dehI*	transposon	ND	ND
	B2	*dehII*	transposon	ND	ND
P.cepacia MBA4	B1	*hdlIVa*	chromosome	25.9	45
Xanthobacter autotrophicus GJ10	B1	*dhlB*	chromosome	27.4	ND
***Rhizobium* sp.**	**B1**	***hadL***	**chromosome**	**29.3**	**58**
	B4	***hadD***	**chromosome**	**30.8**	**60**

Table 6.1

Comparison of known 2-haloacid dehalogenase genes and their gene products with the dehalogenase genes cloned and analysed during the course of this research. Mechanism refers to the nomenclature of Leisinger and Bader (1993). Molecuar weights have only been given for genes that have been sequenced.

Chapter 7

References

7.1 References

Allison, N., (1981). Bacterial degradation of Halogenated aliphatic acids, PhD, Trent Polythecnic.

Allison, N., Skinner, A. J. and Cooper, R. A., (1983). The dehalogenase of a 2,2-dichloropropionate-degrading bacterium. J. G. M., **129**, 1283-1293.

Ashton, F. M. and Crafts, A. S., (1973). Mode of Action of Herbicides, Wiley and Sons, U.S.A.

Barth, P. T., Bolton, L. and Thomson, J. C., (1992). Cloning and Partial Sequencing of an Operon Encoding Two *Pseudomonas putida* Haloalkanoate Dehalogenases of Opposite Stereospecificity. J. Bact., **174.**, 2612-2619.

Bergman, J. G. and Sanik, J., (1957). Determination of trace amounts of chlorine in naptha. Anal. Chem., **29**, 241-243.

Berry, E. K. M., Allison, N., Skinner, A. J. and Cooper, R. A., (1979). Degradation of the selective herbicide 2,2-dichloropropionate (Dalapon) by a soil bacterium. J. G. M., **110**, 39-45.

Bollag, J. M. and Alexander, M., (1971). Bacterial Dehalogenation of chlorinated aliphatic acids. Soil Biol. Biochem., **3**, 91-96.

Brokamp, A. and Schmidt, F. R. J., (1991). Surival of *Alcaligenes xylosoxidans* degrading 2,2-Dichloropropionate and horizontal transfer of its halidohydrolase gene in a soil microcosm. Cur. Micro., **22**, 299-306.

Brown, T. A., (1991). Molecular Laboratory Labfax, Bios Scientific Publishers, Oxford, U.K.

Busto, M. D., Smith, P. P., Perez-Mateos, M. and Burns, R. G., (1992). Degradation of aliphatic halogen substituted pesticides by dehalogenase isolated from *Pseudomonas alcaligenes*. Identification and properties of the enzyme. Sci. Tot. Env., **123/124**, 267-277.

Castro, C. E. and Bartnicki, E. W., (1965). Biological cleavage of carbon-halogen bonds. Metabolism of 3-bromopropanol by *Pseudomonas* sp. Biochim. Biophys. Acta, **100**, 384-392.

Castro, C. E. and Belser, N. O., (1968). Biodehalogenation. Reductive dehalogenation of the biocides ethylene dibromide, 1,2-dibromo-3-chloropropane and 2,3-dibromobutane in soil. Env. Sci. and Tech., **2**, (10), 779-783.

Chaudry, G. R. and Chapalamadugu, S., (1991). Biodegradation of Halogenated Compounds. Micro. Rev., **55**, (1), 59-79.

Davis, J. L. and Evans, W. C., (1962). Elimination of halide ions from aliphatic halogen-substituted organic acids by an enzyme preparation from *Pseudomonas dehalogenans*. Biochem. J., **82**, 50-51.

De Maagd, R. A., Mulders, I. H. M., Canter Cremers, H. C. J. and Lugtenberg, B. J. J., (1992). Cloning, Nucleotide Sequence, and Expression in *Escherichia coli* of a *Rhizobium leguminosarum* Gene Encoding a Symbiotically Repressed Outer Memebrane Protein. J. Bact, **174**, (1), 214-221.

Feinberg, A. P. and Vogelstein, B., (1983). A Technique for Radiolabelling DNA Restriction Fragments to High Specific Activity. Anal. Biochem, **132**, 6-13.

Franken, S. M., Rozeboom, H. J., Kalk, K. H. and Dijkstra, B. W., (1991). Crystal structure of haloalkane dehalogenase: an enzyme to detoxify halogenated alkanes. EMBO J., **10**, (6), 1297-1302.

Gaunt, J. K. and Evans, W. C., (1971). Metabolism of 4-chloro-2-methyl-phenoxyacetate by a soil pseudomonad. Biochem. J., **122**, 533-542.

Ghosal, D., You, I., Chatterjee, D. K. and Chakrabarty, A. M., (1985). Microbial Dehalogenation of Halogenated Compounds. **228**, (4696), 135-142.

Goldman, P., (1965). The enzymatic cleavage of the carbon-fluorine bond in fluoroacetate. J. Bio. Chem., **240**, 3434-3438.

Goldman, P., (1972). Enzymology of carbon-halogen bonds, Washington, National Academy of Sciences, Washington, 147-163.

Golman, P., Milne, G. W. A. and Keister, D. B., (1968). Carbon-fluorine bond cleavage III. Studies on a bacterial halidohydrolase. J. Biol. Chem., **243**, 428-434.

Gornall, A. C., Bardawill, C. S. and David, M. M., (1949). Determination of Serum Proteins by Means of the Biuret Reaction. J. Biol. Chem, **166**, 1-19.

Gough, J. A. and Murray, N. E., (1983). Sequence Diversity among Related Genes for recognition of Specific Targets in DNA Molecules. J. Mol. Biol., **166**, 1-19.

Hardman, D. J. and Slater, J. H., (1981). The Dehalogenase Complment of a Soil Pseudomonad Grown in Closed and Open Cultures on Haloalkanoic Acids. J.G.M., **127**, 399-405.

Hardman, D. J. and Slater, J. H., (1981). Dehalogenases in Soil Bacteria. J.G.M., **123**, 117-128.

Hardman, D. J., Gowland, P. C. and Slater, J. H., (1986). Large Plasmids from Soil Bacteria Enriched on Halogenated Alkanoic Acids. Appl. Env. Micro., **51**, (1), 44-51.

Hardman, D. J., (1991). Biotransformation of Halogenated Compounds. Crit. Rev. Biotech., **11**, (1), 1-40.

Janssen, D. B., Preis, F., Van der Ploeg, J., Kazemeir, B., Terpstra, P. and Witholt, B., (1989). Cloning of 1,2-dichloroethane degradation genes of *Xanthobacter autotrophicus* GJ10 and expression and sequencing of the *dhlA* gene. J. Bact., **171**, (12), 6791-6799.

Jensen, H. L., (1957). Decomposition of chloro-substituted aliphatic acids by soil bacteria. Can. J. Micro., **3**, 151-164.

Jensen, H. L., (1960). Decomposition of chloroacetates and chloropropionates by bacteria. Acta Agric. Scand., **10**, 83-103.

Jones, D. H. A., Barth, P. T., Byrom, D. and Thomas, C. M., (1992). Nucleotide sequence of the structural gene encoding a 2-haloalkonic acid dehalogenase of *Pseudomonas putida* strain AJ1 and purification of the encoded protein. J.G.M., **138**, 675-683.

Kawasaki, H., Tone, N. and Tonomura, K., (1981). Plasmid-determined dehalogenation of Haloactetates in *Moraxella* Species. Agric Biol Chem, **45**, (1), 29-34.

Kawasaki, H., Yahahra, H. and Tonomura, K., (1981). Isolation and Characterization of Plasmid pUO1 Mediating Dehalogenation of Haloacetate And Mercury Resistance in *Moraxella* sp. B. Agric. Biol. Chem, **45**, (6), 1477-1481.

Kawasaki, H. and Tonomura, K., (1989). Gene Diversity of Bacterial Haloalkanoic Acid Dehalogenases. Rec. Adv. Micro. Ecol., **Proceedings of the 5th International Symposium on Microbial Ecology**, 600-604.

Kawasaki, H., Tsuda, K., Matsushita, I. and Tonomura, K., (1992). Lack of homology between two haloacetate dehalogenase gens encoded on a plasmid from *Moraxella* sp. strain B. J.G.M., **138**, 1317-1323.

Kearney, P. C., Kaufman, D. D. and Beall, M. L., (1964). Enzymic dehalogenation of 2,2-dichloropropionate. Biochem. Biophys. Res. Comm., **14**, 29-33.

Kearney, P. C., Harris, C. I., Kaufman, D. D. and Sheets, J. J., (1965). Behavior and fate of chlorinated aliphatic acids in soils. Adv. Pest. Res., **6**, 1-29.

Keuning, S., Janssen, D. B. and Witholt, B., (1985). Purifcation and characterisation of hydrolytic haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10. J. Bact., **163**, (2), 635-639.

Klages, U., Krauss, S. and Lingens, F., (1983). 2-Haloacid Dehalogenase from a 4-Chlorobenzoate-Degrading *Pseudomonas spec.* CBS3. Hoppe-Seyler's Z. Physiol. Chem., **364**, (5), 529-535.

Kocabiyik, S., (1989). Isolation and Preliminary Characterization of Dehalogenase I Enzyme from Soil Pseudomonas. **13**, (2), 88-96.

Kocabiyik, S. and Turkoglu, S., (1989). A Thermo-stable Dehalogenase in the Extracts from *Pseudomonas* sp. (Strain 19S). Biotech. Lett., **11**, (6), 397-400.

Kohler-Staub, D. and Leisinger, T., (1985). Dichloromethane Dehalogenase of *Hyphomicrobium* sp. Strain DM2. J. Bact., **162**, 676-681.

Kohler-Staub, D. and Kohler, H. P. E., (1989). Microbial degradation of ß-chlorinated four-carbon aliphatic acids. J. Bact., **171**, (3), 1428-1434.

Kraft, R., Tardiff, J., Krauter, K. S. and Leinwant, L. A., (1988). Using Mini-prepped DNA for sequencing Double-stranded Templates with Sequenase. Biotechniques, **6**, 544-547.

Kreig, N. R. and Holt, J. G., (1984). Bergey's Manual of Systematic Bacteriology, Williams and Wilkins, Baltimore, U.S.A.

Kushner, S. R., (1978). An Improved Method for the Transformation of *Escherichia coli* with ColE1-derived Plasmids.In Genetic Engineering pp 17-23, Eds. H. B. Boyer and S. Nicosia, Elsevier/North Holland Publishing Co., Amsterdam.

La Roche, S. D. and Leisinger, T., (1990). Sequence Analysis and Expression of the Bacterial Dichloromethane Dehalogenase Structural Gene, a Member of the Glutathione *S*-Transferase Supergene Family. J. Bact., **172**, (1), 164-171.

Layne, E., (1959). Spectrophotometric and Turbidimetric Methods for measuring Proteins. Methods in Enzymol., **3**, 447-454.

Leasure, J. K., (1964). The halogenated aliphatic acids. Agric. Food Chem., **12**, (1), 40-43.

Leigh, J. A., (1986a). Studies on Bacterial Dehalogenases, PhD, Trent Polythecnic.

Leigh, J. A., Skinner, A. J. and Cooper, R. A., (1986b). Isolation and partial characterisation of dehalogenase-deficient mutants of a *Rhizobium* species. FEMS Micro. Letts., **36**, 163-166.

Leigh, J. A., Skinner, A. J. and Cooper, R. A., (1988). Partial purification, stereospecificity and stoichiometry of three dehalogenases from a *Rhizobium* species. FEMS Micro. Letts., **49**, 353-356.

Leisinger, T., (1983). Microorganisms and Xenobiotics Compounds. Experientia, **39**, 1183-1191.

Leisinger, T. and Bader, R., (1993). Microbial Dehalogenation of Synthetic Organohalogen Compounds:Hydrolytic Dehalogenases. Chimia, **47**, 116-121.

Little, M. and Williams, P. A., (1971). A bacterial halidohydrolase; its purification, some properties and its modification by specific amino acid reagents. Eur. J. Bio., **21**, 99-109.

Magee, L. A. and Colmer, A. R., (1955). The effect of herbicides on soil micro-organisms III. The effect of some herbicides on the respiration of *Azotobacter*. Appl. Micro., **3**, 288-292.

Martinez, E., Romero, D. and Palacios, R., (1990). The *Rhizobium* Genome. Crit. Rev. Plant Sci., **9**, (1), 59-93.

Morsberger, F. M., Muller, R., Otto, M. K., Lingens, F. and Klube, K. D., (1991). Purification and Characterization of 2-Halocarboxylic Acid Dehalogenase II from *Pseudomonas spec.* CBS3. Biol. Chem. Hoppe-Seyler`, **372**, 915-922.

Motosugi, K., Esaki, N. and Soda, K., (1982). Bacterial Assimilation of D- and L-2-Chloropropionates and Occurence of a new Dehalogenase. Acta. Microbiol., **131**, 179-183.

Motosugi, K., Esaki, N. and Soda, K., (1982). Purification and Properties of a New Enzyme, DL-2-Haloacid Dehalogenase, from a *Pseudomonas* sp. J. Bact., **150**, (2), 522-527.

Motosugi, K. and Soda, K., (1983). Microbial Degradation of Synthetic Organochlorine Compounds. Experientia, **39**, 1214-1220.

Motosugi, K., Esaki, N. and Soda, K., (1984). Enzymatic Preparation of D- and L-Lactic Acid from Racemic 2-Chloropropionic Acid. Biotech. Bioeng., **26**, 805-806.

Muller, R. and Lingens, F., (1986). Microbial Degradation Of Halogenated Hydrocarbons:A Biological Solution to Pollution Problems? Angew. Chem. Int. Ed. Engl., **25**, 779-789.

Murdiyatmo, U., Asmara, W., Tsang, J. S. H., Baines Bull, A. J. and Hardman, D., J., (1992). Molecular biology of the 2-Haloacid halidohydrolase IVa from *Pseudomonas cepacia* MBA4. J. Biochem., **284**, 87-93.

Neilson, A. H., (1990). The Biodegradation of Halogenated Organic Compounds. J. Appl. Bact., **69**, 445-470.

Onda, M., Motosugi, K. and Nakajima, H., (1990). A New Approach for Enzymatic Synthesis of D-3-Chlorolactic Acid from Racemic 2,3-Dichloropropionic Acid by Halo Acid Dehalogenase. Agric. Biol. Chem., **54**, (11), 3031-3033.

Sambrook, J., Fritsch, E. F. and Maniatis, T., (1989). Molecular Cloning. A Laboratory Manual., 2nd, Cold Spring Harbour Press, U.S.A.,

Sanger, F., Nicklen, S. and Coulson, A. R., (1977). DNA Sequencing with chain-terminating Inhibitors. Proc. Natl. Acad. Sci. USA, **74**, 5463-5467.

Schneider, B., Müller, R., Frank, R. and Lingens, F., (1991). Complete nucleotide sequence and comparison of the structural genes of two 2-haloalkanoic acid dehalogenases from *Pseudomonas* sp. strain CBS3. J. Bact., **173**, 1530-1535.

Senior, E., Bull, A. T. and Slater, J. H., (1976). Enzyme evolution in a microbial community growing on the herbicide Dalapon. Nat., **263**, 476-479.

Shanker, R. and Robinson, J. P., (1991). Anaerobic degradation of dalapon in wateerlogged soil and river sediment. Lett. Appl. Micro., **12**, 8-10.

Slater, J. H., Lovatt, D., Weightman, A. J., Senoir, E. and Bull, A. T., (1979). The Growth of *Pseudomonas putida* on Chlorinated Aliphatic Acids and its Dehalogenase Activity. J.G.M., **114**, 125-136.

Slater, J. H., Weightman, A. J. and Hall, B. G., (1985). Dehalogenase Genes of *Pseudomonas putida* PP3 on Chromosomally Located Transposable Elements. Mol. Biol. Evol., **2**, (6), 557-567.

Smith, J. M., Harrison, K. and Colby, J., (1990). Purification and characterization of D-2-haloacid dehalogenase from *Pseudomonas putida* AJ1/23. J.G.M., **136**, 881-886.

Southern, E. M., (1975). Detection of Specific Sequences among DNA Fragments Separated by Gel Electrophoresis. J. Mol. Biol., **98**, 503-508.

Suida, J. F. and DeBernardis, J. F., (1973). Naturally Occuring Halogenated Organic Compounds. **36**, (2), 107-143.

Tabor, S. and Richardson, C. C., (1987). DNA Sequence Analysis with a Modified Bacteriophage T7 DNA Polymerase. Proc. Natl. Acad. Sci. USA, **84**, 4767-4771.

Tardif, G., W., G. C., Labbé, D. and Lau, P. C. K., (1991). Involvement of a large plasmid in the degradation of 1,2-dichloroethane by *Xanthobacter autotrophicus*. Appl. and Env. Micro., **57**, (6), 1853-1857.

Taylor, S. C., (1984). U.K. Patent 8431923.

Taylor, S. C., (1985). U.S. Patent 4758518.

Thomas, A. W., Slater, J. H. and Weightman, A. J., (1992a). The Dehalogenase Gene *dehI* from *Pseudomonas putida* PP3 Is Carried on an Unusual Mobile Genetic Element Designated *DEH*. J. Bact, **174**, (6), 1932-1940.

Thomas, A. W., W., T. A., Slater, J. H. and Weightman, A. J., (1992b). Localization and Functional Analysis of Structual and Regulaton Dehalogenase Genes Carried on *DEH* from *Pseudomonas putida* PP3. J. Bact., **174**, (6), 1941-1947.

Thomas, A. W., Lewington, J., Hope, S., Topping, A. W., Weightman, A. J. and Slater, J. H., (1992c). Environmentally directed mutations in the dehalogenase system of *Pseudomonas putida* strain PP3. Arch Microbiol, **158**, 176-182.

Tsang, J. S. H., Sallis, P. J., Bull, A. T. and Hardman, D. J., (1988). A monobromoacetate dehalogenase from *Pseudomonas cepacia* MBA4. Arch. Microbiol., **150**, 441-446.

Van den Wijngaard, A. J., Janssen, D. B. and Witholt, B., (1989). Degradation of epichlorohydrin and halohydrins by bacterial cultures isolated from freshwater sediment. J.G.M., **135**, 2199-2208.

Van den Wijngaard, A. J., Reuvekamp, P. T. W. and Janssen, D. B., (1991). Purification and characterisation of haloalcohol dehalogenase from *Arthrobacter* sp. strain AD2. J. Bact., **173**, (1), 124-129.

Van den Wijngaard, A. J., Van den Kamp, K. W. H. J., Van den Ploeg, J., Preis, F., Kazemier, B. and Janssen, D. B., (1992). degradation of 1,2-dichloroethane by *Ancylobacter aquaticus* and other facultative methylotrophs. Appl. and Env. Micro., **58**, (3), 976-983.

Van Der Meer, J. R., De Vos, W. M., Harayama, S. and Zehnder, A. J. B., (1992). Molecular Mechanisms of Genetic Adaptation to Xenobiotic Compounds. Micro. Rev., **56**, (4), 677-694.

Van der Ploeg, J., Van Hall, G. and Janssen, D. B., (1991). Characterization of the Haloacid Dehalogenase from *Xantobacter autotrophicus* GJ10 and Sequencing of the *dhl*B Gene. J. Bact., **173**, (24), 7925-7933.

Wackett, L. P., (1991). Dehalogenation of organohalide pollutants by bacterial enzymes and coenzymes, Texas A&M University, College Station, TX D910318-21, Plenum Press, 191-200.

Weightman, A. J., Slater, J. H. and Bull, A. T., (1979). The Partial Purification of Two Dehalogenases from *Pseudomonas putida* PP3. FEMS Micro. Lett., **6**, 231-234.

Weightman, A. J. and Slater, J. H., (1980). Selection of *Pseudomonas putida* Strains with Elevated Dehalogenase Activities by Continuous Culture Growth on Chlorinated Alkanoic Acids. J.G.M., **121**, 187-193.

Weightman, A. J., Weightman, A. L. and Slater, J. H., (1982). Stereospecificity of 2-Monochloropropionate Dehalogenation by the Two Dehalogenases of *Pseudomonas putida* PP3:Evidence of Two Different Dehalogenation Mechanisms. J.G.M., **128**, 1755-1762.

Weightman, A. J., Weightman, A. L. and Slater, J. H., (1985). Toxic Affects of Chorinated and Brominated Alkanoic Acids on *Pseudomonas putida*PP3: Selection at High Frequencies of Mutations in Genes encoding Dehalogenases. Appl. Env. Micro., **49**, (6), 1494-1501.

Weightman, A. L., Weightman, A. J. and Slater, J. H., (1992). Microbial dehalogenation of trichloroacetic acid. World J. Micro. and Biotech., **8**, 512-518.

Williams, S. G., Greenwood, J. A. and Jones, C. W., (1990). Isolation of novel strains of *Agrobacterium radiobacter* with altered capacities for lactose metabolism and succinoglucan production. J.G.M., **136**, 2179-2188.

Wilson, K., (1990). Preparation of Genomic DNA from Bacteria. Cur. Protocols in Mol. Biol., **1**, 2.4.1-2.4.5.

Yanisch-Perron, C., Vieria, J. and Messing, J., (1985). Improved M13 Cloning Vectors and Host Strains: Nucleotide Sequences of the M13 mp18 and pUC19 Vectors. Gene, **33**, 103-119.

Yokota, T., Omori, T. and Kodama, T., (1987). Purification and properties of haloalkane dehalogenase from *Corynebacterium* sp. strain m15-3. J. Bact., **169**, (9), 4049-4054.

Appendix I

Appendix I

Abstract of poster presented at the 6th European Congress on Biotechnology, Florence, Italy, (June 1993).

Cloning and Analysis of Rhizobial Dehalogenase Genes

S. Cairns and R. A. Cooper.

The production of specific stereo-isomers from racemic mixtures of chiral compounds is an important process in the manufacture of certain herbicides and pharmaceuticals. We are studying ways in which individual stereo-isomers of D/L-2-chloropropionic acid (D/L-2CP) can be prepared by a biotransformation process involving the use of stereospecific dehalogenase enzymes.

A *Rhizobium* sp. capable of growth on 2,2-dichloropropionic acid (DCP), and D/L-2CP had previously been isolated, and was shown to contain 3 distinct, inducible dehalogenase enzymes, with individual substrate specificities.

To investigate the genes encoding these enzymes a plasmid based Rhizobial genomic library was constructed in *Escherichia coli* NM522 and screened for the ability to grow on D/L-2CP. Two clones, SC1 and SC2, were obtained and have been shown to contain, respectively, a 6.5kb and 3.0kb *Eco*RI insert. These inserts also conferred the ability to grow on the separate stereo-isomers of 2CP. Restriction mapping and Southern Blotting analysis has shown that the two clones are unrelated to each other, however both hybridize to Rhizobial genomic DNA.

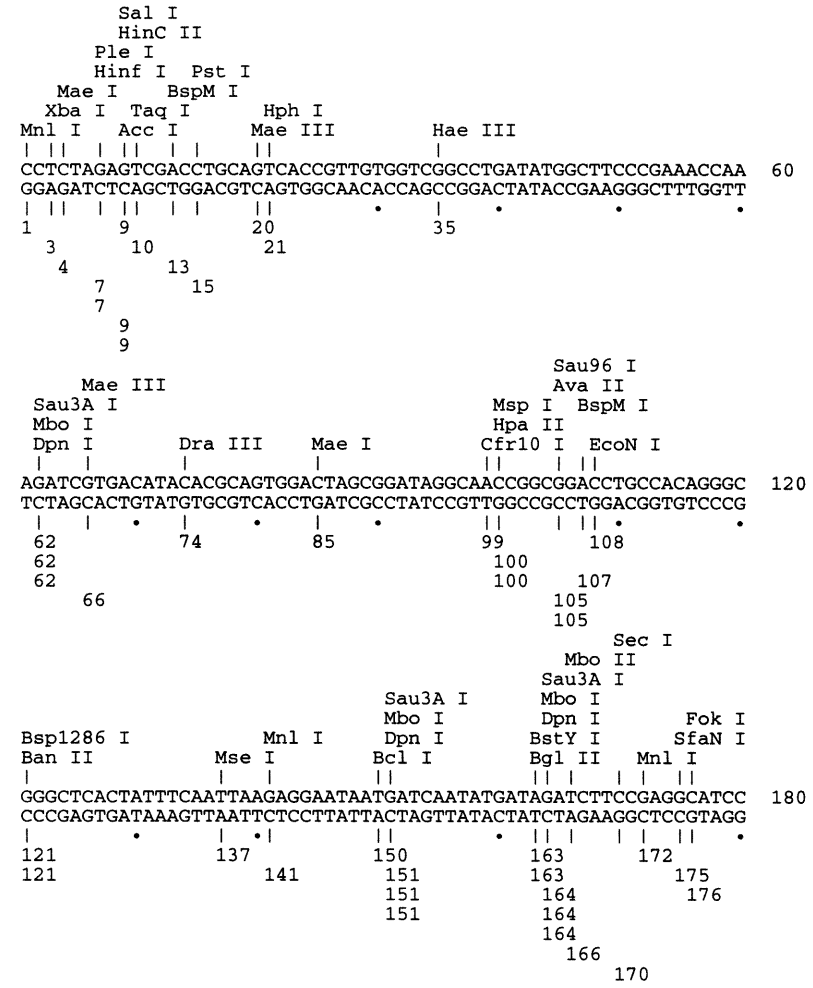
A 2.0kb subclone of SC1 retained the original growth characteristics and the cloned genes were over-expressed in *E.coli* to give 20% of total cell protein. This resulted in two major proteins of 31kDa and 33kDa. These proteins were partially purified and their N-terminal amino acid sequences

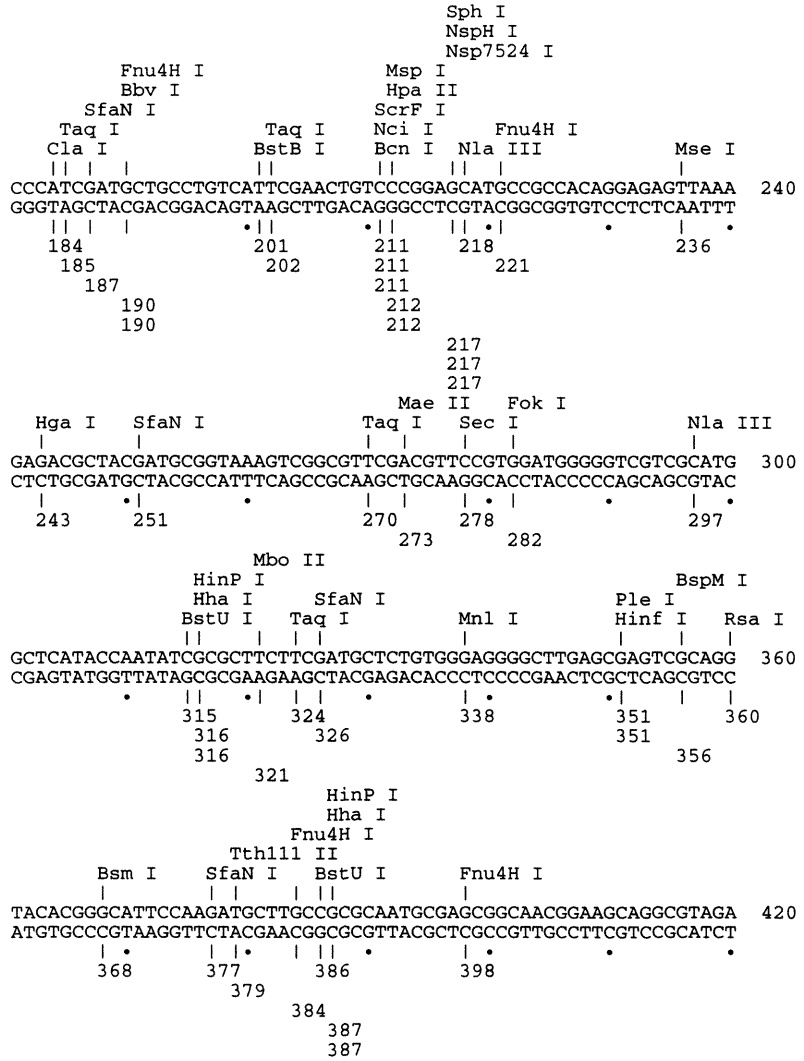
determined. These sequences showed no homology to those of published dehalogenase N-terminal sequence. Nucleotide sequencing of SC1 is underway, as are further studies of SC2. Investigation into the inter-relatedness of the clones and their gene products is being carried out

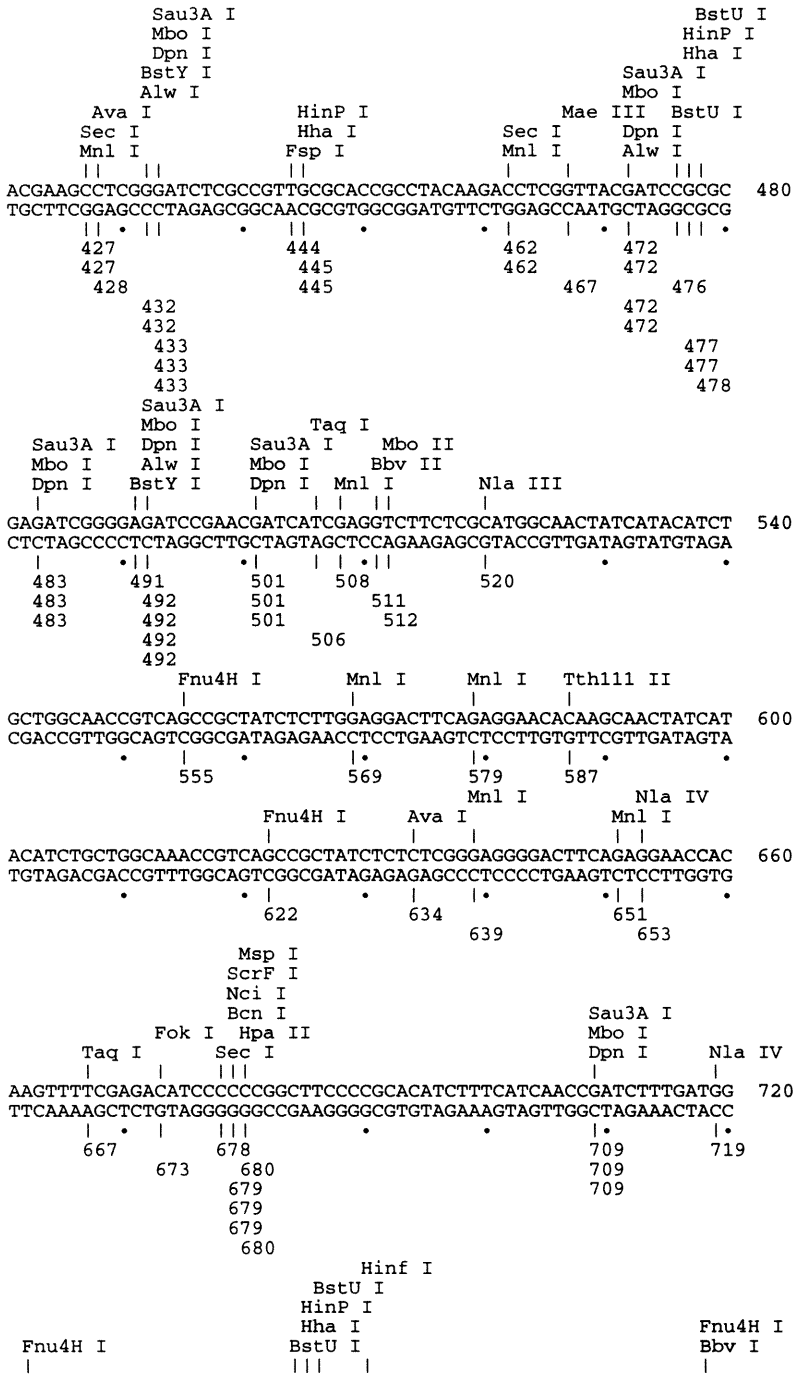
Appendix II

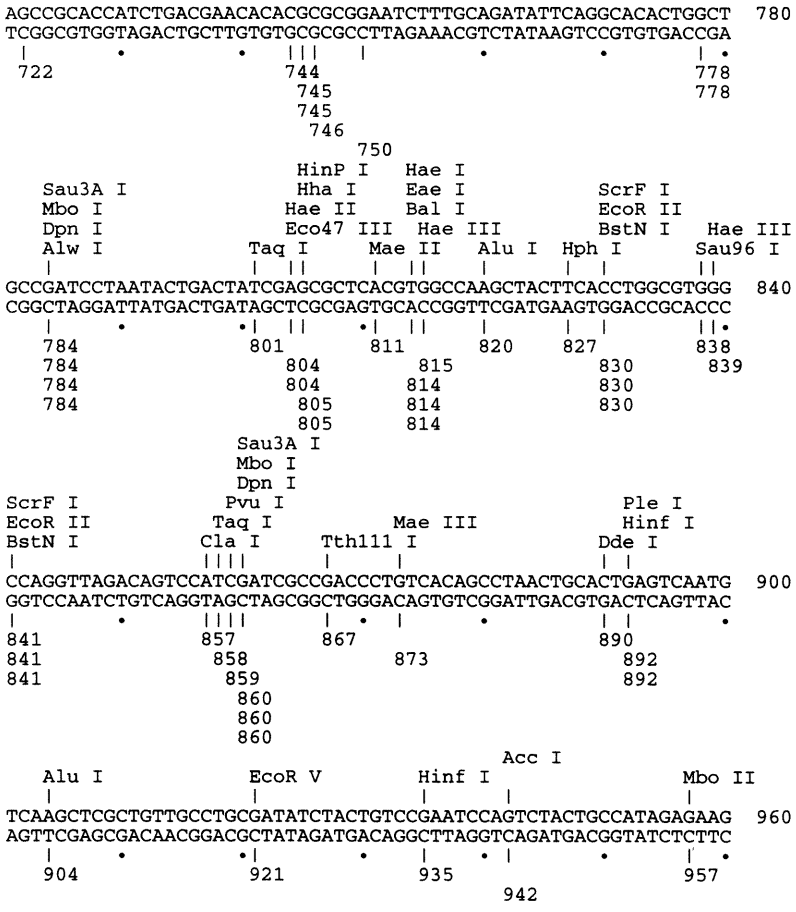
Appendix II

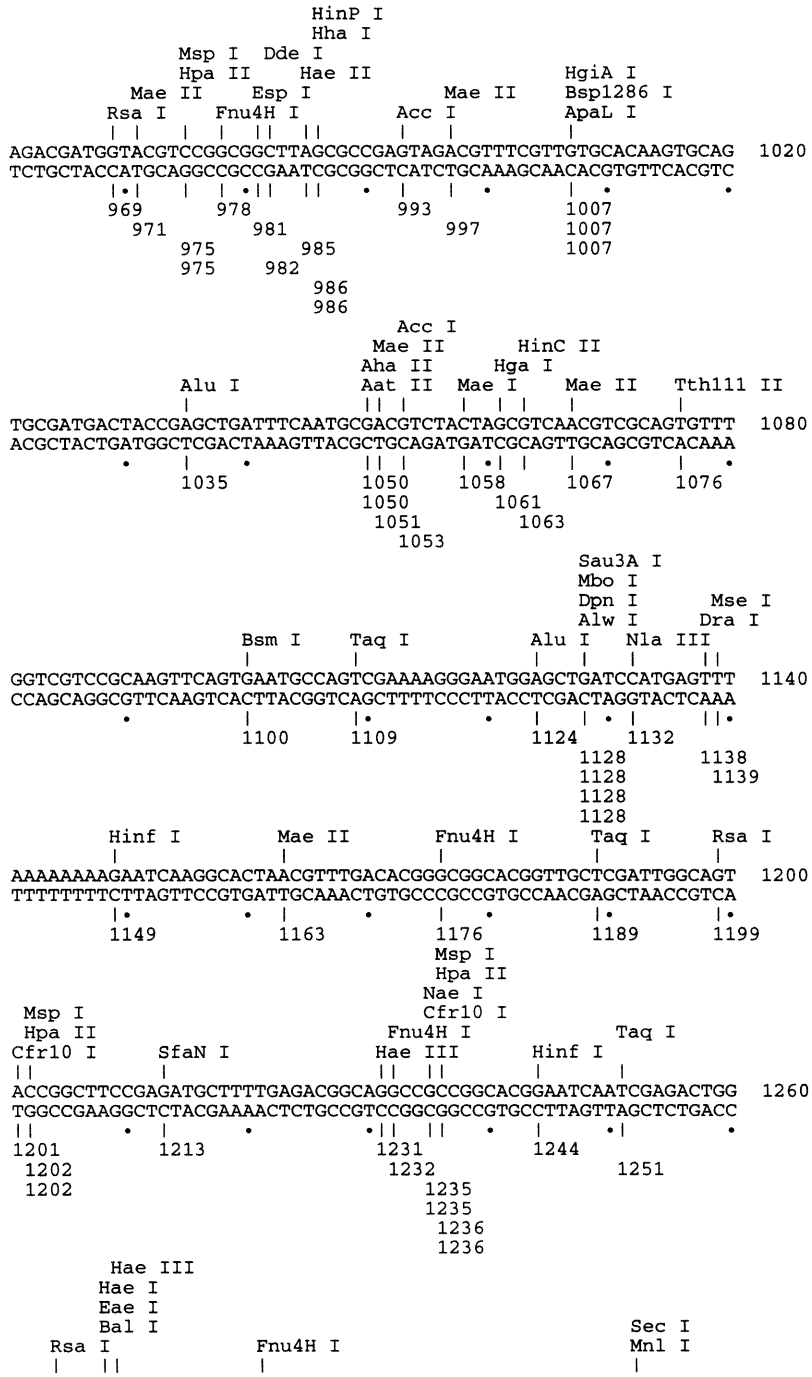
Complete nucleotide sequence obtained from pSC2.

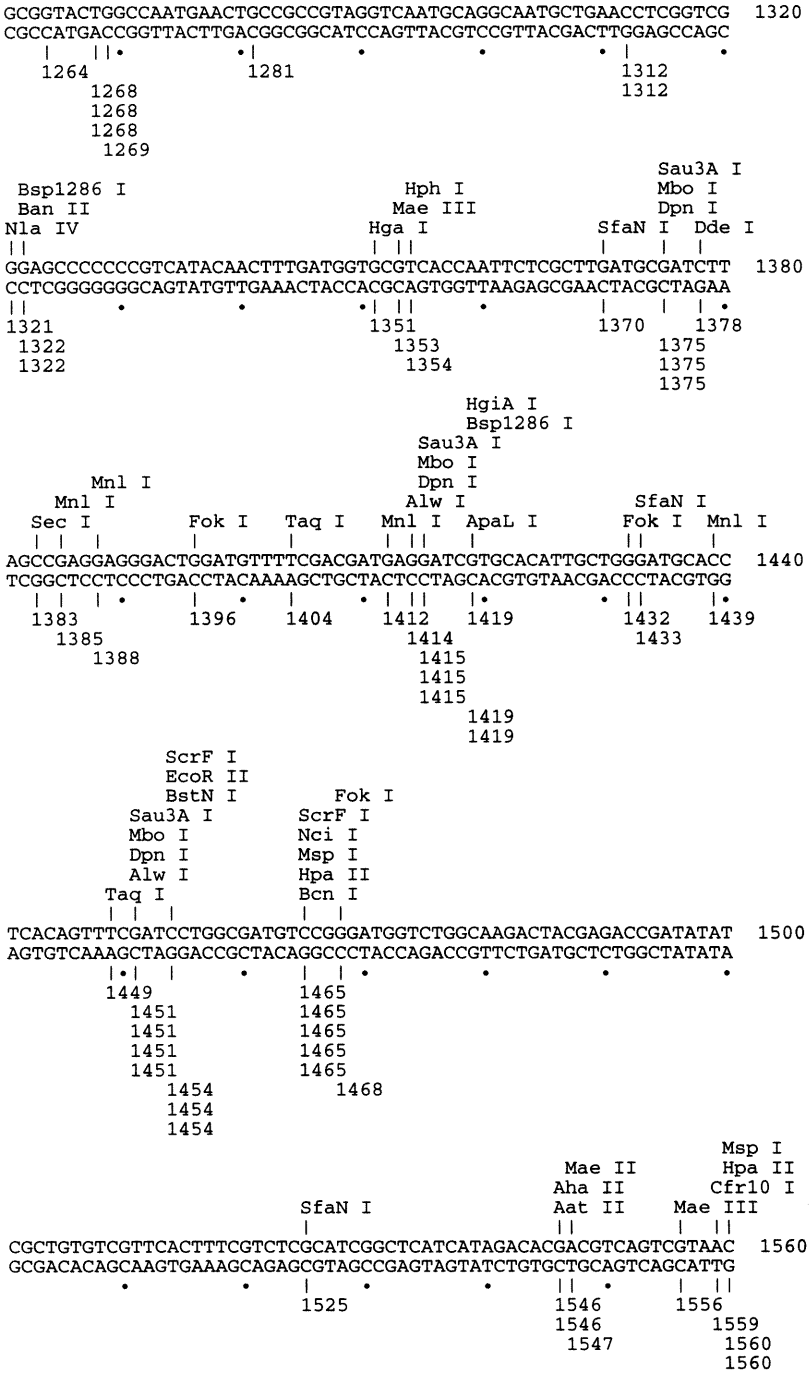


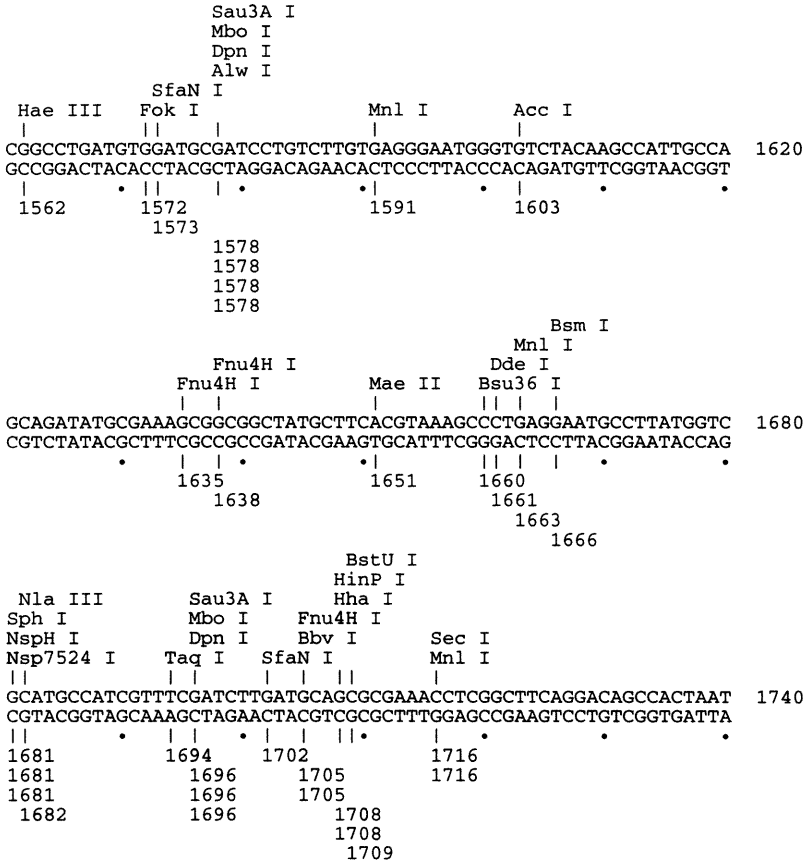

Sph I
NspH I
Nsp7524 I
Fnu4H I
Bbv I
Msp I
Hpa II
SfaN I
ScrF I
Taq I
Taq I
Nci I
Fnu4H I
Cla I
BstB I
Bcn I
Nla III
Mse I
CCCATCGATGCTGCCTGTCATTCGAACTGTCCCGGAGCATGCCGCCACAGGAGAGTTAAA 240
GGGTAGCTACGACGGACAGTAAGCTTGACAGGGCCTCGTACGGCGGTGTCCTCTCAATTT
184
201
211
218
236
185
202
211
221
187
211
190
212
190
212
217
217
217
Mae II
Fok I
Hga I
SfaN I
Taq I
Sec I
Nla III
GAGACGCTACGATGCGGTAAAGTCGGCGTTCGACGTTCCGTGGATGGGGGTCGTCGCATG 300
CTCTGCGATGCTACGCCATTTCAGCCGCAAGCTGCAAGGCACCTACCCCCAGCAGCGTAC
243
251
270
278
297
273
282
Mbo II
HinP I
Hha I
SfaN I
BspM I
Ple I
BstU I
Taq I
Mnl I
Hinf I
Rsa I
GCTCATACCAATATCGCGCTTCTTCGATGCTCTGTGGGAGGGGCTTGAGCGAGTCGCAGG 360
CGAGTATGGTTATAGCGCGAAGAAGCTACGAGACACCCTCCCCGAACTCGCTCAGCGTCC
315
324
338
351
360
316
326
351
316
356
321
HinP I
Hha I
Fnu4H I
Tth111 II
Bsm I
SfaN I
BstU I
Fnu4H I
TACACGGGCATTCCAAGATGCTTGCCGCGCAATGCGAGCGGCAACGGAAGCAGGCGTAGA 420
ATGTGCCCGTAAGGTTCTACGAACGGCGCGTTACGCTCGCCGTTGCCTTCGTCCGCATCT
368
377
386
398
379
384
387
387

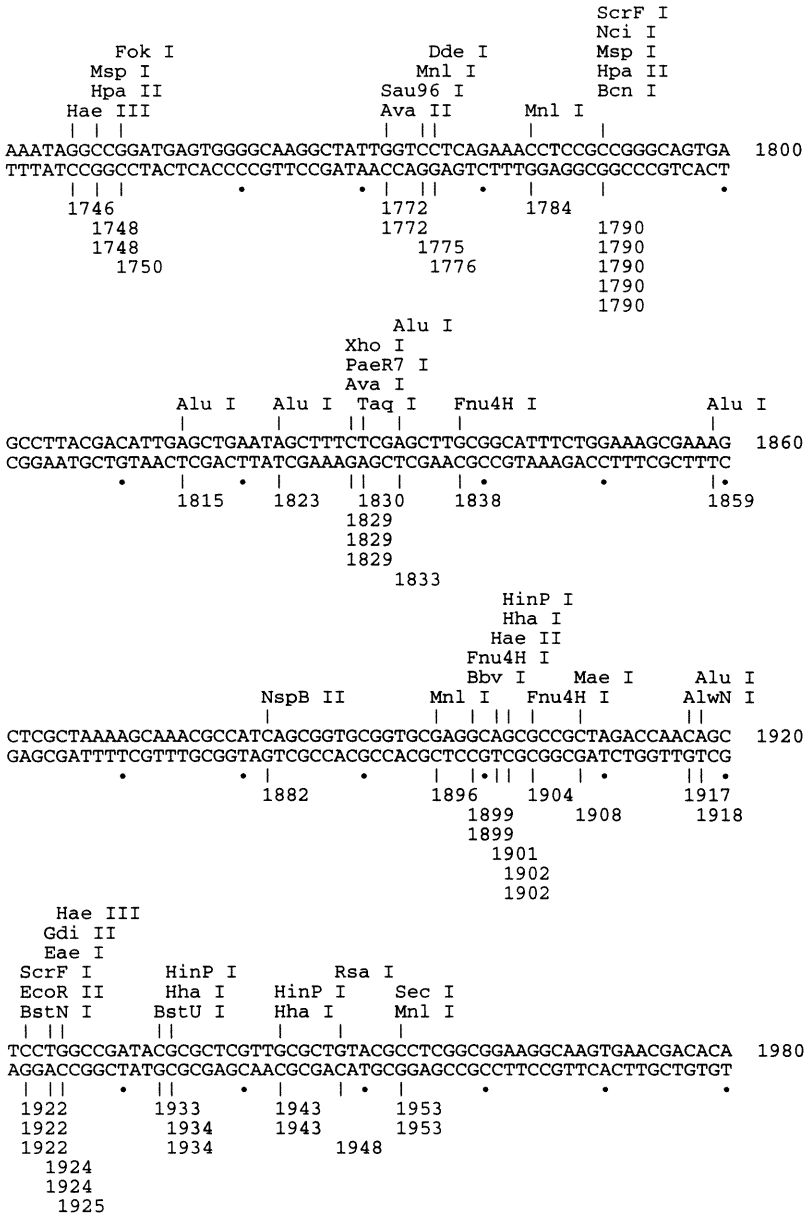

Sau3A I
Mbo I
Dpn I
BstY I
Alw I
BstU I
HinP I
Hha I
Sau3A I
Mbo I
Ava I
Sec I
Mnl I
HinP I
Hha I
Fsp I
Mae III
BstU I
Sec I
Dpn I
Mnl I
Alw I
ACGAAGCCTCGGGATCTCGCCGTTGCGCACCGCCTACAAGACCTCGGTTACGATCCGCGC 480
TGCTTCGGAGCCCTAGAGCGGCAACGCGTGGCGGATGTTCTGGAGCCAATGCTAGGCGCG
427
427
428
432
432
433
433
433
477
477
478
Sau3A I
Mbo I
Taq I
Sau3A I
Dpn I
Sau3A I
Mbo II
Mbo I
Alw I
Mbo I
Bbv II
Dpn I
BstY I
Dpn I
Mnl I
Nla III
GAGATCGGGGAGATCCGAACGATCATCGAGGTCTTCTCGCATGGCAACTATCATACATCT 540
CTCTAGCCCCTCTAGGCTTGCTAGTAGCTCCAGAAGAGCGTACCGTTGATAGTATGTAGA
483
483
483
491
492
492
492
501
501
501
508
511
512
506
520
Fnu4H I
Mnl I
Mnl I
Tth111 II
GCTGGCAACCGTCAGCCGCTATCTCTTGGAGGACTTCAGAGGAACACAAGCAACTATCAT 600
CGACCGTTGGCAGTCGGCGATAGAGAACCTCCTGAAGTCTCCTTGTGTTCGTTGATAGTA
555
569
579
587
Mnl I
Nla IV
Fnu4H I
Ava I
Mnl I
ACATCTGCTGGCAAACCGTCAGCCGCTATCTCTCTCGGGAGGGGACTTCAGAGGAACCAC 660
TGTAGACGACCGTTTGGCAGTCGGCGATAGAGAGAGCCCTCCCCTGAAGTCTCCTTGGTG
622
634
639
651
653
Msp I
ScrF I
Nci I
Bcn I
Sau3A I
Mbo I
Fok I
Hpa II
Taq I
Sec I
Dpn I
Nla IV
AAGTTTTCGAGACATCCCCCCGGCTTCCCCGCACATCTTTCATCAACCGATCTTTGATGG 720
TTCAAAAGCTCTGTAGGGGGGCCGAAGGGGCGTGTAGAAAGTAGTTGGCTAGAAACTACC
667
673
678
680
679
679
679
680
709
709
709
719
Hinf I
BstU I
HinP I
Hha I
Fnu4H I
BstU I
Fnu4H I
Bbv I

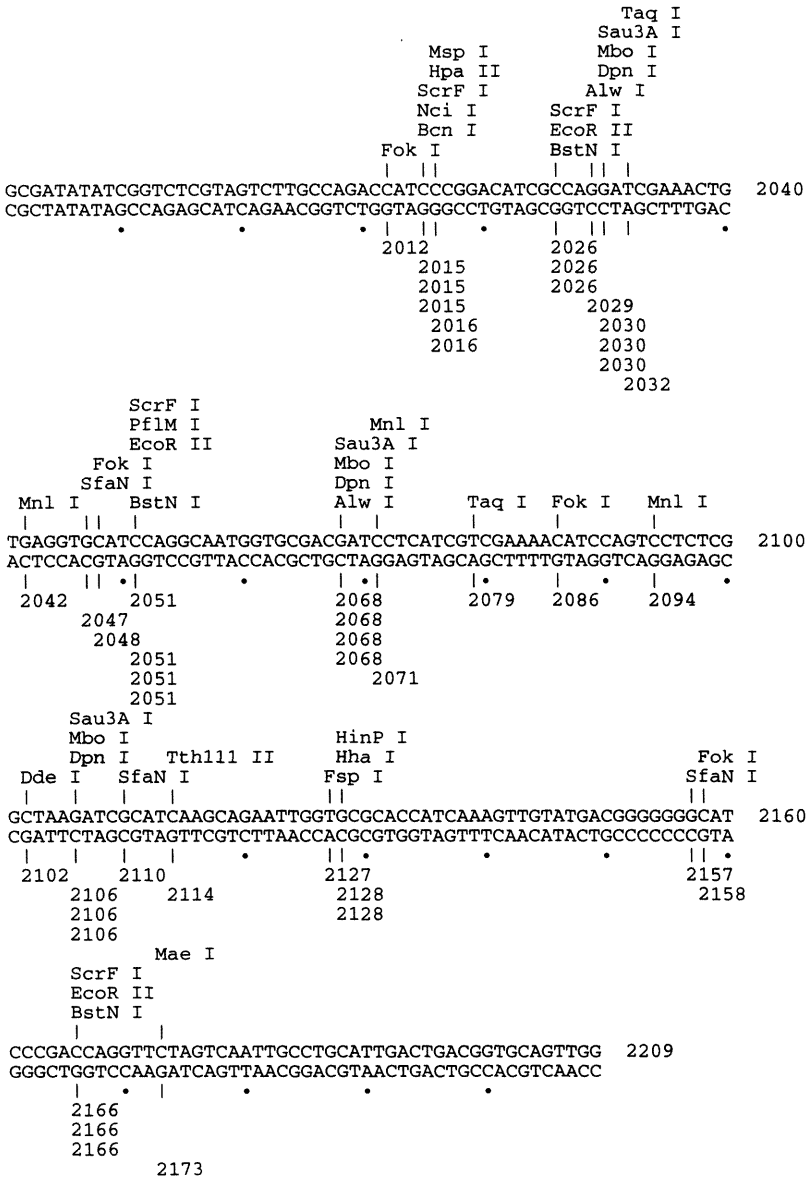

AGCCGCACCATCTGACGAACACACGCGCGGAATCTTTGCAGATATTCAGGCACACTGGCT 780
TCGGCGTGGTAGACTGCTTGTGTGCGCGCCTTAGAAACGTCTATAAGTCCGTGTGACCGA
722
744
778
745
778
745
746
750
HinP I Hae I
Sau3A I Hha I Eae I ScrF I
Mbo I Hae II Bal I EcoR II
Dpn I Eco47 III Hae III BstN I Hae III
Alw I Taq I Mae II Alu I Hph I Sau96 I
GCCGATCCTAATACTGACTATCGAGCGCTCACGTGGCCAAGCTACTTCACCTGGCGTGGG 840
CGGCTAGGATTATGACTGATAGCTCGCGAGTGCACCGGTTCGATGAAGTGGACCGCACCC
784
801
811
820
827
838
784
804
815
830
839
784
804
814
830
784
805
814
830
805
814
Sau3A I
Mbo I
Dpn I
ScrF I Pvu I Ple I
EcoR II Taq I Mae III Hinf I
BstN I Cla I Tth111 I Dde I
CCAGGTTAGACAGTCCATCGATCGCCGACCCTGTCACAGCCTAACTGCACTGAGTCAATG 900
GGTCCAATCTGTCAGGTAGCTAGCGGCTGGGACAGTGTCGGATTGACGTGACTCAGTTAC
841
857
867
890
841
858
873
892
841
859
892
860
860
860
Acc I
Alu I EcoR V Hinf I Mbo II
TCAAGCTCGCTGTTGCCTGCGATATCTACTGTCCGAATCCAGTCTACTGCCATAGAGAAG 960
AGTTCGAGCGACAACGGACGCTATAGATGACAGGCTTAGGTCAGATGACGGTATCTCTTC
904
921
935
957
942


HinP I
Hha I
Msp I Dde I
Hpa II Hae II
HgiA I
Mae II Esp I Mae II Bsp1286 I
Rsa I Fnu4H I Acc I ApaL I
AGACGATGGTACGTCCGGCGGCTTAGCGCCGAGTAGACGTTTCGTTGTGCACAAGTGCAG 1020
TCTGCTACCATGCAGGCCGCCGAATCGCGGCTCATCTGCAAAGCAACACGTGTTCACGTC
969 978 993 1007
971 981 997 1007
975 985 1007
975 982
986
986
Acc I
Mae II HinC II
Aha II Hga I
Alu I Aat II Mae I Mae II Tth111 II
TGCGATGACTACCGAGCTGATTTCAATGCGACGTCTACTAGCGTCAACGTCGCAGTGTTT 1080
ACGCTACTGATGGCTCGACTAAAGTTACGCTGCAGATGATCGCAGTTGCAGCGTCACAAA
1035 1050 1058 1067 1076
1050 1061
1051 1063
1053
Sau3A I
Mbo I
Dpn I Mse I
Alw I Dra I
Bsm I Taq I Alu I Nla III
GGTCGTCCGCAAGTTCAGTGAATGCCAGTCGAAAAGGGAATGGAGCTGATCCATGAGTTT 1140
CCAGCAGGCGTTCAAGTCACTTACGGTCAGCTTTTCCCTTACCTCGACTAGGTACTCAAA
1100 1109 1124 1132
1128 1138
1128 1139
1128
1128
Hinf I Mae II Fnu4H I Taq I Rsa I
AAAAAAAAGAATCAAGGCACTAACGTTTGACACGGGCGGCACGGTTGCTCGATTGGCAGT 1200
TTTTTTTTCTTAGTTCCGTGATTGCAAACTGTGCCCGCCGTGCCAACGAGCTAACCGTCA
1149 1163 1176 1189 1199
Msp I
Hpa II
Nae I
Cfr10 I
Msp I
Hpa II
Fnu4H I Taq I
Cfr10 I SfaN I Hae III Hinf I
ACCGGCTTCCGAGATGCTTTTGAGACGGCAGGCCGCCGGCACGGAATCAATCGAGACTGG 1260
TGGCCGAAGGCTCTACGAAAACTCTGCCGTCCGGCGGCCGTGCCTTAGTTAGCTCTGACC
1201 1213 1231 1244
1202 1232 1251
1202 1235
1235
1236
1236
Hae III
Hae I
Eae I
Bal I
Sec I
Rsa I Fnu4H I Mnl I


GCGGTACTGGCCAATGAACTGCCGCCGTAGGTCAATGCAGGCAATGCTGAACCTCGGTCG 1320
CGCCATGACCGGTTACTTGACGGCGGCATCCAGTTACGTCCGTTACGACTTGGAGCCAGC
1264
1281
1312
1312
1268
1268
1268
1269
Sau3A I
Mbo I
Dpn I
Bsp1286 I
Ban II
Nla IV
Hph I
Mae III
Hga I
SfaN I
Dde I
GGAGCCCCCCCGTCATACAACTTTGATGGTGCGTCACCAATTCTCGCTTGATGCGATCTT 1380
CCTCGGGGGGGCAGTATGTTGAAACTACCACGCAGTGGTTAAGAGCGAACTACGCTAGAA
1321
1351
1370
1378
1322
1353
1375
1322
1354
1375
1375
HgiA I
Bsp1286 I
Sau3A I
Mbo I
Dpn I
Alw I
Mnl I
Mnl I
Sec I
Fok I
Taq I
Mnl I
ApaL I
SfaN I
Fok I
Mnl I
AGCCGAGGAGGGACTGGATGTTTTCGACGATGAGGATCGTGCACATTGCTGGGATGCACC 1440
TCGGCTCCTCCCTGACCTACAAAAGCTGCTACTCCTAGCACGTGTAACGACCCTACGTGG
1383
1396
1404
1412
1419
1432
1439
1385
1414
1433
1388
1415
1415
1415
1419
1419
ScrF I
EcoR II
BstN I
Fok I
Sau3A I
Mbo I
Dpn I
Alw I
Taq I
ScrF I
Nci I
Msp I
Hpa II
Bcn I
TCACAGTTTCGATCCTGGCGATGTCCGGGATGGTCTGGCAAGACTACGAGACCGATATAT 1500
AGTGTCAAAGCTAGGACCGCTACAGGCCCTACCAGACCGTTCTGATGCTCTGGCTATATA
1449
1465
1451
1465
1451
1465
1451
1465
1451
1465
1454
1468
1454
1454
Msp I
Hpa II
Mae II
Aha II
Cfr10 I
Aat II
Mae III
SfaN I
CGCTGTGTCGTTCACTTTCGTCTCGCATCGGCTCATCATAGACACGACGTCAGTCGTAAC 1560
GCGACACAGCAAGTGAAAGCAGAGCGTAGCCGAGTAGTATCTGTGCTGCAGTCAGCATTG
1525
1546
1556
1546
1559
1547
1560
1560


Sau3A I
Mbo I
Dpn I
Alw I
SfaN I
Hae III
Fok I
Mnl I
Acc I
CGGCCTGATGTGGATGCGATCCTGTCTTGTGAGGGAATGGGTGTCTACAAGCCATTGCCA 1620
GCCGGACTACACCTACGCTAGGACAGAACACTCCCTTACCCACAGATGTTCGGTAACGGT
1562
1572
1573
1591
1603
1578
1578
1578
1578
Bsm I
Mnl I
Fnu4H I
Dde I
Fnu4H I
Mae II
Bsu36 I
GCAGATATGCGAAAGCGGCGGCTATGCTTCACGTAAAGCCCTGAGGAATGCCTTATGGTC 1680
CGTCTATACGCTTTCGCCGCCGATACGAAGTGCATTTCGGGACTCCTTACGGAATACCAG
1635
1638
1651
1660
1661
1663
1666
BstU I
HinP I
Nla III
Sau3A I
Hha I
Sph I
Mbo I
Fnu4H I
NspH I
Dpn I
Bbv I
Sec I
Nsp7524 I
Taq I
SfaN I
Mnl I
GCATGCCATCGTTTCGATCTTGATGCAGCGCGAAACCTCGGCTTCAGGACAGCCACTAAT 1740
CGTACGGTAGCAAAGCTAGAACTACGTCGCGCTTTGGAGCCGAAGTCCTGTCGGTGATTA
1681
1694
1702
1716
1681
1696
1705
1716
1681
1696
1705
1682
1696
1708
1708
1709


ScrF I
Nci I
Fok I
Dde I
Msp I
Mnl I
Msp I
Hpa II
Sau96 I
Hpa II
Hae III
Ava II
Mnl I
Bcn I
AAATAGGCCGGATGAGTGGGGCAAGGCTATTGGTCCTCAGAAACCTCCGCCGGGCAGTGA 1800
TTTATCCGGCCTACTCACCCCGTTCCGATAACCAGGAGTCTTTGGAGGCGGCCCGTCACT
1746
1748
1748
1750
1772
1772
1775
1776
1784
1790
1790
1790
1790
1790
Alu I
Xho I
PaeR7 I
Ava I
Alu I
Alu I
Taq I
Fnu4H I
Alu I
GCCTTACGACATTGAGCTGAATAGCTTTCTCGAGCTTGCGGCATTTCTGGAAAGCGAAAG 1860
CGGAATGCTGTAACTCGACTTATCGAAAGAGCTCGAACGCCGTAAAGACCTTTCGCTTTC
1815
1823
1830
1838
1859
1829
1829
1829
1833
HinP I
Hha I
Hae II
Fnu4H I
Bbv I
Mae I
Alu I
NspB II
Mnl I
Fnu4H I
AlwN I
CTCGCTAAAAGCAAACGCCATCAGCGGTGCGGTGCGAGGCAGCGCCGCTAGACCAACAGC 1920
GAGCGATTTTCGTTTGCGGTAGTCGCCACGCCACGCTCCGTCGCGGCGATCTGGTTGTCG
1882
1896
1904
1917
1899
1908
1918
1899
1901
1902
1902
Hae III
Gdi II
Eae I
ScrF I
HinP I
Rsa I
EcoR II
Hha I
HinP I
Sec I
BstN I
BstU I
Hha I
Mnl I
TCCTGGCCGATACGCGCTCGTTGCGCTGTACGCCTCGGCGGAAGGCAAGTGAACGACACA 1980
AGGACCGGCTATGCGCGAGCAACGCGACATGCGGAGCCGCCTTCCGTTCACTTGCTGTGT
1922
1922
1922
1924
1924
1925
1933
1934
1934
1943
1943
1948
1953
1953


Taq I
Sau3A I
Msp I
Mbo I
Hpa II
Dpn I
ScrF I
Alw I
Nci I
ScrF I
Bcn I
EcoR II
Fok I
BstN I
GCGATATATCGGTCTCGTAGTCTTGCCAGACCATCCCGGACATCGCCAGGATCGAAACTG 2040
CGCTATATAGCCAGAGCATCAGAACGGTCTGGTAGGGCCTGTAGCGGTCCTAGCTTTGAC
2012
2026
2015
2026
2015
2026
2015
2029
2016
2030
2016
2030
2030
2032
ScrF I
PflM I
Mnl I
EcoR II
Sau3A I
Fok I
Mbo I
SfaN I
Dpn I
Mnl I
BstN I
Alw I
Taq I
Fok I
Mnl I
TGAGGTGCATCCAGGCAATGGTGCGACGATCCTCATCGTCGAAAACATCCAGTCCTCTCG 2100
ACTCCACGTAGGTCCGTTACCACGCTGCTAGGAGTAGCAGCTTTTGTAGGTCAGGAGAGC
2042
2051
2068
2079
2086
2094
2047
2068
2048
2068
2051
2068
2051
2071
2051
Sau3A I
Mbo I
HinP I
Dpn I
Tth111 II
Hha I
Fok I
Dde I
SfaN I
Fsp I
SfaN I
GCTAAGATCGCATCAAGCAGAATTGGTGCGCACCATCAAAGTTGTATGACGGGGGGGCAT 2160
CGATTCTAGCGTAGTTCGTCTTAACCACGCGTGGTAGTTTCAACATACTGCCCCCCCGTA
2102
2110
2127
2157
2106
2114
2128
2158
2106
2128
2106
Mae I
ScrF I
EcoR II
BstN I
CCCGACCAGGTTCTAGTCAATTGCCTGCATTGACTGACGGTGCAGTTGG 2209
GGGCTGGTCCAAGATCAGTTAACGGACGTAACTGACTGCCACGTCAACC
2166
2166
2166
2173