WARWICK
THE UNIVERSITY OF WARWICK

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

**Permanent WRAP URL:**
http://wrap.warwick.ac.uk/131941

**Copyright and reuse:**
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

**warwick.ac.uk/lib-publications**

# The Application of Molecular Biology Techniques to the Taxonomy of the Pink Pigmented Facultative Methylotrophs

by

Derek W. Hood, B.Sc. (Edinburgh)

This thesis is presented for the degree of Ph.D. in the Department of Biological Sciences, University of Warwick, Coventry.

December 1985

TO MY FAMILY

Contents

Page No.

List of Contents I

List of Figures VIII

List of Tables X

Declaration XI

Acknowledgements XII

Abbreviations XIII

Summary XIV

1. Introduction

1.1 General Introduction 1

1.2 Categorisation 2

1.2.1 Obligate methane utilisers 2

1.2.2 Obligate methanol utilisers 3

1.2.3 Facultative methanol utilisers 4

1.3 Carbon metabolism 5

1.3.1 Ribulose monophosphate pathway 5

1.3.2 Serine pathway 7

1.3.3 Ribulose bisphosphate pathway 7

1.3.4 Energy metabolism 10

1.4 Methylotrophic yeasts 11

1.5 Pink Pigmented Facultative Methylotrophs 11

1.5.1 Characterisation 12

1.5.2 Carbohydrate metabolism 13

Page No.


1.5.3 Carbon assimilation 16

1.5.4 Regulation of carbon assimilation 21

1.5.5 C1 Oxidation and energy metabolism 21

1.5.5.1 Methane mono-oxygenase 22

1.5.5.2 Methanol dehydrogenase 23

1.5.5.3 Formaldehyde dehydrogenase 25

1.5.5.4 Formate dehydrogenase 26

1.5.5.5 Methylamine dehydrogenase 26

1.5.5.6 Electron transport and energy transduction systems 27

1.5.6 Regulation of C1 metabolism 29

1.5.7 Pigments of the PPFMs 32

1.5.8 Genetics of the PPFMs 34

1.5.8.1 Mutagenesis 34

1.5.8.2 Plasmids 35

1.5.8.3 Gene transfer systems 35

1.5.8.4 Gene cloning systems 36

1.5.9 Classification and Taxonomy 38

1.5.10 Aims 43

2. Introduction to Methods and Bacterial Classification Systems 44

2.1 Numerical Taxonomy 45

2.2 DNA Base Composition 47

2.2.1 Thermal denaturation method 47

Page No.


| | | | Page No. |
|---|---|---|---|
| | 2.2.2 | Density gradient method | 48 |
| | 2.2.3 | DNA restriction analysis | 49 |
| 2.3 | DNA-DNA hybridisation | | 49 |
| | 2.3.1 | Matrix methods | 51 |
| | 2.3.2 | Free solution methods | 53 |
| | 2.3.3 | Thermal denaturation of DNA hybrids | 54 |
| 2.4 | Polyacrylamide gel electrophoresis | | 54 |
| 2.5 | Other taxonomic methods | | 58 |
| 2.6 | Phylogenetic techniques | | 58 |
| | 2.6.1 | Protein studies | 58 |
| | 2.6.2 | rRNA studies | 59 |
| | | a) DNA-rRNA hybridisation | 59 |
| | | b) rRNA oligonucleotide sequencing | 61 |
| 3. | Materials and Methods | | 63 |
| 3.1 | Strains and culture conditions | | 63 |
| | 3.1.1 | Strains | 63 |
| | 3.1.2 | Culture and maintenance of strains | 63 |
| 3.2 | Cell lysis and DNA isolation | | 66 |
| | 3.2.1 | Cell lysis | 66 |
| | 3.2.2 | DNA isolation | 69 |
| | | 3.2.2.1 Method 1 | 70 |
| | | 3.2.2.1 Method 2 | 72 |
| | 3.2.3 | Storage of DNA | 72 |
| | 3.2.4 | Large scale plasmid preparation | 72 |

Page No.


| | | |
|---|---|---|
| 3.3 | DNA purity | 73 |
| | 3.3.1 Spectrophotometric determination | 74 |
| | 3.3.2 Burton diphenylamine assay | 74 |
| | 3.3.3 Agarose gel electrophoresis | 75 |
| 3.4 | Determination of % G+C content of DNA | 76 |
| | 3.4.1 Sample preparation | 76 |
| | 3.4.2 Calculation of % G+C content | 77 |
| 3.5 | Preparation of soluble protein cell-free extracts | 77 |
| | 3.5.1 Cell culture and lysis | 77 |
| | 3.5.2 Determination of protein concentration | 78 |
| 3.6 | Polyacrylamide gel electrophoresis | 79 |
| | 3.6.1 Solutions | 79 |
| | 3.6.2 Gel formation | 82 |
| | 3.6.3 Sample preparation | 83 |
| | 3.6.4 Electrophoresis | 83 |
| | 3.6.5 Gel staining | 84 |
| 3.7 | Densitometry of stained gels | 85 |
| 3.8 | Analysis of densitometer traces | 86 |
| | 3.8.1 Normalisation | 86 |
| | 3.8.2 Peak search and position analysis | 87 |
| | 3.8.3 Correlation coefficient analysis | 89 |
| | 3.8.4 Matrix sorting | 90 |
| 3.9 | Preparation of labelled DNA | 91 |
| | 3.9.1 *In vivo* labelling | 91 |
| | 3.9.2 *In vitro* labelling | 91 |
| 3.10 | DNA-DNA hybridisation | 93 |

Page No.

| | | |
|---|---|---|
| | 3.10.1 Multi-blot method | 94 |
| | 3.10.2 Individual filter method | 96 |
| | 3.10.3 Use of formamide | 98 |
| | 3.10.4 Filter melts | 98 |
| 3.11 | DNA restriction analysis | 99 |
| 3.12 | Plasmid detection | 100 |
| | 3.12.1 In gel lysis method | 100 |
| | 3.12.2 Alkali lysis method | 101 |
| | 3.12.3 Conjugative transfer of plasmids | 102 |
| 3.13 | Isolation of ribosomal RNA | 103 |
| | 3.13.1 Bulk RNA isolation | 103 |
| | 3.13.2 Separation of rRNAs | 104 |
| | 3.13.3 Electroelution | 105 |
| 3.14 | Blotting and hybridisation of restricted DNA fragments | 105 |
| | 3.14.1 Transfer of DNA to nitrocellulose | 105 |
| | 3.14.2 Hybridisation of filters | 106 |
| 3.15 | DNA-rRNA Hybridisation | 106 |
| | Results and Discussion | |
| 4. | DNA Isolation and Base Composition | 107 |
| 4.1 | Cell lysis | 107 |
| 4.2 | DNA purification | 108 |
| 4.3 | DNA base composition | 109 |

Page No.


| | | | Page No. |
|---|---|---|---|
| 5. | PAGE Analysis of Soluble Proteins | | 113 |
| 5.1 | Preparation of soluble protein extracts | | 113 |
| 5.2 | Gel system | | 113 |
| | 5.2.1 | Staining methods | 113 |
| | 5.2.2 | Gradient gels | 113 |
| | 5.2.3 | Linear gels | 114 |
| | 5.2.4 | Protein molecular weight markers | 114 |
| 5.3 | Reproducibility | | 116 |
| | 5.3.1 | Scanning | 116 |
| | 5.3.2 | Preparation of soluble proteins | 117 |
| | 5.3.3 | Culture medium | 117 |
| | 5.3.4 | Culture age | 119 |
| | 5.3.5 | Gels | 122 |
| | 5.3.6 | Staining | 125 |
| 5.4 | Peak search and position analysis | | 125 |
| | 5.4.1 | COMP 4 | 125 |
| | 5.4.2 | DENDO | 128 |
| | 5.4.3 | AUTO CA | 128 |
| 5.5 | Correlation coefficient analysis | | 131 |
| 5.6 | Discussion | | 133 |
| 6. | Nucleic Acid Hybridisation Analyses | | 139 |
| 6.1 | DNA-DNA hybridisation analysis | | 139 |
| | 6.1.1 | DNA purity and denaturation | 139 |
| | 6.1.2 | DNA filter retention | 147 |

Page No.

6.1.3 Radiolabelled probes 151

6.1.4 Hybridisation conditions 152

6.1.5 Comparison of results 159

6.1.6 DNA hybrid dissociation 162

6.1.7 Taxonomic implications 166

6.2 DNA-rRNA hybridisation 168

6.3 DNA-rDNA hybridisation 170

7. Restriction Analysis and Detection of Plasmids in the PPFMs 173

7.1 Restriction Analysis 173

7.1.1 Hybridisation of restricted DNAs 175

7.2 Detection of plasmids 177

7.2.1 In-gel lysis method 177

7.2.2 Alkali method 177

8. Discussion and Conclusions 180

9. References 189

10. Appendix: Computer Programs 214

List of Figures

Page No.

| | | |
|---|---|---|
| Figure 1 | Route for the oxidation of C1 substrates | 6 |
| Figure 2 | The ribulose monophosphate pathway | 6 |
| Figure 3 | The serine pathway | 8 |
| Figure 4 | The ribulose bisphosphate pathway | 9 |
| Figure 5 | Pathways of glucose metabolism in the PPFMs | 15 |
| Figure 6 | The serine pathway in the PPFMs | 20 |
| Figure 7 | The electron transport chain proposed for *Pseudomonas* AM1 | 30 |
| Figure 8 | Model for the genes coding for C1 specific functions in strain XX | 30 |
| Figure 9 | Protein profiles for PPFM strains cultured on various substrates | 121 |
| Figure 10 | Protein profiles for PPFM strains harvested at various culture turbidities | 121 |
| Figure 11 | Soluble protein extracts from representatives of the PPFMs | 124 |
| Figure 12 | Peak search and positioning on a PPFM electrophoretogram | 126 |
| Figure 13 | Comparison of PPFM soluble proteins using program COMP 4 | 127 |
| Figure 14 | Dendrogram after AUTO CA analysis of PPFM electrophoretograms | 130 |
| Figure 15 | Dendrogram after AUTO EF analysis of PPFM electrophoretograms | 132 |

| | | Page No. |
|---|---|---|
| Figure 16 | Dendrogram after AUTO EF analysis of PPFM and *Rhodospirillaceae* electrophoretograms | 136 |
| Figure 17 | Dendrogram after DNA homology studies on the PPFMs | 160 |
| Figure 18 | Correlation between DNA homology and $-\Delta Tm_{(e)}$ values | 164 |
| Figure 19a | Thermal denaturation profiles for PPFM DNA duplexes | 165 |
| Figure 19b | Elution profiles for PPFM DNA duplexes | 165 |
| Figure 20 | Plasmid screening by the in-gel lysis method | 178 |
| Figure 21 | Correlation between DNA homology and PAGE-protein data | 169 |
| Figure 22 | Restriction endonuclease fragment patterns of PPFM DNAs | 174 |
| Figure 23 | Autoradiograph of restricted PPFM DNAs hybridised against a probe from pCOI | 176 |

Page No.


| | | |
|---|---|---|
| Figure 16 | Dendrogram after AUTO EF analysis of PPFM and *Rhodospirillaceae* electrophoretograms | 136 |
| Figure 17 | Dendrogram after DNA homology studies on the PPFMs | 160 |
| Figure 18 | Correlation between DNA homology and $-\Delta Tm_{(e)}$ values | 164 |
| Figure 19a | Thermal denaturation profiles for PPFM DNA duplexes | 165 |
| Figure 19b | Elution profiles for PPFM DNA duplexes | 165 |
| Figure 20 | Plasmid screening by the in-gel lysis method | 178 |
| Figure 21 | Correlation between DNA homology and PAGE-protein data | 169 |
| Figure 22 | Restriction endonuclease fragment patterns of PPFM DNAs | 174 |
| Figure 23 | Autoradiograph of restricted PPFM DNAs hybridised against a probe from pCOI | 176 |

List of Tables


| | | Page No. |
|---|---|---|
| Table 1 | Differences between major PPFM clusters | 14 |
| Table 2 | C1 substrate range of PPFMs | 16 |
| Table 3 | PPFM strains used in this study | 64 |
| Table 4 | Non-PPFM strains used in this study | 65 |
| Table 5 | Base composition of PPFM DNAs | 112 |
| Table 6 | Effect of growth substrate on protein profile comparison | 118 |
| Table 7 | Effect of growth cycle stage on protein profile comparison | 120 |
| Table 8 | DNA hybridisation and $-\Delta Tm_{(e)}$ results | 140 |
| Table 9 | DNA retention by membrane filter types | 149 |
| Table 10 | Retention of DNA in various solvents | 150 |
| Table 11 | Quenching of radioactivity by components of the hybridisation mixture | 153 |
| Table 12 | Homologous reassociation and various hybridisation conditions | 155 |
| Table 13 | Hybridisation of PPFM DNAs at 53°C and 60°C | 156 |
| Table 14 | Hybridisation with members of the *Enterobacteriaceae* | 158 |

Declaration

The work contained in this thesis is the result of original research conducted by myself under the supervision of Dr. C. S. Dow, unless otherwise stated. All sources of information have been acknowledged by means of reference.

None of the work contained in this thesis has been submitted for any previous degree.

Derek Hood

Derek Hood

## Acknowledgements

I would like to sincerely thank my supervisors Dr. C. S. Dow, Dr. P. N. Green and Dr. D. M. Gibson for their help and encouragement over the last three years, and the Science and Engineering Research Council for financing the research. Thanks also to J. De Ley, M. Gillis and R. Tijtgat for advice on nucleic acid hybridisation techniques and to Melvyn Whiteside for his patience and help in computer programming and the many enjoyable hours of badminton.

My thanks also to all members, past and present, of the Microbiology II Lab for the pleasant working atmosphere and their lasting friendship. Special thanks to Hazel for her support during the composition of this thesis.

Finally, my thanks to Miss Kay Pollitt for her excellent typing of the manuscript.

## Abbreviations

| | | |
|---|---|---|
| BSA | - | bovine serum albumen |
| Cm | - | chloramphenicol |
| cpm | - | counts per minute |
| DMSO | - | dimethylsulphoxide |
| DNA | - | deoxyribonucleic acid |
| dATP | - | deoxyadenosine triphosphate |
| dCTP | - | deoxycytidine triphosphate |
| dGTP | - | deoxyguanine triphosphate |
| dTTP | - | deoxythymidine triphosphate |
| EDTA | - | ethylenediaminetetra acetic acid |
| GP | - | glycerol peptone |
| kb | - | kilobase |
| M | - | molar |
| mM | - | millimolar |
| mol | - | mole |
| μCi | - | microcurie |
| MS | - | mineral salts |
| NB | - | nutrient broth |
| PAGE | - | polyacrylamide gel electrophoresis |
| RNA | - | ribonucleic acid |
| RNAase | - | ribonuclease |
| rRNA | - | ribosomal ribonucleic acid |
| SDS | - | sodium dodecyl sulphate |
| SSC | - | standard saline citrate |
| Tc | - | tetracycline |
| TEMED | - | N,N,N,N'-tetramethylethylenediamine |
| Tris | - | tris (hydroxymethyl) aminomethane |
| v/v | - | volume to volume |
| w/v | - | weight to volume |

Summary

Molecular biology techniques were employed to help resolve the taxonomic confusion associated with a group of bacteria, the pink-pigmented facultative methylotrophs (PPFMs). One technique was the computer assisted comparison of traces from bacterial soluble protein extracts electrophoresed through polyacrylamide then scanned with a microdensitometer. A reliable quantitative system was developed for the automated analysis of the protein extracts by two fundamentally different modes. These were comparison of protein profiles by peak search and position analysis of coincident bands, and correlation coefficient analysis of the complete trace contour.

A second technique involved hybridisation of the DNAs isolated from a number of PPFM strains. A multi-blot system was developed for efficient analysis of multiple DNA samples and various parameters affecting filter preparation and the hybridisation reaction were investigated to optimise conditions. The DNA homology data was compared with that obtained after the protein analysis and from previous investigations and the relative merits of each technique were discussed.

Also investigated were the base composition of the PPFM DNAs, the prevalence of indigenous plasmids and attempts were made towards a phylogenetic analysis of the PPFMs by DNA-rRNA and DNA-rDNA hybridisations. The results suggested that the PPFMs comprise a series of closely related species groups and single member species related at a taxonomical level approximating to that of the genus. The most appropriate generic assignment may be the genus *Methylobacterium*. Evidence was discussed suggesting that the PPFMs may be most closely related, in an evolutionary sense, to the purple non-sulphur bacteria.

# Introduction

## 1.1 General Introduction

Methylotrophs are microorganisms defined by their ability to utilise as the sole carbon, and often energy, source for growth, carbon compounds more reduced than $CO_2$ and which contain no carbon-carbon bonds. Such compounds are termed C1 compounds. Obligate methylotrophs grow only on such compounds whereas facultative methylotrophs can grow additionally on a wide variety of multicarbon compounds.

Methane, methanol, formate and the N-methyl compounds (methylamines) are the most common reduced C1 substrates formed in the natural environment. Of these methane, produced during the decomposition of organic matter, is the most abundant. Methanol, formed by the oxidation of methane (Figure 1) and the breakdown of structural components of plant materials, is the next most widespread, but only accumulates to significant levels in specialised environments. Methylamines and formate, again produced from biological degradation processes, are found at levels equivalent to those of methanol but only the N-methyl compounds form readily utilisable substrates. Formaldehyde is found only as an intermediate in the oxidation of methanol and rarely accumulates in nature. With such an abundance, it is not surprising that a considerable number of microorganisms have developed an ability to utilise these C1 compounds for growth. In addition to the substrates mentioned above, bacteria able to grow on carbon monoxide (CO), i.e. carbodoxybacteria (Colby *et al.*, 1979; Whittenbury and Dalton, 1981) have also been considered as methylotrophs, as they obtain energy by oxidising CO to $CO_2$ and then assimilate the $CO_2$ to form cellular carbon.

It is important, ecologically, that reduced C1 compounds are utilised to help maintain the carbon cycle but the real contribution of methylotrophic processes has yet to be fully determined. The first well described methylotroph was reported as early as 1906 by Sohngen (*Bacillus methanicus*) but only in the last 20-25 years has any concerted effort been made to isolate and study these organisms. The burst of activity in the field of C1 metabolism in recent years, much of which has been stimulated by potential commercial exploitation of these organisms has resulted in the isolation of a whole spectrum of new strains and to the reinvestigation of strains previously identified.

## 1.2 Categorisation

C1 utilising organisms can be grouped and categorised under different criteria varying with the bias of the investigator. One such general scheme, based on the obligateness of the requirement for C1 substrates, divides the methylotrophs into 3 major groups. These are:

1. Obligate methane utilisers - methanotrophs
2. Obligate methanol (methylamine) utilisers
3. Facultative methanol (methylamine) utilisers

### 1.2.1 The Obligate Methane Utilisers

The obligate methane utilisers, or methanotrophs, have been the focal point of much of the recent research on methylotrophs. The biology of these organisms has been reviewed in depth by Colby *et al.*, 1979; Hanson, 1980; Higgins *et al.*, 1981 and Anthony, 1982. Most

methanotrophic bacteria are obligate for reduced C1 substrates and are obligately aerobic gram negative rods, vibrios or cocci. Whittenbury and his colleagues isolated more than 100 of these organisms (Whittenbury *et al.*, 1970) and formed the basis of the current classification into Type I (*Methylomonas*, *Methylobacter*, *Methylococcus*) and Type II (*Methylosinus*, *Methylocystis*) strains. This division is based, largely, upon differences in intracytoplasmic membrane arrangement, carbon assimilation pathways and the type of resting cell produced (Galchenko and Andreev, 1984). Recently, the suggestion has been made (Whittenbury and Dalton, 1981) that the rather atypical *Methylococcus capsulatus* (Bath) be removed from group I and placed into a third major group, type X.

The only facultative members of the methanotrophs are strains of the genus *Methylobacterium*, accommodated as a sub-group of Type II organisms (Patt *et al.*, 1976; Patel *et al.*, 1978). Recent evidence, however, questions the validity of this grouping (Hanson, 1980; this thesis).

1.2.2 The Obligate Methanol Utilisers

The obligate methanol utilisers form the smallest of the three major groups of methylotrophs. Some members can utilise N-methyl compounds in place of methanol. All are gram-negative, aerobic, non-spore forming rods which superficially resemble the pseudomonads. These are the most recent group of methylotrophic bacteria to emerge and have been studied largely because of interest from industry in their high potential yields in single cell protein (SCP) production. Taxonomic studies on this group of organisms have just been completed (Byrom, 1981; Jenkins *et*

*al.*, 1984) and indicate that members of this group are distinct from all other methylotrophs and pseudomonads. At least two distinct taxa are thought to be evident within this group.

### 1.2.3 The Facultative Methanol Utilisers

The facultative methanol utilisers are the most disparate group of methylotrophs sharing certain metabolic pathways involved in C1 metabolism, but differing radically in other properties. Many members of this group are also facultative with respect to methylamine (or higher N-methyl amines) and occasionally formate. All are capable of heterotrophic growth on more complex organic carbon sources. This diverse assemblage of organisms has been grouped into a wide variety of taxa encompassing both gram-negative and gram-positive genera. Examples are: *Pseudomonas*, *Vibrio*, *Bacillus*, *Protaminobacter*, *Achromobacter*, *Hyphomicrobium*, *Methylobacterium*, *Mycoplana*, *Streptomyces* and *Corynebacterium*. Without doubt, some organisms have been correctly placed into established, well characterised taxa such as *Streptomyces*, *Bacillus* and *Hyphomicrobium* but correspondingly, many of the 'pseudomonad type' organisms have been mis-named and placed in taxa which are not correct. The gram-negative pink-pigmented facultative methylotrophs (PPFMs), exemplified by *Pseudomonas extorquens*, form a sizeable part of this group and are the subject of this study. These organisms will be discussed in greater detail later (1.5).

## 1.3 Carbon Metabolism

A common link between all methylotrophic organisms is the remarkable similarity in their pathways of reduced C1 compound utilisation. In most instances, growth on C1 compounds does not involve the central catabolic pathways normally associated with the metabolism of aerobes. Prior to incorporation C1 substrates must first be oxidised by a series of special oxidation steps as shown in Figure 1. Formaldehyde occupies a central position in this scheme being a common intermediate for subsequent assimilation into cellular materials (Attwood and Quayle, 1984). Assimilation of C1 compounds is by one of three major pathways as described below.

### 1.3.1 Ribulose Monophosphate Pathway

The Ribulose Monophosphate Pathway (RMP) (Figure 2) allows a condensation reaction between the C1 moiety (formaldehyde) and a C5 acceptor molecule (ribulose monophosphate) to produce fructose-6-phosphate. This C6 molecule is then split to produce two C3 compounds, the exact nature of which varies between organisms. At least four variant mechanisms are known for this stage of the pathway (Anthony, 1982; Dijkhuizen, 1983). Finally, two molecules of fructose-6-phosphate and one molecule of glyceraldehyde are used to regenerate three molecules of ribulose-5-phosphate.

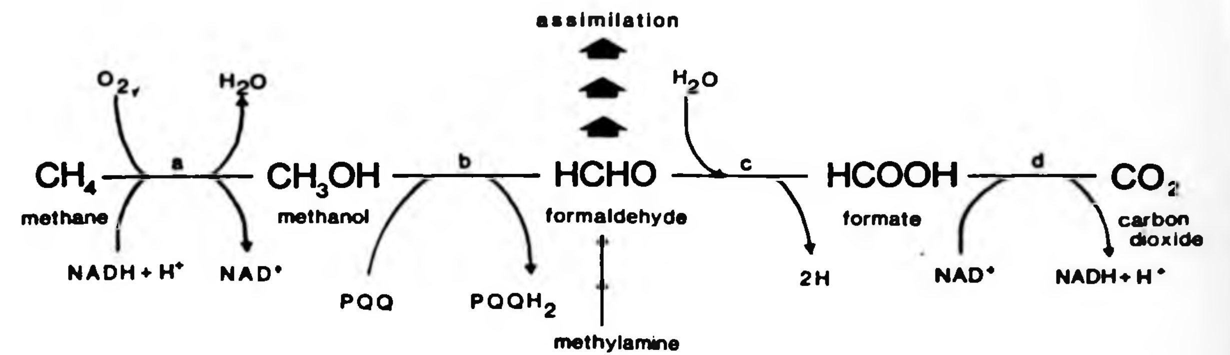
This pathway is favoured by type I (and type X) methanotrophs, obligate methanol utilisers and a minority of facultative methylotrophs (Zatman, 1981).

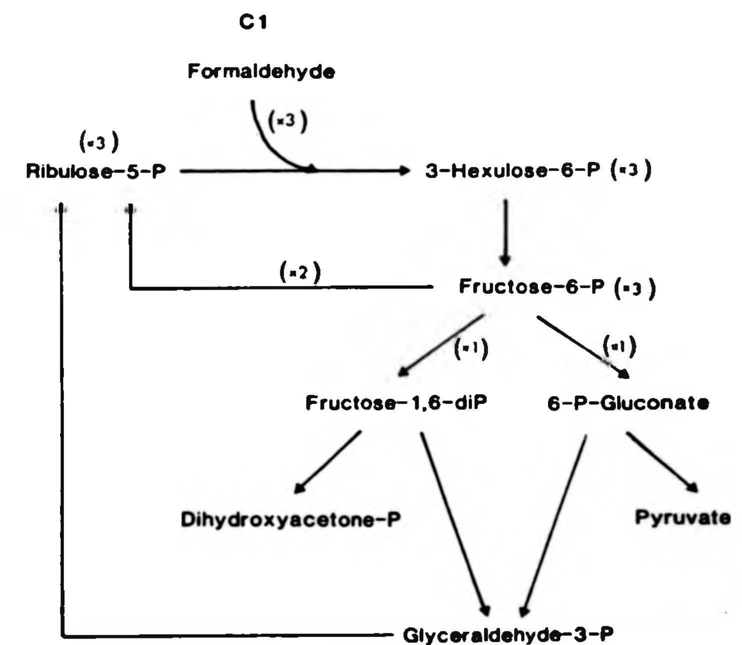
Figure 1 Route for the Oxidation of C1 Substrates

a, methane monooxygenase; b, methanol dehydrogenase ; c, formaldehyde dehydrogenase; d, formate dehydrogenase; NAD, nicotinamide adenine dinucleotide; PQQ, pyrroloquinoline quinone.

Figure 2 The Ribulose Monophosphate Pathway for C1 Assimilation

Incorporating two variant routes of glyceraldehyde-3-P formation; a, 3-hexulosephosphate synthase; b, phospho-3-hexuloisomerase; c, 6-phosphofruktokinase; d, glucose phosphate isomerase; e, fructose diphosphate aldolase; f, 6-phosphogluconate dehydratase.



### 1.3.2 Serine Pathway

The serine pathway was the first C1 assimilatory pathway to be elucidated in methylotrophs (Large and Quayle, 1963). The C1 moiety (formaldehyde) reacts with a C2 acceptor molecule (glycine) to form serine (Figure 3). Serine is converted by several enzymatic steps and carboxylation with $CO_2$ to the C4 compound malate. Malate is then split into two C2 compounds, glyoxylate and acetyl-CoA and the cycle continues. The main variation between organisms is in their mechanisms for the oxidation of acetyl-CoA to glyoxylate to regenerate a second glycine acceptor molecule. Figure 3 shows the mechanism found in some methylotrophs where isocitrate lyase and some TCA cycle enzymes catalyse this regeneration (- $ICl^+$ strains). Many methylotrophs favour a more direct conversion of acetyl-CoA to glyoxylate ($ICl^-$ strains), often via homocitrate and homoisocitrate (Kortstee, 1980, 1981). The serine pathway for C1 assimilation is favoured by type II methanotrophs and the majority of facultative methylotrophs.

### 1.3.3 Ribulose Bisphosphate Pathway

Discussion has been limited so far to organisms which convert the C1 substrates to, and assimilate them at, the level of formaldehyde. Some C1 utilising microorganisms, however, oxidise C1 growth substrates to $CO_2$ (often producing energy) then assimilate the $CO_2$ so formed by the Ribulose Bisphosphate (RBP) pathway or Calvin cycle (Quayle and Ferenci, 1978; Colby *et al.*, 1979) (Figure 4). This pathway is favoured by a minority of facultative methylotrophs growing on formate (or methanol) but may also be present in some obligate methanotrophs such as *Methylococcus capsulatus* (Bath) where it functions in some minor

Figure 3 The Serine Pathway for C1 Assimilation

Isocitrate lyase variant.

a, serine transhydroxymethylase; b, serine glyoxylate amino transferase.

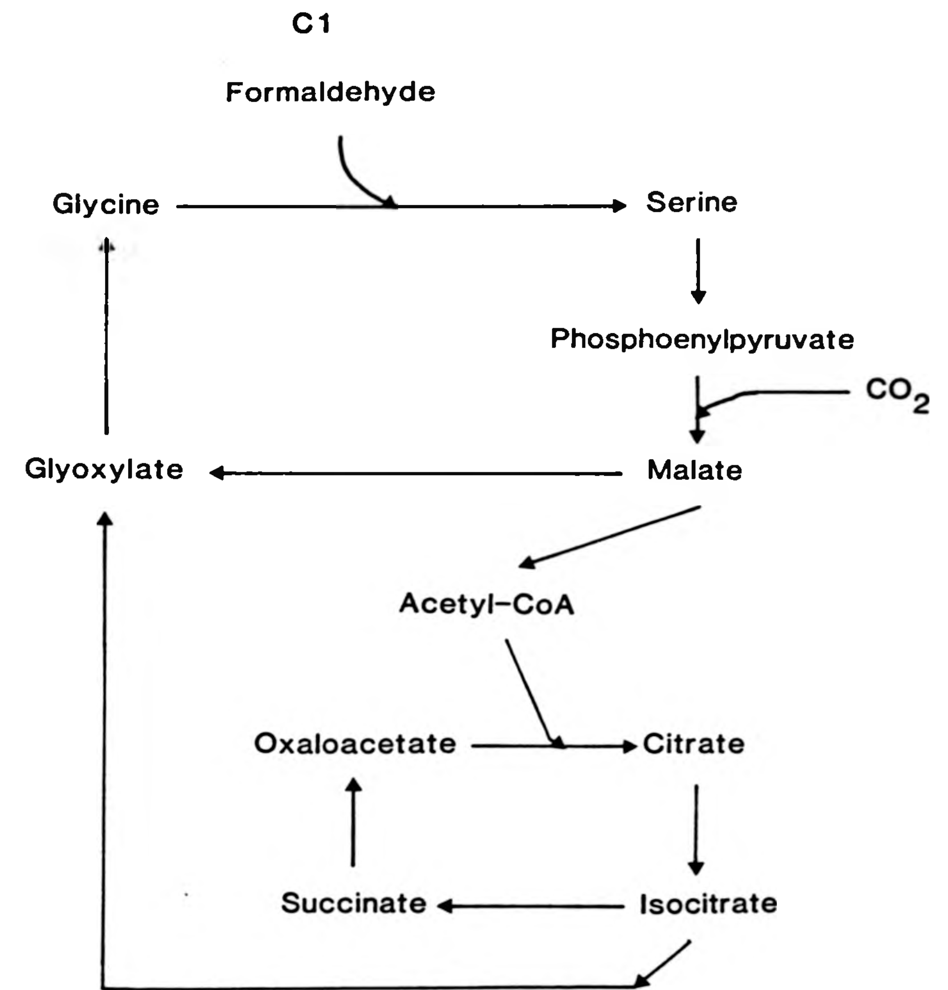
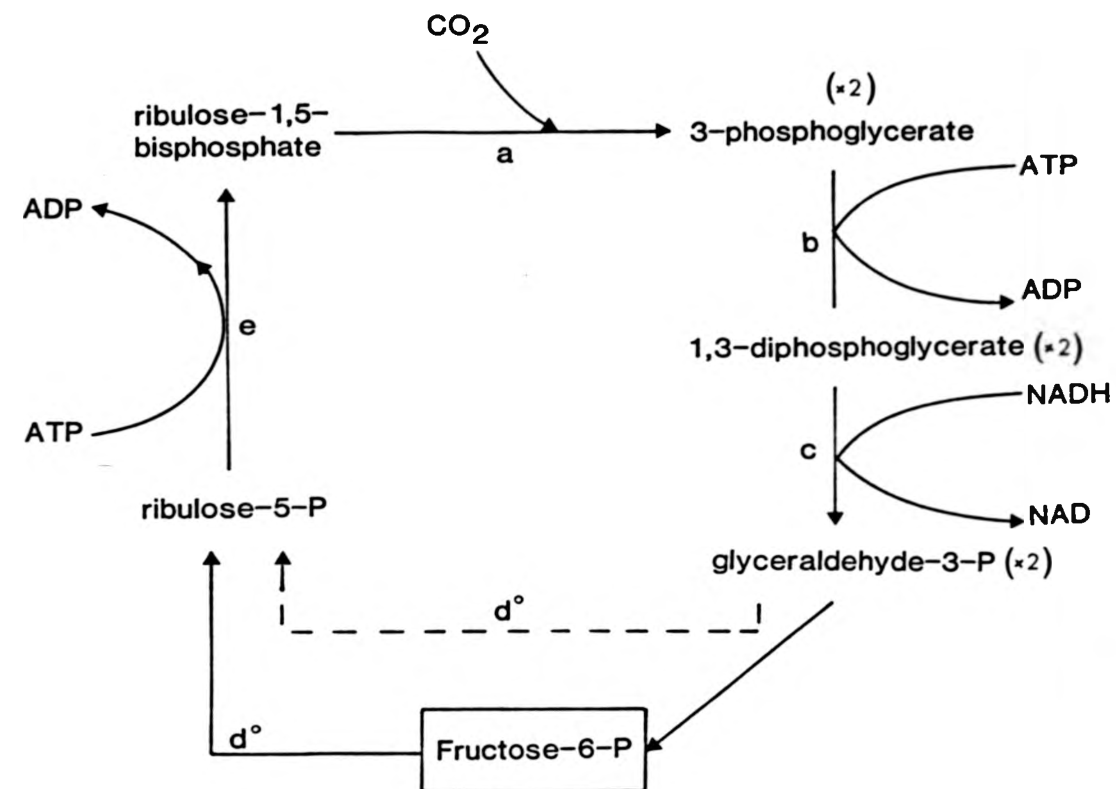


Figure 4 The Ribulose Bisphosphate Pathway for $CO_2$ Assimilation

The variants of the formation of ribulose-5-phosphate from three and six carbon sugars are not detailed.

a, ribulose bisphosphate carboxylase; b, phosphoglycerate kinase; c, glyceraldehyde phosphate dehydrogenase; d, variety of transketolase, aldolase and other enzymes; e, phosphoribulokinase.



capacity (Stanley and Dalton, 1982). The RBP pathway is the pathway of $CO_2$ assimilation found in most photo- and chemolithoautotrophs. Many such organisms appear to have evolved extra nutritional versatility by the acquisition of the few enzymes required to oxidise C1 substrates such as formate, methanol (and CO) to $CO_2$. Zatman (1981) has suggested that these organisms be termed pseudo-methylotrophs.

### 1.3.4 Energy Metabolism

Of the three C1 assimilation pathways, studies have shown that the RMP pathway is less energetically stressful to the cell than the serine pathway, which is less energy demanding than the RBP pathway (Quayle and Ferenci, 1978). It is very difficult to assess the relevance of these findings in isolation without full knowledge of possible ulterior benefits to the cell provided by each pathway. In general, methylotrophs are considered to have, exclusively, only one of the three assimilation pathways but recent work has indicated that some methylotrophs may possess two, or even all three, occupying different roles under varying growth conditions.

The net effect of each C1 utilisation pathway is the conversion of reduced C1 substrate to a variety of central metabolic intermediates, the majority of carbon being further assimilated at the C2 or C3 level. Studies on the concommitant mechanism of ATP formation (Anthony, 1981; Anthony, 1982; Zatman, 1981) indicate that ATP is produced by the reoxidation of NAD(P)H and the prosthetic groups of dehydrogenases reduced during C1 oxidation processes. This is by way of electron transport chains similar to those found in most aerobic bacteria.

## 1.4 Methylotrophic Yeasts

The utilisation of reduced C1 compounds is not a property restricted to bacteria. Eukaryotes, mainly yeasts and a few species of mycelial fungi, are also capable of growth on C1 compounds, predominantly methanol.

All methylotrophic yeasts divide by budding. Most require vitamins as cofactors for growth and have an optimum temperature for growth of 28°C. Representatives of only six or seven of the thirty-nine genera of yeasts exhibit methylotrophic capabilities. A few yeasts have been reported to be able to grow on methane but generation times are in excess of 48 hrs indicating that this property can be of litle importance in nature (Wolf and Hanson, 1979).

The C1 oxidative metabolism of yeast differs in many ways from that in bacteria. Methanol oxidation to formaldehyde is compartmentalised in specialised peroxisomes, with the subsequent oxidation to $CO_2$ and assimilation occurring in the cytoplasm. Unlike bacteria, the initial oxidation of methanol is catalysed by an alcohol oxidase. Assimilation of methanol into cellular carbon is at the level of formaldehyde by the dihydroxyacetone (DHA) cycle (= xylulose monophosphate (XMP) pathway) and requires only one unique enzyme for the process.

## 1.5 Pink Pigmented Facultative Methylotrophs

As indicated previously (1.2.3) the facultative methylotrophs exhibit the greatest physiological and ecological diversity of all

methylotrophic groups. This present study is concerned with representatives of the sub-group of pink-pigmented facultative methanol utilisers (PPFMs). Members have been studied previously in isolation but only recently have they been studied as a group of potentially related organisms. Historically, studies have concentrated on their biochemistry at the expense of any concerted effort to explore their taxonomy. Consequently, as expressed by Quayle in 1972, the current position on the taxonomy and nomenclature of the PPFMs is confused. It is one of the main purposes of this study to help resolve this enigma.

Some comparative studies have been made on the PPFMs (Stocks and McCleskey, 1964; Kouno and Ozaki, 1975; Urakami and Komagata, 1979, 1981) but it was not until the study of Green and Bousfield (1982) that much of the available comparative data was obtained. Green and Bousfield found that 150 strains of PPFMs fell into two closely related clusters. These clusters were distinct from all non-pigmented facultative methanol utilising and other reference strains used.

### 1.5.1 Characterisation

All PPFM strains are gram negative or gram variable rods, 1-8 μm in length and often are pleomorphic with branching or L-shaped cell forms. Cells are motile, by polar or sub-polar flagella, and have sudanophilic inclusion bodies. Colonies on agar are pink to red in colour with a diameter of <1-3 mm. In broth, a pellicle and/or ring is formed and turbidity is rarely uniform. All strains are catalase and urease positive and no growth factors are required. Most strains can grow over the temperature range of 15-37$^{o}$C and over a pH range of between 5 and 9,

the maximum growth rate occurring between 25 and 30°C. Some intra- and inter-group variability is found when considering the physiological growth ranges and also the sensitivities to antibiotics. In addition to methanol and formate, most strains can utilise as the sole carbon source: glycerol, malonate, succinate, fumarate, 2-oxoglutarate, lactate, malate, pyruvate, ethanol, propylene, glycol and ethylamine. DNA base compositions are in the range of 60 to 70 moles % (G + C) (Green and Bousfield, 1982). A problem found during this work was that all PPFMs were negative with the majority of diagnostic tests commonly applied to other bacterial groups (P. Green, personal communication). Consequently, features such as substrate utilisation ranges had to be examined in depth to allow for discrimination between strains. Table 1 lists the main features used to differentiate between the two PPFM clusters. Cluster A corresponds to the major cluster 10 of Green and Bousfield (1982) and Cluster B to their minor cluster 11.

### 1.5.2 Carbohydrate Metabolism

Members of Cluster B are nutritionally more versatile and can utilise carbohydrates and carbohydrate derivatives not utilised by strains of Cluster A. Some recent work (Green and Gibson, 1984) investigated possible carbohydrate utilisation pathways within the PPFMs and any enzymatic lesions which might help explain the inability of group A organisms to grown on aldose sugars such as glucose. Complete elucidation of these pathways requires further study, but a scheme for glucose metabolism was proposed as shown in Figure 5. All of the PPFMs examined had the key enzymic activities associated with the Entner-Doudoroff Pathway (but not the analogous enzymes of the KDG variant

Table 1 Features which Differentiate between clusters A and B (from Green and Bousfield, 1981)

| Character | A | B |
|---|---|---|
| Utilisation of | | |
| L-Arabinose | - | + |
| D-Xylose | - | + |
| D-Fucose | - | + |
| D-Glucose | - | + |
| D-Galactose | - | + |
| 2-Deoxyglucose | - | + |
| L-Arabonate | - | + |
| L-Arabonic acid lactone | - | + |
| D-Gluconate | - | + |
| D-Galactonate | - | + |
| DL-Aspartate | - | + |
| L-Glutamate | V | + |
| Adipic Acid | - | + |
| Sebacic Acid | - | + |
| Citrate | - | + |
| Citraconate | - | + |
| Saccharate | - | + |
| Suberate | - | + |
| Pimelate | - | + |
| Azeleate | - | + |
| Adipate | - | + |

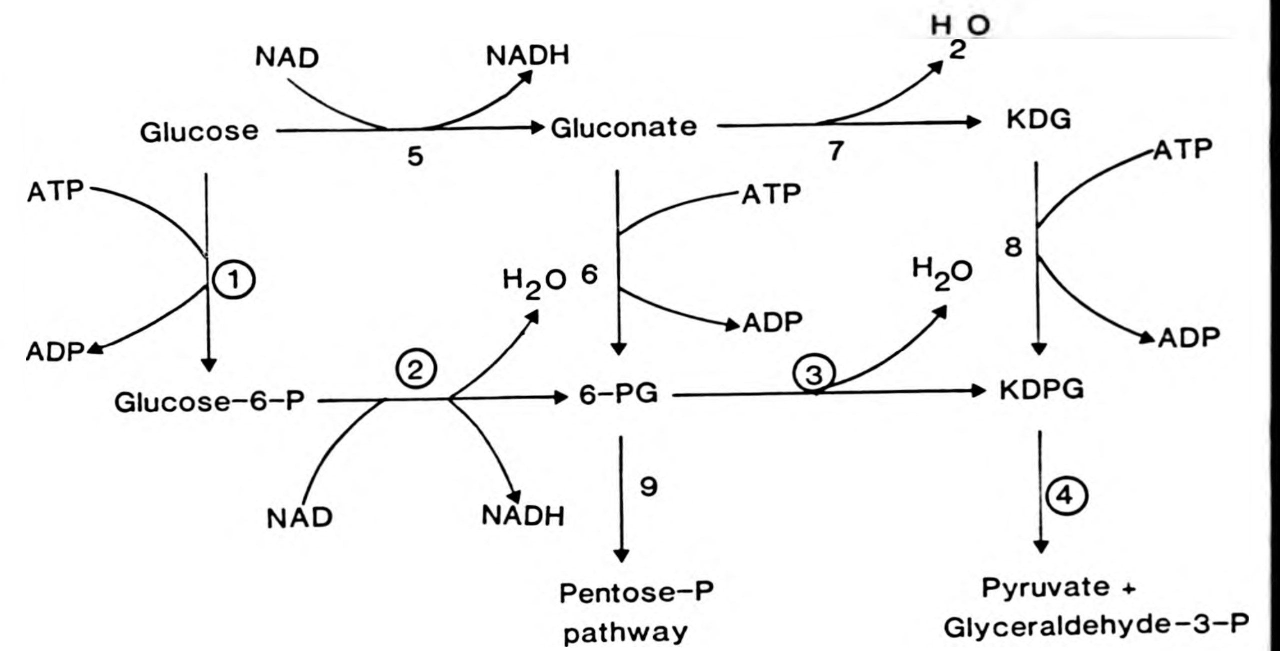
\+ = >90% of strains positive
\- = >90% of strains negative
V = variable result.

Figure 5 Proposed Pathways of Glucose Metabolism in some PPFMs

(from Green and Gibson, 1984)

1, glukokinase; 2, glucose-6-P dehydrogenase; 3, 6-P-gluconate dehydratase; 4, KDPG aldolase; 5, glucose dehydrogenase; 6, glucokinase; 7, gluconate dehydratase; 8, KDG kinase; 9, 6-P-gluconate dehydrogenase; KDG, 2-keto-3-deoxygluconate; KDPG, 2-keto-3-deoxyphosphogluconate; 6PG, 6-P-gluconate.

O steps of the Entner Doudoroff pathway.



pathway). Non-glucose utilising members of cluster A must use this pathway solely for the production of metabolic intermediates. The inability to utilise glucose for growth may lie with their inability to take up aldoses, i.e. the absence of a permease system (Green and Gibson, 1984). The absence of FBP aldolase indicates that the glycolytic pathway is probably absent in the strains examined. The presence of 6-P-glucose and its dehydrogenase in 2 strains, *Methylobacterium organophilum* strain XX (NCIB 11278) and strain 58, indicates that at least part of the pentose phosphate cycle is present.

### 1.5.3 Carbon Assimilation

Of the 150 PPFM strains used in the study of Green and Bousfield (1982), virtually all can utilise methanol and formate, methylotrophically, as the sole carbon and energy source. None can use methane. Table 2 lists the range of C1 substrates shown to be utililsed by each of the main PPFM clusters.

Table 2 Utilisation of C1 Compounds by the PPFMs

| Sole Carbon Source | Cluster A | Cluster B |
|---|---|---|
| Methanol | + | + |
| Methylamine | + | v |
| Dimethylamine | v | v |
| Trimethylamine | v | - |
| Trimethylamine-N-oxide | - | - |
| Formate | + | + |
| Formamide | v | v |

+, >90% of strains positive; -, <10% of strains positive; v, 10-90% of strains positive.

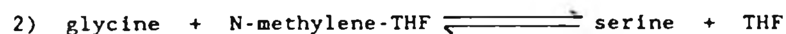
All members of the PPFMs so far examined have been found to convert the Cl substrates to formaldehyde to be assimilated by an ICl$^-$ variant of the serine pathway. This pathway must provide precursors for the biosynthesis of all cell components from the Cl substrates (Large and Quayle, 1963; Colby *et al.*, 1979; Zatman, 1981; Kortstee, 1980; Hanson, 1980).

A representative of the PPFMs, namely *Pseudomonas* AM1 NCIB 9133 was the organism used by Quayle and coworkers in the original investigations of this pathway in methylotrophs. Radiotracer studies of short term incubations of *Pseudomonas* AM1 with $^{14}$C-methanol resulted in the early labelling of serine followed by glycine, malate and aspartate (Large *et al.*, 1961). Such a distribution of label was consistent with the fixation of Cl formaldehyde by a serine transhydroxymethylase enzyme. THF is tetrahydrofolate.

(chemical)

1) HCHO + THF N-methylene-THF

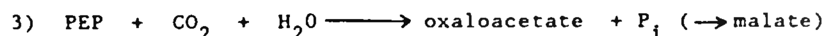
b

2) glycine + N-methylene-THF  serine + THF

a

These studies also implied the formation of malate via a C3 compound derived from serine. The involvement of folate derivatives in this pathway was also demonstrated with *Protaminobacter ruber* NCIB 2879 and *Pseudomonas (Vibrio) extorquens* NCIB 9399 using folate antagonists. Two serine transhydroxymethylase isoenzymes have been shown to be present in

*M. organophilum* XX (NCIB 11278) (O'Connor and Hanson, 1975; Hanson, 1980). Isoenzyme (a), containing a single subunit of $10^5$ molecular weight, is found in cells grown on succinate and is believed to catalyse the synthesis of glycine from serine during heterotrophic growth. An additional isoenzyme (b), of 2 x $10^5$ molecular weight, is induced during growth on methanol and appears to participate in the assimilation of C1 units produced from methanol (Hanson, 1980). This isoenzyme is activated by glyoxylate. Some circumstantial evidence suggests that two such enzymes may also occur in *Pseudomonas* AM1 (Harder and Quayle, 1971a). The glycine produced by isoenzyme (a) is required for protein and purine biosynthesis and the methylene tetrahydrofolate serves as a one carbon donor for other biosynthetic pathways. Serine enters the pathway described in Figure 6. Further radiotracer and enzymological studies on cell free extracts of *Pseudomonas* AM1 (Large *et al.*, 1962; Large and Quayle, 1963; Harder and Quayle, 1971a) confirmed many of the intervening reactions of the pathway between serine and malate. Malate formation was found to be catalysed by phosphoenolpyruvate (PEP) carboxylase.



3) $PEP + CO_2 + H_2O \longrightarrow oxaloacetate + P_i\ (\longrightarrow malate)$

The irreversible carboxylation of PEP is a critical step in the pathway. Fifty percent of carbon derived from methanol is assimilated as $CO_2$ (Large *et al.*, 1961). Two isoenzymes have been shown to exist in *M. organophilum* but the exact mechanisms of the reaction are still unclear. Harder and Quayle (1971b) also established that the serine biosynthesis pathway plays a role in amino acid replenishment.

An important unsolved problem was still the method of regeneration of the glycine acceptor molecules. Large *et al.* (1962) proposed that C2 glycine could be generated by the cleavage of a C4 compound such as malate in a cyclic type pathway. It was only after further radiotracer studies with methylamine utilising organisms by Quayle and coworkers in 1972 that the crucial C4 cleavage reaction was determined (Dunstan *et al.*, 1972).

4) Malate + ATP +CoA ⟶ glyoxylate + acetyl-CoA + ADP + Pi

Subsequent work (Hersh, 1973) showed the cleavage reaction to be separable into two components catalysed by malate thiokinase (reaction 5) and malyl-CoA lyase (reaction 6).

5) Malate + ATP + CoA ⟶ malyl-CoA + ADP + Pi

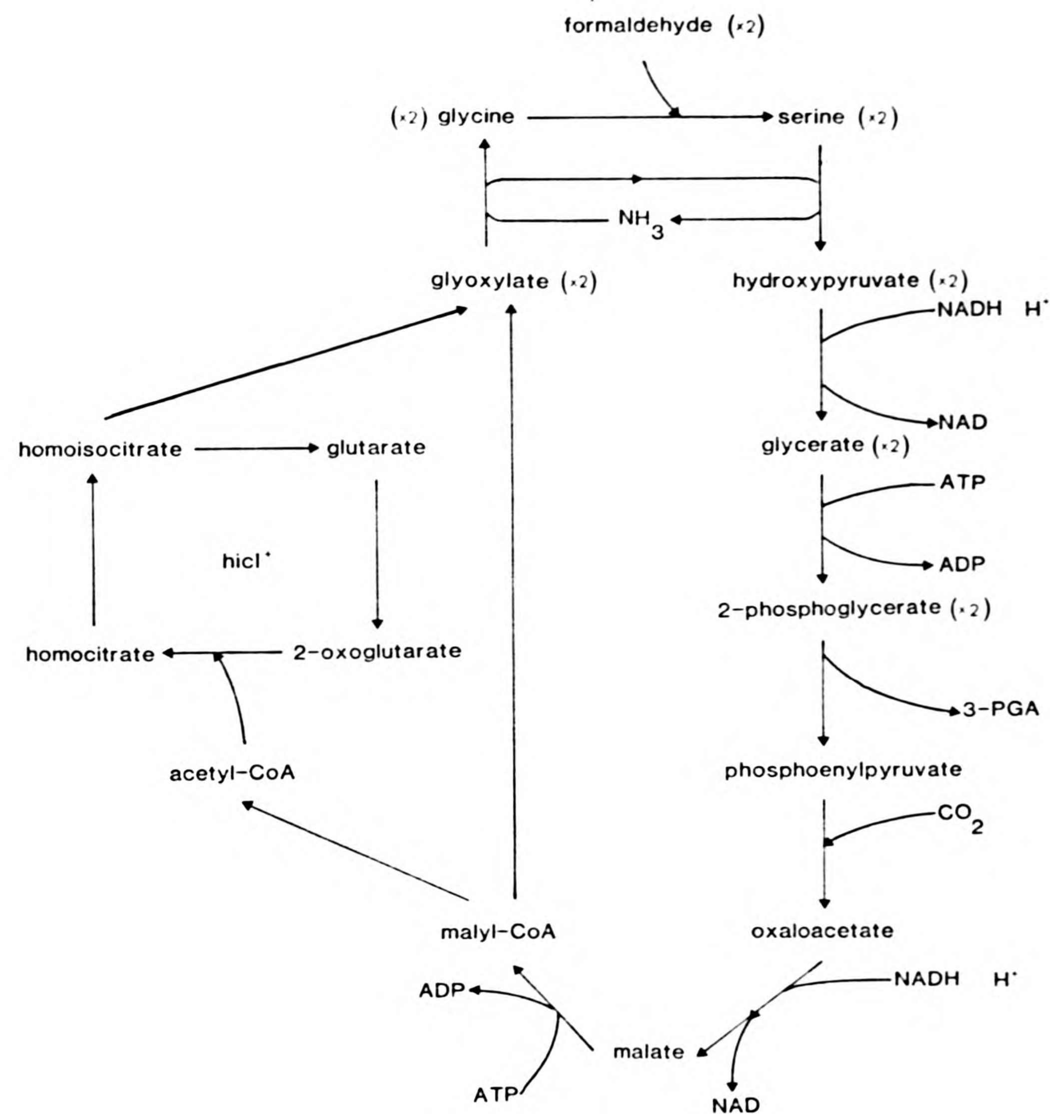
6) Malyl-CoA ⟶ glyoxylate + acetyl-CoA

In *Pseudomonas* AM1 malyl-CoA lyase but no malate thiokinase activity could be detected (Salem *et al.*, 1973). The mechanism for making malyl-CoA in the PPFMs is still unknown but mutant studies with *Pseudomonas* AM1 have confirmed the importance of this C4 cleavage reaction (Salem *et al.*, 1974). Malyl-CoA lyase in serine pathway organisms generates one C2 glyoxylate acceptor molecule but there is still the problem of how to generate a second glyoxylate molecule from acetyl-CoA. Some methylotrophs are known to employ a mechanism involving isocitrate lyase and some TCA cycle enzymes (= ICl$^{+}$ strains) as shown in figure 3 (Quayle, 1972). The PPFMs *Pseudomonas* AM1 (Large and Quayle, 1963) and

Figure 6 The Serine Pathway for C1 Assimilation Proposed for the PPFMs

a, serine transhydroxymethylase; b, serine glyoxylate; c, hydroxypyruvate reductase; d, glycerate kinase; e, phosphopyruvate hydratase; f, phosphoenol pyruvate carboxylase; g, malate dehydrogenase; h, malate thiokinase; i, malyl-CoA lyase; j, isocitrate lyase; k, homocitrate synthase; l, homoisocitrate lyase; m, glutarate dehydrogenase.

Net reaction - $CO_2$ + 2HCHO + 2NADH + $2H^+$ + 3ATP → 3 phosphoglycerate + $2NAD^+$ + 3ADP + 2 Pi + $FPH_2$.



*M. organophilum* XX (Hanson, 1980) contain no significant isocitrate lyase activity during methylotrophic growth. This ICl$^-$ pathway has been studied in detail in *Pseudomonas* AM1 and *Pseudomonas* 80 by Kortstee (1980, 1981). The pathway proposed involves the formation of glyoxylate from acetyl-CoA via homoisocitrate and is tentatively called the homoisocitrate lyase variant (hicl$^+$) of the serine pathway (Figure 6).

### 1.5.4 Regulation of Carbon Assimilation

It has been shown in *Pseudomonas* AM1 that certain serine pathway enzymes uniquely required for growth on C1 compounds (hydroxypyruvate reductase, serine glyoxylate aminotransferase, glycerate kinase) are induced in the presence of C1 substrates and repressed by growth with succinate (Large and Quayle, 1963; Dunstan *et al.*, 1972). In succinate grown cells the enzymes in question are present, if at all, at low activities being required only for the synthesis of particular cellular constituents, such as glycine and serine, and not bulk carbon assimilation. The same key enzymes have also been shown to be induced in methanol grown *M. organophilum* XX but do not appear to be strongly repressed on addition of succinate to the medium (O'Connor and Hanson, 1977).

### 1.5.5 C1 Oxidation and Energy Metabolism

In typical aerobically grown heterotrophs the assimilatory power (mainly NAD(P)H and ATP) required for the biosynthesis of cell components is derived largely from the reducing equivalents generated by the tricarboxylic acid (TCA) cycle. In most methylotrophs, however, growth on C1 compounds is independent of such central catabolic pathways. Evidence from the PPFMs indicates that the TCA cycle is in some way

impaired during methylotrophic growth and can only play a very minor role in energy generation. The primary cause seems to be the severe, if not complete, repression of 2-oxoglutarate dehydrogenase as seen with *Pseudomonas* AM1 and *Pseudomonas* 1 (Large, 1981). Two TCA cycle enzymes (citrate synthase and 2-oxoglutarate dehydrogenase) are also repressed in methanol grown *Methylobacterium organophilum* XX (O'Connor and Hanson, 1977). During growth on C1 compounds, the reducing equivalents for assimilatory power are generated, largely, during the oxidation of the reduced C1 substrate via formaldehyde to $CO_2$, as shown in Figure 1. This pathway involves a series of NAD(P)H dependent and independent C1 specific enzymes, many of which have been characterised from representatives of the PPFMs.

1.5.5.1 Methane Mono-Oxygenase

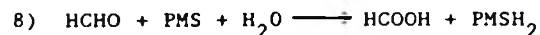
The ability to utilise methane facultatively has been reported in only one well characterised PPFM, namely *Methylobacterium organophilum* strain XX (Patt *et al.*, 1974, 1976). Methane oxidation is induced 4-5 hours after glucose grown cells are transferred to a minimal salts medium and incubated under an atmosphere of methane (Hanson, 1980). Methane mono-oxygenase (MMO) has been detected only in cells grown with methane. MMOs have been purified and characterised from several obligate methanotrophs (Dalton, 1981) but to date, little physical data is available on the MMO from strain XX because of difficulties in its purification (Haber *et al.*, 1983). However, it does appear to have a similar varied substrate range as is found with MMOs from other organisms (Hou *et al.*, 1979).

A point of interest is the wholesale refusal of strain XX to grow on methane outwith the United States of America (Hanson, personal communication). Many researchers have tried this repeatedly to no avail. One possible explanation is in the observation that the unstable methane utilising phenotype (and MMO activity) may be linked with the presence of an unstable plasmid (Haber *et al.*, 1983). This plasmid appears to be lost in unison with methane oxidising ability and may explain the "Trans-Atlantic curing" experienced by workers outwith the U.S.A.

### 1.5.5.2 Methanol (Alcohol) Dehydrogenase

Methanol (or alcohol) dehydrogenases have been found in all methane and methanol utilising bacteria. A broad specificity NAD(P)H-independant enzyme was first described in the PPFM *Pseudomonas* M27 (NCIB 9686) (Anthony and Zatman, 1964) and has subsequently been isolated from a number of methylotrophs. All appear to be closely similar with a pH optimum for activity of 9 to 11. *In vitro* the enzyme can catalyse the reactions

7) $RCH_2OH + PMS \longrightarrow RCHO + PMSH_2$



8) $HCHO + PMS + H_2O \longrightarrow HCOOH + PMSH_2$

where PMS = phenazine methosulphate. *In vitro* reactions are PMS-dependant. Reaction (8) can occur as formaldehyde exists in the hydrated form in aqueous solution.

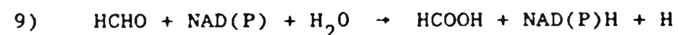
According to the scheme of Colby *et al.* (1979), the enzymes from representatives of the PPFMs fall into two categories. Group 1 enzymes, such as those found in *Pseudomonas* M27, *Protaminobcter ruber*, *Pseudomonas extorquens* and *Pseudomonas* AM1, catalyse the oxidation of only primary alcohols, require ammonia or methylamine for activity and are dimeric proteins of molecular weight about 120,000. Group 2 enzymes such as that from *Methylobacterium organophilum* XX, resemble those of group 1 but are capable also of oxidising some secondary alcohols. In addition the alcohol dehydrogenase from strain XX can slowly oxidise branched and cyclic alcohols (Wolf and Hanson, 1978). In *Pseudomonas* M27 the enzyme can account for as much as 10% of the total soluble protein (Quayle, 1972). Methanol dehydrogenases from both *M. organophilum* XX and *Pseudomonas* AM1 are inducible by methanol and are not repressed on the addition of succinate (O'Connor and Hanson, 1977; Dunstan *et al.*, 1972). Serological studies using antiserum raised against the methanol dehydrogenase from *M. organophilum* XX, have shown cross reactivity with extracts from other obligate and facultative methylotrophs (Wolf and Hanson, 1978).

Relatively little is known about the molecular mechanisms of methanol dehydrogenase action. *In vivo*, the enzyme appears to be coupled to the electron transport chain at the level of cytochrome c (Figure 7) and its activity, at least in some methylotrophs, has been shown to require the novel coenzyme pyrroloquinoline quinone (PQQ) (Duine and Frank, 1981). The methanol dehydrogenase from *Pseudomonas* AM1 was found largely in the cytoplasmic fraction of broken cells but when assayed the membrane fraction also contained some activity (Anthony, 1982). It is probable

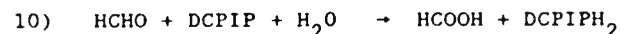
from its function that methanol dehydrogenase is capable of binding to the membrane where components of the electron transport chain are located. A distinct secondary alcohol dehydrogenase activity has been reported recently in *Pseudomonas* AM1 induced to some extent during growth on methanol but its function remains unclear (McNerney and O'Connor, 1980).

### 1.5.5.3 Formaldehyde Dehydrogenase

As stated above, NAD(P) independant methanol dehydrogenases can catalyse the conversion of formaldehyde to formate. In addition, methylotrophs can also contain either NAD(P) or dye linked formaldehyde dehydrogenases catalysing the *in vitro* conversion to formate as follows



9) $HCHO + NAD(P) + H_2O \rightarrow HCOOH + NAD(P)H + H$



10) $HCHO + DCPIP + H_2O \rightarrow HCOOH + DCPIPH_2$

Both enzymes have been shown in methanol grown *Protaminobacter ruber*, *Pseudomonas extorquens* and *Pseudomonas* strains 1 and 135 (Rock *et al*., 1976) but only the NAD(P) independant enzyme in *Pseudomonas* AM1 (Johnson and Quayle, 1964). The enzyme from *Pseudomonas* AM1 has been found to have a broad specificity towards aliphatic aldehydes.

Comparative studies on the specific activities of methanol and formaldehyde dehydrogenases from methanol grown *Pseudomonas* AM1, *Protaminobacter ruber* and *Pseudomonas extorquens* have been made (Zatman, 1981). Low activities for the dye linked enzyme seem to exclude it from having an important physiological role. When grown on methanol, the

organisms have a high level of methanol dehydrogenase activity and a low or zero level of NAD(P)H formaldehyde dehydrogenase activity. This would indicate that most formaldehyde oxidation is mediated by methanol dehydrogenase. Mutant studies with *Pseudomonas* AM1, however, have shown that a methanol dehydrogenase$^-$ strain has unimpaired oxidation of formaldehyde and no increase in the level of formaldehyde dehydrogenase. It has been suggested that the serine pathway could mediate the conversion of formaldehyde to acetyl-CoA with the TCA cycle catalysing the oxidation of acetyl-CoA to $CO_2$ (Zatman, 1981) but this seems unlikely due to the complete repression of 2-oxoglutarate dehydrogenase activity during methylotrophic growth.

1.5.5.4 Formate Dehydrogenase

The oxidation of formate to $CO_2$ is the final and probably the least studied step in the C1 dissimilatory pathway. Most methylotrophs appear to oxidise formate via a specific soluble NAD-dependent formate dehydrogenase as has been found in the PPFMs *Pseudomonas* AM1, *Protaminobacter ruber*, *Pseudomonas extorquens* (Johnson and Quayle, 1964) and *M. organophilum* XX (O'Connor and Hanson, 1977). The enzyme is induced to higher levels during methylotrophic growth of *Pseudomonas* AM1 (Johnson and Quayle, 1964). Attempts to purify the enzyme from this organism have been made but it was found to be unstable *in vitro* (Quayle, 1972).

1.5.5.5 Methylamine Dehydrogenase

Growth on methylamine and higher methylated amines involves their ultimate oxidation to ammonia and the key C1 intermediate formaldehyde.

The oxidation of methylamine is generally attributed to a methylamine dehydrogenase enzyme which catalyses the reaction

$$11) \quad CH_3NH_3^+ + H_2O + Dye \rightarrow HCHO + NH_4^+ + Dye\ H_2$$

The enzyme from *Pseudomonas* AM1 has been studied extensively (Eady and Large, 1971; Large, 1981), has a molecular weight of 105,000 made up of 4 non-identical subunits, and a pH optimum of 7.5. Similar enzymes have been shown to exist in *Pseudomonas* M27 and *Pseudomonas* strains 1 and 135. Methylamine dehydrogenase from *Pseudomonas* AM1 is known to have a quinone type prosthetic group similar to the PQQ molecule of methanol dehydrogenase (Large, 1981). Recent work has led to the isolation of a blue copper protein, amicyanin, which acts as the primary electron acceptor linking this enzyme to the electron transport pathway.

Some PPFMs are also capable of utilising dimethylamine (DMA) and even trimethylamine (TMA) as the sole carbon source. Little information is available on the mechanisms by which these substrates are utilised but it is likely that it comprises a series of oxidative N-demethylation steps as in (11) with formaldehyde being released at each stage.

1.5.5.6 Electron Transport and Energy Transduction Systems

In methylotrophs, the energy available from C1 oxidation reactions is harnessed as ATP by the subsequent oxidation of NAD(P)H and the reduced prosthetic groups of dehydrogenases by way of the proton translocating electron transport chains. These, in common with other aerobes, consist of flavoproteins, iron sulphur proteins, quinones (coenzyme Q),

cytochromes and cytochrome oxidases. The systems in the PPFMs are not yet fully understood but representative strains, namely *Pseudomonas* AM1, *Pseudomonas extorquens* and *Methylobacterium organophilum* XX have been shown to possess cytochromes b and c, usually cytochrome a/a3 (oxidase) and coenzyme $Q_{10}$ when growing on methanol (Anthony, 1975; O'Connor and Hanson, 1978; Urakami and Komagata, 1979; Anthony, 1981). The relative concentrations of each would appear to vary with methylotrophic and heterotrophic modes of growth. The most notable increase is that of soluble cytochrome c during methylotrophic growth, presumably indicating its particular importance during growth on C1 substrates. Conversely, the concentration of all membrane cytochromes (a, b and c) is markedly lower during methylotrophic growth and must be related to a greater capacity for NAD(P)H oxidation required during heterotrophic growth.

Electrons resulting from the oxidation of methanol appear to be donated to the electron transport chain largely by methanol dehydrogenase at the level of cytochrome c and after cytochrome b. *Pseudomonas* AM1 has been shown to have two soluble cytochromes c, $c_L$ and $c_H$, with differing isoelectric points. Cytochrome $c_L$ is the one immediately reduced by methanol dehydrogenase but the functional significance of this is not yet completely clear (O'Keeffe and Anthony, 1980; Beardmore-Gray and Anthony, 1984). The coupling of methanol oxidation to the electron transport chain enables proton translocated generation of a proton motive force (PMF) and ATP synthesis. To achieve this the methanol dehydrogenase and cytochrome c appear to react on the outer side of the periplasmic membrane with $O_2$ being consumed by an oxidase on the inner side of the membrane (O'Keeffe and Anthony, 1980; Beardmore-Gray and

Anthony, 1984).

During growth of *Pseudomonas* AM1 on methylamine, the methylamine dehydrogenase is also known to feed electrons into the electron transport chain (Large, 1981). Recently a type I blue copper protein, amicyanin, was found to be the primary electron acceptor for the enzyme with the ability to serve as an electron carrier between methylamine dehydrogenase and c-type cytochromes. Cytochrome $c_L$ is again thought to be immediately reduced by the enzyme.

Mutants of *M. organophilum* XX (O'Connor and Hanson, 1978) and *Pseudomonas* AM1 (Anthony, 1975) lacking cytochrome c and unable to grow on methanol can still oxidise most other substrates. Figure 7 shows a proposed electron transport chain for *Pseudomonas* AM1 growing on C1 compounds (Tobari, 1984). Formate dehydrogenase is the only NAD(P) linked enzyme. Growth thus tends to be NAD(P)H limited and less than 5% of electron transport to $O_2$ during growth on methanol is from NAD(P)H; 50% is from methanol dehydrogenase and the remainder from other dehydrogenases and the flavoproteins.

### 1.5.6 Regulation of C1 Metabolism

Much of the work on the physiological and genetic regulation of enzymes involved in C1 metabolism has been carried out with representatives of the PPFMs. Coordinate regulation of C1 specific enzymes in *Pseudomonas* AM1 was suggested by Dunstan *et al.* (1972) and supported later by work on the specific activities of the C1 enzymes under various growth

ure 7 The Electron Transport Chain proposed for Pseudomonas AM1 Growing on C1 Compounds

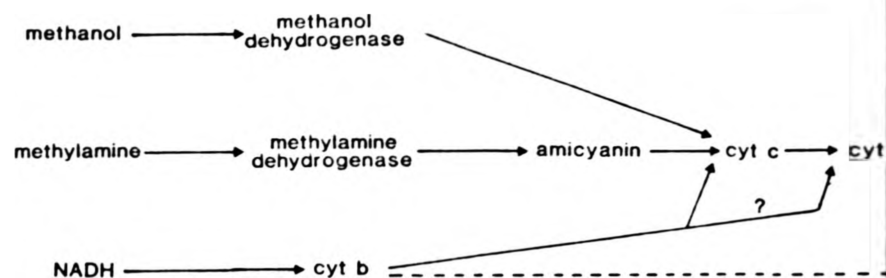
(from Tobari, 1984; Anthony, 1982)

---→ pathway for NADH oxidation under carbon excess conditions. It not yet clear whether cytochrome b feeds directly to cytochrome $c_H$ or cytochrome $c_L$ mediates the transfer under normal conditions.

ure 8 Model for the Arrangement of Genes Coding for C1 Specific Functions in Methylobacterium organophilum XX

(from O'Connor, 1981b)

promoter; o, operator; gly.kin., glycerate kinase; H. Pred, roxypyruvate reductase; SGAT, serine glyoxylate amino transferase; M, serine transhydroxymethylase; MCL, malyl-CoA lyase; MDH, hanol dehydrogenase; PEP carb., PEP carboxylase; cyt c, cytochrome R, 'regulatory molecule(s)'.



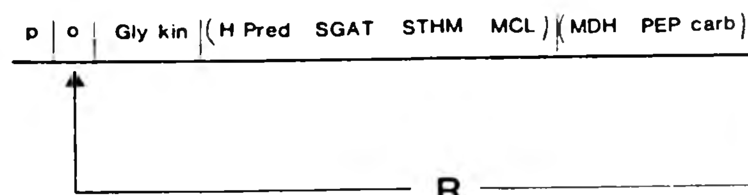



Figure 7 The Electron Transport Chain proposed for *Pseudomonas* AM1 Growing on C1 Compounds

(from Tobari, 1984; Anthony, 1982)

------→ pathway for NADH oxidation under carbon excess conditions. It is not yet clear whether cytochrome b feeds directly to cytochrome $c_H$ or if cytochrome $c_L$ mediates the transfer under normal conditions.

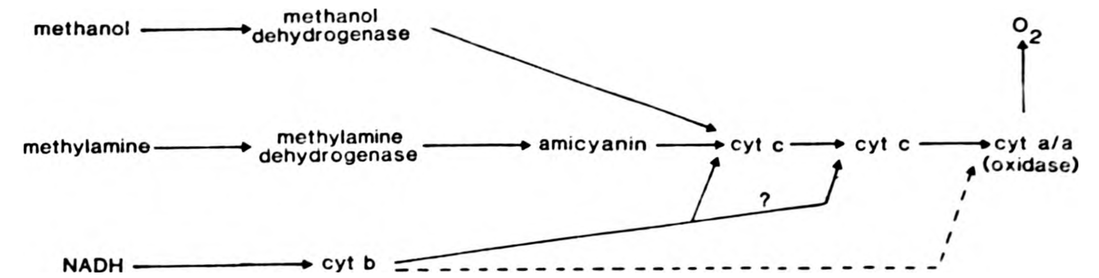
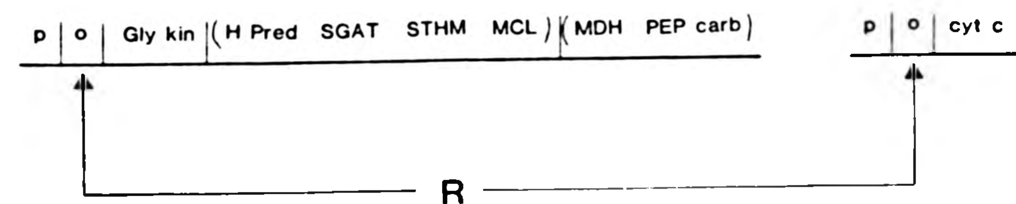



Figure 8 Model for the Arrangement of Genes Coding for C1 Specific Functions in *Methylobacterium organophilum* XX

(from O'Connor, 1981b)

p, promoter; o, operator; gly.kin., glycerate kinase; H. Pred, hydroxypyruvate reductase; SGAT, serine glyoxylate amino transferase; STHM, serine transhydroxymethylase; MCL, malyl-CoA lyase; MDH, methanol dehydrogenase; PEP carb., PEP carboxylase; cyt c, cytochrome c; R, 'regulatory molecule(s)'.



conditions (McNerney and O'Connor, 1980). Coordinate induction or repression of C1 pathway enzymes in the presence and absence of methanol has now been shown for methanol dehydrogenase and 6 serine pathway enzymes in *Pseudomonas* AM1 and *M. organophilum* XX (McNerney and O'Connor, 1980; O'Connor and Hanson, 1977). All of these C1 utilisation enzymes are coordinately induced in *M. organophilum* XX but methanol dehydrogenase, PEP carboxylase and serine transhydroxymethylase appear to be induced separately in *Pseudomonas* AM1. A further difference between these two organisms is that four of the key assimilatory enzymes (hydroxypyruvate reductase, serine glyoxylate aminotransferase, malyl-CoA lyase and glycerate kinase) are strongly repressed by growth on methanol plus succinate in *Pseudomonas* AM1, but not in *M. organophilum* XX. Succinate, or a product of its metabolism is therefore acting as a catabolic repressor of the serine pathway enzymes. Methanol dehydrogenase is not repressed by succinate in either organism. In strain XX the methane mono-oxygenase activity has also been shown to be induced only in the presence of methane (Hanson, 1980; Patt and Hanson, 1978). Coordinate regulation of some or all of the C1 specific enzymes suggests possible operonic control. To test this, a genetic transformation system has been developed in *M. organophilum* XX (O'Connor *et al.*, 1977) and using several pleiotropic mutants a linkage map for the C1 genes has been proposed. All genes coding for C1 enzymes are thought to be in an operon with the gene for cytochrome c being separate but still coordinately controlled by an unlinked regulatory molecule (O'Connor, 1981b) (Figure 8). The gene for formate dehydrogenase is also separate and appears to be regulated outwith this scheme. In strain XX methanol itself appears to be the operonic inducer and a

metabolite of it the inducer of formate dehydrogenase. Formate alone is unable to induce the C1 specific enzymes in strain XX (O'Connor, 1981) but does so in *Pseudomonas* AM1 (Large and Quayle, 1963). It is likely that there is also some regulation of C1 metabolism at the level of enzyme activity but, as yet, there is little data available on this aspect of control.

### 1.5.7 Pigments of the PPFMS

One of the most striking characteristics of the PPFMs is their pink or pink-red colouration, yet remarkably little is known of the nature of these pigments. In the pioneering work of Peel and Quayle (1961) they reported that methanol grown *Pseudomonas* AM1 contained red intracellular pigments assumed to be highly unsaturated carotenoids. A similar report was made by Anthony and Zatman (1964) during work with *Pseudomonas* M27. Little more was known until Downs and Harrison (1974) studied factors affecting the formation of pigment in *Pseudumonas extorquens*. They showed that the levels of pigmentation increased through the growth cycle and tentatively identified the main pigment as an oxo-carotenoid similar to rhodoxanthin. Pigmentless mutants of *Pseudomonas extorquens* had the same growth rate as the wild-type suggesting that the pigment plays no central metabolic role. Merabtine and coworkers (1982) carried out more detailed studies on the selective advantage conferred by the carotenoid pigments. They compared the growth and survival of the wild type with a pigmentless mutant, under varying conditions. The only evidence found suggested a photoprotective role for the pigment against ultra violet light. No evidence was found to support the suggestion

(Downs and Harrison, 1974) that carotenoids may protect the organism against the inhibitory effects of methanol by altering membrane fluidity. Similar studies have been carried out on *Protaminobacter ruber* (Shimizu *et al.*, 1982).

A curious phenomenon associated with the PPFMs is the presence of secondary pigments identified as bacteriochlorophyll. These pigments, normally found exclusively in photosynthetic bacteria (with the exception of a few marine bacteria (Shiba and Simidu, 1982)), have been identified in the PPFMs *Protaminobacter ruber* and *Pseudomonas* AM1 (Sato, 1978). The bacteriochlorophyll appears to be photoinducible by intermittent light (Sato and Shimizu, 1979) and the probable existence of a chromatophore-like structure has been indicated by electron microscopic studies. Bacteriochlorophyll formation has also been shown in the PPFM *Pseudomonas radiora* (Nishimura *et al.*, 1981). Recent work by Takamiya and Okamura (1984) has shed some light on the significance of bacteriochlorophyll in these organisms. With membrane preparations from *Protaminobacter ruber* they claim to have demonstrated light-dependent ATP formation via a cyclic photosynthetic electron transfer system. If correct, this has serious ramifications when considering the case of methylotrophy versus autotrophy, as distinct modes of existence and arouses speculation regarding the phylogeny of the organisms concerned.

### 1.5.8 Genetics of the PPFMs

A complete understanding of the methylotrophic processes in the PPFMs requires a knowledge of the genetical basis of C1 metabolism.

#### 1.5.8.1 Mutagenesis

A prerequisite for any genetical study is the isolation of suitable mutant strains. As facultative methylotrophs are capable of growth on multi-carbon compounds, non-lethal mutations can be induced in the C1 assimilatory enzymes. Standard techniques of UV irradiation and chemical mutagenesis have successfully generated a wide range of mutants in *Methylobacterium organophilum* XX (O'Connor *et al.*, 1977) and *Pseudomonas* AM1 (Anthony, 1982). Interestingly, mutants in *M. organophilum* XX with metabolic lesions were almost entirely confined to genes coding for enzymes that enable growth on methanol. Very few stable auxotrophs were isolated. Recently transposons (DNA sequences that code for resistance to an antibiotic and can insert with a degree of randomness into bacterial DNA - transposition) have been employed for more widespread mutant generation. Tranposons are usually introduced into the cell via a plasmid with subsequent transposition into cellular DNA causing insertional mutation at that site. Such a transposon delivery system using the suicide plasmid pUW964 (RK2 derivative) containing the transposons Tn5 (kanamycin resistance) and Tn7 (trimethoprim and streptomycin resistance) has been used successfully in *M. organophilum* XX (Haber *et al.*, 1983). Recently a similar system has been developed in *Pseudomonas* AM1 using the plasmid R91-5 with a Tn5 insertion (Whitta *et al.*, 1985).

1.5.8.2 Naturally Occurring Plasmids

Many methylotrophs have been shown to contain indigenous plasmids which are cryptic in that nothing is known of the functions they encode. The PPFMs *Pseudomonas extorquens* (NCIB 9399) and *Pseudomonas* AM1 (NCIB 9133) have been examined for plasmid DNA (Warner and Higgins, 1977). Only strain AM1 was found to contain 3 plasmids of molecular weights 15, 22 and 27 x $10^6$, all of which are entirely cryptic. A larger plasmid with a molecular weight of 8 x $10^7$ has been reported in *Methylobacterium organophilum* XX and its presence has been implicated with the organisms ability to utilise methane (Hanson, 1980; Haber *et al.*, 1983). Plasmid encoded methanotrophic functions would help to explain the instability of the methane-oxidising phenotype. Only circumstantial evidence, in that methane$^-$ cells contain no detectable plasmid DNA, supports this hypothesis.

1.5.8.3 Gene Transfer Systems

Until recently, genetic analysis of C1 functions in the PPFMs has been hampered by the lack of classical genetic systems, i.e. conjugation (transfer of DNA between cells by mating); transformation (transfer of naked DNA as duplex fragments) and transduction (transfer of DNA fragments via a bacteriophage) available in these organisms. Attempts to transform both chromosomal and plasmid DNA have failed in *Pseudomonas* AM1 (Fulton *et al.*, 1984) and *M. organophilum* can be transformed only by chromosomal DNA (O'Connor *et al.*, 1977). Several methanol mutants of *M. organophilum* have been mapped successfully by transformation studies and a model for the operonic control of C1 specific genes has been proposed (Figure 8).

The most successful method for the genetic investigation of PPFMs has involved either conjugative or mobilisable foreign plasmids and cloning vectors. Some broad host range plasmids are capable of mobilising the chromosomes of many gram negative bacteria. One such plasmid R68.45 (Holloway, 1981) has been shown to be transferred from a *Pseudomonas aeruginosa* PAO donor strain to the PPFM recipients *Pseudomonas extorquens*, *Pseudomonas* AM1 and *M. organophilum* XX (Jeyaseelan and Guest, 1979; Warner *et al.*, 1980). In all cases at least some of the antibiotic resistance markers are expressed in the recipient but no conjugative transfer of chromosomal markers could be demonstrated. More recently R68.45 was found to mobilise the chromosome of *Pseudomonas* AM1 and has been used to determine linkage relationships among a variety of mutations (Tatra and Goodwin, 1983, 1984). Linkage was observed between drug resistance markers and the genes for methanol dehydrogenase, cytochrome c, malyl-CoA lyase and possibly PEP carboxylase and glycerate kinase at frequencies of cotransfer greater than 95%.

1.5.8.4 Gene Cloning Systems

Further progress in the study of C1 utilisation genes has been made using cloning vectors and recombinant DNA techniques. Chromosomal DNA fragments produced by restriction endonuclease digestion, can be inserted into plasmid vectors cut by the same enzyme. Such recombinant plasmids can then be cloned in a suitable recipient to produce many copies of the vector and inserted fragment. A biological function can be assigned to the cloned DNA fragment by transfer into methylotrophic mutants and scoring for complementation of the mutant phenotypes. Initial attempts to clone genes from *M. organophilum* XX used the plasmid

R68.45. R68.45 can form R-prime (R') plasmids having incorporated fragments of chromosomal DNA by illegitimate recombination events. R' plasmids constructed in strain XX, were used to complement methanol mutants with limited success.

More recently, cosmid vectors (i.e. a cloning vehicle (plasmid) containing a ColE1 type replicon joined to the cos site of phage $\lambda$) have been applied to the cloning of *M. organophilum* XX DNA. This system combines the advantageous properties of both plasmid and phage transduction systems. Allen and Hanson (1985) constructed several new cosmid vector derivatives of the broad host range cloning vector pRK290. One of these, pLA2917 proved successful for constructing a gene library of strain XX DNA. This gene library has been used to complement mutants impaired in the ability to utilise methanol in *M. organophilum* strains XX and 761 and *Pseudomonas* AM1. The results indicate that five distinct complementation groups exist. The gene clones able to complement methanol dehydrogenase mutants were investigated further by mutagenising them with Tn5 in *E. coli* and scoring for the loss of ability to complement with PPFM mutants. The data obtained suggests that three genes, each separated by 1.25 to 5.5 Kb, are necessary for the expression of active methanol dehydrogenase.

In general, the results from complementation analyses support the model proposed for the arrangement of C1 genes in strain XX (figure 8). The cytochrome $c_{552}$ and regulatory genes appear not to be linked and the gene encoding malyl-CoA lyase is unlinked to any of the other C1 genes.

Gene libraries have also recently been constructed in *Pseudomonas* AM1 with the broad host range cosmid vector pVK100 (Fulton *et al.*, 1984). These have been used to detect and clone the malyl-CoA lyase gene from *Pseudomonas* AM1. The use of such recombinant DNA technology allows a possible study of the biological activities of cloned C1 gene products which have proved difficult to isolate and purify. Unfortunately, no methylotrophic DNA has yet been found to be recognised by *E. coli* transcriptional and translational systems (Haber *et al.*, 1983; Fulton *et al.*, 1984).

### 1.5.9 Classification and Taxonomy of the PPFMs

Interest in the biochemistry and potential industrial applications of methylotrophs has led, in recent years, to the isolation of many red or pink-pigmented facultative strains capable of growth on methanol (PPFMs). Although some attempts to identify and classify these strains have been made, their taxonomy has received insufficient attention. PPFMs have been isolated from a wide range of habitats and assigned variously to the genera *Bacillus*, *Flavobacterium*, *Chromobacterium*, *Vibrio*, *Protaminobacter*, *Mycoplana*, *Pseudomonas* and *Methylobacterium*. Many are also unnamed.

The first PPFM to be described in detail was that of Bassalik in 1913, who isolated an organism from the contents of an earthworm and named it '*Bacillus extorquens*'. This organism has led a clandestine existance being subsequently renamed '*Vibrio extorquens*' in 1948, '*Pseudomonas extorquens*' in 1949, '*Flavobacterium extorquens*' in 1960 and was given as '*Vibrio extorquens*' in the 7th Edition of Bergeys Manual of

Determinative Bacteriology (1957). In the 8th Edition of the manual (1974) it was re-listed as '*Pseudomonas extorquens*'. The strain used in this study is not the original strain, but is rather the strain re-isolated by Janota-Bassalik and Pedyck (1961). Such confusion of nomenclature typifies many of the PPFMs so far isolated.

The oldest existing PPFM strain is probably '*Protaminobacter rubrum*' isolated by den Dooren de Jong (1927) and subsequently emended to '*Protaminobacter ruber*' (Hayward, 1960).

In the 1950s, De Vries and Derx isolated some pink-pigmented organisms from plant root nodules and leaf surfaces and found them to be very similar to de Jongs original isolates of *P. ruber*. De Vries and Derx grouped their isolates, along with *P. ruber*, in the genus *Mycoplana* as *Mycoplana rubra*. None of these strains appear to have survived to the present day. Associations with plant surfaces are common amongst methanol-utilising bacteria (Corpe and Basile, 1982).

Five red-pigmented pseudomonads isolated by Hayward (1960) were compared with *P. ruber* and the aerial contaminant isolate *Pseudomonas* AM1 of Peel and Quayle (1961). All were found to be morphologically similar. Haywards isolates include the strains NCIB 9141, 9142 and 9145 used in this study. Peel and Quayle (1961) also noted the close similarity of their isolate *Pseudomonas* AM1 with the methylotrophic bacteria *P. ruber*, *V. extorquens* and *Pseudomonas methanica* (an obligate methanotroph later renamed *Methylomonas methanica*) and questioned their classification.

Stocks and McCleskey (1964) used growth substrate utilisation and serology studies to compare *V. extorquens*, *P. ruber*, *Pseudomonas* AM1 and *P. methanica* (a non-methanotrophic pink strain) with some of their own isolates. They concluded that they were all sufficiently similar to be designated strains of *V. extorquens* with the reservation that another genus may be more appropriate. In the 8th Edition of Bergeys Manual the genus *Vibrio* was restricted to fermentative organisms thus supporting the view that a change of genus would be appropriate. Anthony and Zatman (1964) isolated a further PPFM *Pseudomonas* M27 from soil and found that it too was closely related to the *extorquens* group.

Kouno and Ozaki (1975) isolated from soil and aquatic environments, some 223 methanol-utilising strains of which 59 were PPFMs. Data from the morphological and physiological characterisation of these PPFMs agreed closely with that from studies on existing PPFMs. Again it was noted that the taxonomic position of these bacteria was still obscure. Also in 1975, Chandra and Shethna isolated 41 strains of oxalate decomposing bacteria from chicken dung, 6 of which were PPFMs. These strains (including strains 385B and 602 used in this study) were again noted as being similar to the *V. extorquens* group.

Austin and Goodfellow (1979) isolated PPFMs from the leaf surfaces of perennial rye grass (*Loli um perenne*) and showed by numerical taxonomic methods, that some (their phenon 1) were distinct from other PPFMs such as *P. ruber*, *Pseudomonas* AM1 and *Pseudomonas rhodos* (Heumann, 1962). Consequently, a new species *Pseudomonas mesophilica* (NCIB 11561) was proposed for them.

Patt, Cole, Bland and Hanson (1974) reported the first PPFM able to utilise methane and it was subsequently included in a new genus *Methylobacterium* (Patt *et al.*, 1976). Only the type strain XX (NCIB 11278) of the sole species *M. organophilum* has been described in detail. Other organisms isolated subsequently have been assigned to the genus, often solely because of their facultative methane-utilising phenotype. As well as the type strain used in this study, the genus has included variously the isolate R6 (Patel *et al.*, 1978), now shown to be a mixed culture, and the isolate CRL-26 (Patel, 1984). It is unlikely that this latter strain is a bonafide PPFM. The acumen of classifying isolates with such as bias towards a single unstable character will be discussed later (Chapter 7). Two further species of facultative methanotrophs *Methylobacterium hypolimneticum* and *Methylobacterium ethanolicum* have been described (Lynch *et al.*, 1980). *M. hypolimneticum* has since lost its ability to utilise methane (P. Green, personal communication) and *M. ethanolicum* has been shown to be a mixed culture of an obligate methane utiliser and a *Xanthobacter* sp. (Lidstrom-O'Connor *et al.*, 1983).

Nishimura and coworkers (1983) noted the unusual phenomenon of bacteriochlorophyll formation in the PPFM *Pseudomonas radiora* (Ito and Iizuka, 1971) (NCIB 10815) and carried out comparative studies with representatives of the purple non-sulphur photosynthetic bacteria (*Rhodospirillaceae*) (Pfennig, 1977). They proposed that *P. radiora* was more closely linked to *Rhodopseudomonas palustris* and *Rhodopseudomonas sphaeroides* than to other gram-negative species. For this reason, representatives of the *Rhodospirillaceae* have been included in this study.

The most comprehensive study on the PPFMs, to date, has been the numerical taxonomic grouping of 150 strains by Green and Bousfield (1981, 1982). This study grouped the PPFMs into two closely related clusters distinct from other non-pigmented methanol utilising strains. This led to the proposal by Green and Bousfield (1983) for the emendation of the genus *Methylobacterium* and the transfer of all PPFMs to the redefined genus. Suggestions were also made to amalgamate some of the species but were not proposed because of uncertainty about the heterogeneity of the clusters. The spread of moles % (G + C) content in each cluster was 6-7% and DNA-DNA homology studies (Hanson, 1980) had shown only 45-48% homology between two strains in the same cluster.

A more recent numerical taxonomical investigation of methanol utilisers (Jenkins *et al.*, 1984) also included several PPFM strains. The results were in broad general agreement with those of Green and Bousfield (1982) but indicated a higher level of heterogeneity within the group.

About the same time, Urakami and Komagata (1979) examined the cellular fatty acid composition and the coenzyme Q systems in 34 strains of PPFMs and 114 strains of other methanol utilising bacteria. All PPFMs fell into one major group with 3 sub-divisions based on carbon utilisation spectra. In a second study, Urakami and Komagata (1981) used poly-acrylamide gel electrophoresis to examine and compare the electrophoretic mobilities of 15 enzymes from 32 strains of PPFMs and other methanol utilising bacteria. Numerical analysis of the zymograms divided the PPFMs into four groups. One of these (sub-group 4) corresponds to cluster 11 of Green and Bousfield (1982) and the other

three fall within their cluster 10. Subsequently, Urakami and Komagata (1984) proposed the formation of a new genus *Protomonas* to include all of the PPFM strains. This proposal contradicted that of Green and Bousfield (1983), but was valid in its own right as the type species *M. organophilum* was excluded from their study. Attempts to resolve this confusion have now been made (Bousfield and Green, 1985) with the rejection of the latter proposal. For the sake of clarity in this study, the names assigned to strains are those as detailed in the study of Green and Bousfield (1982).

### 1.5.10 Aims

As this brief history has shown, there is still sufficient confusion and uncertainty as to the true taxonomic relationships within the PPFMs to warrant a further in depth investigation. This present study was proposed to allow some of the more searching molecular biology techniques to be developed and applied to further elucidate the taxonomy of this group of organisms.

## Introduction to Methods and Bacterial Classification Systems

## 2. Bacterial Classification Systems

One of the questions most fundamental to any taxonomic study must be "why bother naming microorganisms?" According to the International Code of Nomenclature of Bacteria: "the primary purpose of giving a name to a taxon is to supply a means of referring to it" (a taxon being defined as a taxonomic group of any rank). This answer may seem self-evident but it is a point not always appreciated by microbiologists. Names are made to foster communication and to ensure that the description of a given set of characteristics, with a defined name, means the same to all scientists.

Further to naming microorganisms, there must be some logical system of ordering the names and grouping them by similarities. Thus, classification is defined as the ordering of microorganisms into groups or taxons. Taxonomy is often used synonomously with classification but is moreso the science of classification. To promote comparison between taxons once formed we must also study the evolutionary history of microorganisms. This is often referred to as phylogeny. Systematics is the more general study of diversity of microorganisms and their relationships (Sneath, 1978). As a scientific subject, the taxonomy and phylogeny of bacteria has, until recently, been in unacceptable disarray. Modern scientific methods are now being employed to help resolve this situation, much of which must be attributed to the historical development of classification systems with little or no defined rules to guide workers. From Pasteur's first observations of bacterial particles, early microbiologists adhered rigorously to a

classification scheme as used by fellow botanists and zoologists. This was biased heavily towards morphological characteristics and a few simple biochemical tests. Further problems have arisen in that plant, clinical and pure microbiologists and biochemists have all contributed their own biases and weighting of characters to produce very uneven and unnatural bacterial classification schemes.

## 2.1 Numerical Taxonomy

It was not until the 1960s, with the advent of computers and improved techniques that any great inroads were made in resolving the problems of bacterial systematics. Workers realised that a natural bacterial classification system must include examination of as many features as possible. This led to the development of Adansonian, or numerical taxonomy which involves the comparison of a large number of phenotypic characters of one organism with the same characters of another organism. Each character is given equal weighting, the degree of similarity is computed and on that basis the organisms are clustered into groups. The theoretical basis and application of numerical taxonomy are well documented (Sneath, 1972; Sneath, 1978; Sneath and Sokal, 1973) and a taxonomical investigation of this type, using 140 characters, has been carried out on the Pink Pigmented Facultative Methylotrophs (PPFMs) (Green and Bousfield, 1981, 1982).

Numerical taxonomic investigations have an advantage in that the large number of phenotypic and biochemical tests yield a great deal of basic information on the nature of any organism under study. However, a

problem when using only observable and recordable characteristics, without regard to ancestry, is that relationships established cannot be phylogenetic. Also, the choice of methods, interpretation of results and the inclusion of fastidious strains can pose problems for numerical taxonomists and influence the final result.

The average bacterium contains sufficient DNA to encode about 3000 genes. The most exhaustive numerical analysis can test only 10-20% of this genetic capacity. Consequently, techniques have been developed which examine the maximum amount of information available in the genome or, alternatively, the maximum amount of gene product information available. These are commonly referred to as molecular or chemotaxonomic techniques.

An ideal taxonomic system would be to sequence the complete deoxyribonucleic acid (DNA) from the genome of the organisms under study, then to compare these with standard sequences previously obtained and stored on computer. Closely related organisms would contain large parts of identical genomic sequence whilst more distantly related organisms would have smaller, and less frequent, similar (or homologous) DNA sequences. Although accelerating improvements in nucleic acid sequencing technology have made this dream closer to reality, indirect methods must still be used to estimate the DNA sequence (or gene product) similarities between different sources.

## 2.2 DNA Base Composition

The simplest indirect comparison of DNA sequences is in their base composition. DNA contains four bases: Adenine (A), Thymine (T), Cytosine (C) and Guanine (G). In normal double strand form the DNA strands are base paired by hydrogen bonds A - T and G - C. The ratio of (G - C):(A - T) can vary over a wide range between organisms and is usually quoted as moles % (G+C) content. For bacteria, this range is from 20 to 75%. Although DNA base composition can be determined chemically, after hydrolysis of the DNA and separation of the free bases, it is more readily determined by physical methods.

### 2.2.1 Thermal Denaturation Method

The thermal denaturation method makes use of the fact that the melting temperature of DNA (i.e. the temperature at which it becomes denatured by breakage of the hydrogen bonds) is directly related to the % (G+C) content. Bonding between G - C pairs is stronger than that between A - T pairs. Strand separation is monitored by the resulting increase in absorbance at 260 nm (hyperchromic shift) and is easily measured in a spectrophotometer (Mandel *et al.*, 1968; Marmur and Doty, 1962). The melting temperature (Tm) is quoted as that corresponding to the midpoint of the increase in absorbance. The % (G+C) content is determined from this value using the equation

$$\text{(12)} \quad \%\,(G+C) = 2.44\,(Tm - 69.4)$$

### 2.2.2 Density Gradient Method

The second method is that of caesium chloride (CsCl) buoyant density gradient analysis and is the method of choice in this study (Methods, 3.4). This involves centrifugation of a DNA sample in a CsCl gradient using a Model E analytical ultracentrifuge (Mandel *et al.*, 1968). The precise position of the DNA band in the gradient is determined optically and affords an accurate measure of its buoyant density. As the density ($\rho$) of the DNA in CsCl increases linearly with the % (G+C) content, this value can be calculated using the equation:

$$(13) \qquad \%\,(G+C) = \frac{\rho - 1.660}{0.098}$$

(Mandel *et al.*, 1968)

This method has the advantage that it is less susceptible to errors from contaminated DNA preparations and is less dependant on high molecular weight DNA for accurate analysis. Exhaustive comparisons with the thermal denaturation method reveal excellent correlation of results (De Ley, 1970c).

Further techniques have also been developed recently in an attempt to simplify and speed up the base composition determination. These include a spectrofluorometric method (Lutz and Yayanos, 1985) using A - T and G - C specific fluorescent dyes and a reversed phase high performance liquid chromatography (HPLC) method (Tamaoka and Komagata, 1984) where the base ratio of hydrolysed DNA is analysed directly.

Although information on the % G+C content of DNA is useful and is now

### 2.2.2 Density Gradient Method

The second method is that of caesium chloride (CsCl) buoyant density gradient analysis and is the method of choice in this study (Methods, 3.4). This involves centrifugation of a DNA sample in a CsCl gradient using a Model E analytical ultracentrifuge (Mandel *et al.*, 1968). The precise position of the DNA band in the gradient is determined optically and affords an accurate measure of its buoyant density. As the density ($\rho$) of the DNA in CsCl increases linearly with the % (G+C) content, this value can be calculated using the equation:

$$(13) \quad \% \, (G+C) = \frac{\rho - 1.660}{0.098}$$

(Mandel *et al.*, 1968)

This method has the advantage that it is less susceptible to errors from contaminated DNA preparations and is less dependant on high molecular weight DNA for accurate analysis. Exhaustive comparisons with the thermal denaturation method reveal excellent correlation of results (De Ley, 1970c).

Further techniques have also been developed recently in an attempt to simplify and speed up the base composition determination. These include a spectrofluorometric method (Lutz and Yayanos, 1985) using A = T and G = C specific fluorescent dyes and a reversed phase high performance liquid chromatography (HPLC) method (Tamaoka and Komagata, 1984) where the base ratio of hydrolysed DNA is analysed directly.

Although information on the % G+C content of DNA is useful and is now

considered essential in the description of any new species, it is in itself of limited taxonomic use. Organisms with disparate % (G+C) values are evidently not closely related, but similar, or identical, % (G+C) values are no indication that organisms are necessarily closely related.

### 2.2.3 DNA Restriction Analysis

A further method useful for the characterisation of DNA is restriction analysis. The base sequence in the DNA of any organism is constant and contains specific sequences which can be cut by restriction endonuclease enzymes chosen to give the required number of fragments. The pattern of fragments generated by subsequent gel electrophoresis is highly reproducible furnishing a 'finger-print' for each organism under test. On the basis that closely related strains show similar patterns this technique has recently emerged as an aid to the taxonomy of many genera (Crameri *et al.*, 1983; Razin *et al.*, 1983; Torres *et al.*, 1984). In addition, this technique can discriminate between serovars of strains indistinguishable by their DNA homology values (Manachini and Parini, 1983; Marshall *et al.*, 1984), and even between mutants of the same strain (Hintermann *et al.*, 1981). In all studies analysis of fragmentation patterns is only qualitative being restricted to visual observations and comparisons.

## 2.3 DNA-DNA Hybridisation

DNA-DNA hybridisation, as used for taxonomic studies of microorganisms, has been reviewed by De Ley (1970a); Johnson (1973); Bradley (1980);

Johnson (1981) and Schleifer and Stackebrandt (1983).

On heating, a DNA preparation becomes single stranded by thermal denaturation and, on cooling, at 25-30°C below the melting temperature (Tm), can specifically reassociate (homologous annealing) to reform native double stranded molecules. If DNA preparations from two different organisms are mixed and treated in this manner, reassociation between non-homologous DNA strands (= heterologous annealing or hybridisation) can occur. Under defined conditions, comparison with the level of homologous annealing gives a measure of the degree of difference between the heterologous strands. This principle is common to all nucleic acid hybridisation techniques and it is only the method of determining the degree of reassociation which varies significantly. Each method has its own utility and its own limitations. Most require the radioisotopic labelling of one of the DNAs by either *in vivo* or *in vitro* techniques.

The first technique used to detect DNA hybrids was that of Schildkraut *et al.* (1961). Hybrids formed between normal and heavy deuterium labelled DNAs banded separately when spun on CsCl gradients in an analytical ultracentrifuge. Determination of the amount of this mixed hybrid allowed an estimation of the sequence homology between the DNA strands. This method is expensive in terms of equipment, time consuming and is not amenable to all strains. Interpretation of results is also difficult but this laid the foundations for much of the subsequent work.

2.3.1 Matrix Methods

Later methods involved fixing one of the denatured DNAs to some sort of matrix (to prevent homologous reassociation) and allowing the second DNA to circulate around it in free solution to anneal. The first such method was that of Hoyer, McCarthy and Bolton (1964) where one high molecular weight single stranded DNA was fixed in an agar gel matrix. A small amount of low molecular weight single stranded radiolabelled DNA (= probe) was incubated with it overnight at an ionic strength and temperature suitable for hybridisation. Measurement of the amount of bound radiolabelled DNA allows an estimation of the % homology with the fixed DNA (assuming that the degree of homologous DNA binding = 100% homology) (De Ley, 1970b). This method has been superceded by that of Gillespie and Spiegelman (1965), whereby the agar matrix is replaced by nitrocellulose filter membranes. Although originally designed for DNA:RNA hybridisations the method was modified by Warnaar and Cohen (1966) and Denhardt (1966) for efficient DNA:DNA homology studies. In the DNA filter method, denatured single stranded DNA is immobilised on nitrocellulose filters after filtration and fixed by baking at 80°C under vacuum. Filters are often then coated with a mixture of albumen and polysaccharides (Denhardts solution), to prevent subsequent non-specific binding of probe DNA, and incubated with fragmented radiolabelled DNA under defined conditions. After careful washing, to remove unbound probe DNA, and drying, the amount of bound radioactivity is determined. Factors influencing the denaturation and binding of DNA to filters (Gillespie and Spiegelman, 1965; Fishman and Schiff, 1968; Phillips, 1969; De Ley and Tijtgat, 1970; De Ley and De Smedt, 1975; Baker, 1977) and the reassociation reaction (Denhardt, 1966;

Legault-Demare *et al.*, 1967; Johnson and Ordal, 1968; De Ley and Tijtgat, 1970; Flavel *et al.*, 1974a,b; Gillis and De Ley, 1975), have been studied thoroughly. Due to the high % (G+C) content of the PPFM DNAs, hybridisations were often performed using dimethylsulphoxide (DMSO) (Legault-Demare *et al.*, 1967) as the agent preventing non-specific binding of probe DNA to the filters as recommended by De Ley and Tijtgat (1970). A further problem with high % (G+C) DNAs is that the optimal renaturation temperatures are correspondingly higher. Studies have shown (De Ley and Tijtgat, 1970; De Ley and De Smedt, 1975) that at temperatures much above 65°C, there is an unacceptably high loss of filter bound DNA during incubation. This can be overcome by altering the salt concentration (Gillis *et al.*, 1970) or by adding organic solvents such as formamide (McConaughy *et al.*, 1969) or DMSO (Legault-Demare *et al.*, 1967) to the incubation mixture to depress the required incubation temperature. As a general rule, formamide in the hybridisation mixture reduces the effective Tm by 0.72°C for each 1% (v/v) added and DMSO by 0.6°C for each 1% (v/v) included.

Traditionally, hybridisation experiments have involved separate incubation reactions for each filter bound DNA and probe. Recently, however, there has been a trend towards using multiple spot (or dot blot) hybridisations whereby several DNA samples are bound to unique positions on a common filter matrix and hybridised with the probe DNA in a common solution (Kafatos *et al.*, 1979; Athwal *et al.*, 1984; Seldin and Dubnau, 1985; this thesis, 3.10.1). Stronger and more malleable nylon based membrane filters have also become available more recently for filter hybridisations (Cannon *et al.*, 1985).

### 2.3.2 Free Solution Methods

Alternative methods exist for DNA hybridisation which differ from the filter method in that single stranded DNAs are allowed to form duplexes in free solution. The first of these, developed by De Ley and coworkers measures DNA reassociation from renaturation rates, comparing rates of reassociation of homologous and heterologous DNAs by optical methods (De Ley *et al.*, 1970). This eliminates the need for radiolabelling of DNA but requires more DNA per reaction and a very reliable spectrophotometer with accurate thermocuvette. Analysis of the kinetics of such reassociation does also, however, allow estimation of the complexity and size of the genome (De Ley, 1970c; Gillis *et al.*, 1970; Gillis and De Ley, 1975).

The second method is the hydroxyapatite method. Fragmented radio-labelled DNA is mixed with a large excess of fragmented unlabelled DNA, which is then denatured and allowed to reassociate under suitable conditions. Duplexes are separated from unhybridised DNA by hydroxyapatite (a modified calcium phosphate gel) which specifically binds double stranded DNA (Brenner and Cowie, 1968; Jarvis *et al.*, 1980).

In the third method, reassociation is carried out as above but endonuclease S1 is used to specifically remove non-annealed single stranded DNA. Double stranded DNA is then precipitated with acid (Crosa *et al.*, 1973; Owen and Leaper, 1981).

Comparative studies of the free solution and matrix methods indicate a

good general agreement (De Ley *et al.*, 1970; Grimont *et al.*, 1980; Koops and Harms, 1985) of results.

### 2.3.3 Thermal Denaturation of DNA:DNA Hybrids

After DNA reassociation fragments may be exactly matched, partially matched or may well be matched in one, but unmatched in another region. This varies with the DNAs used as well as the precise hybridisation conditions chosen. The degree of mismatch can be determined directly by electron microscopic analysis of the duplexes (Davis *et al.*, 1971) but more satisfactory is to determine the thermal stability of the duplexes by heat denaturation (see methods 3.10.3). Depression of the melting point of heteroduplexes ($-\Delta Tm(e)$) compared with that of the homoduplex allows decisions about the degree of exact pairing. (The presence of 1% unpaired bases within a heteroduplex lowers its Tm by 1- 1.5$^{\circ}$C (Schleifer and Stackebrandt, 1983)). The stability of heteroduplexes formed under various conditions of hybridisation have been studied (Johnson and Ordal, 1968; Brenner *et al.*, 1969; Kingsbury *et al.*, 1969; De Ley *et al.*, 1973).

## 2.4 Polyacrylamide Gel Electrophoresis (PAGE) of Proteins

The amino acid sequence, molecular weight and charge on proteins are all determined by the nucleotide sequence of the corresponding cistron in the DNA. Although not all DNA codes for proteins, and not all genes are expressed at any given time, up to 2000 proteins can constitute, and function in microbial cells. Thus, an indirect study of comparative DNA coding capacities can be made by analysis of the protein gene products.

Electrophoresis of these mixtures through a gel matrix under standardised conditions, produces protein banding patterns or profiles which are equivalent to 'finger prints' for each bacterial strain. Soluble (Kersters and De Ley, 1975) or cell envelope (Mocca and Frasch, 1982) proteins can be compared, when electrophoresed and stained to visualise the banding pattern (= electrophoretogram). Alternatively, on electrophoresis a stain specific for only certain enzymes can be used (= zymogram). Some workers favour the use of only either phenol soluble or ribosomal proteins for electrophoretic analysis. This method has the advantage over DNA:DNA hybridisation in that it is relatively simpler and much more rapid for the comparison of large numbers of strains. The theoretical and practical aspects of PAGE are well documented (Hames and Rickwood, 1981). Recent reviews on the application of PAGE to bacterial taxonomy can also be found (Feltham and Sneath, 1979; Kersters and De Ley, 1980; Jackman, 1983).

Two forms of PAGE are currently used for examining bacterial proteins: (1) the disc electrophoresis system using tube gels (Kersters and De Ley, 1975) and (2) the vertical slab gel technique (Laemmli, 1970). The latter was favoured in this study as it allows a more direct comparison of multiple samples (see Methods, 3.6). Further variation is found in the staining method used to visualise the protein bands. Traditionally, amido black or coomassie blue have been the stains of choice because of their relative ease of use. This study favoured the silver stain method of Wray *et al*. (1981) because of its greater sensitivity and potentially greater resolution of the protein profiles. Gels were of the denaturing SDS-alkaline type and were run using a discontinuous buffer system (Methods, 3.6).

The degree of resemblance between protein band patterns is often estimated visually either by directly comparing stained gels or gel tracks (Moore *et al.*, 1980) or indirectly by visual comparison of gel photographs (Biavati *et al.*, 1982), densitometric tracings or schematic drawings (Hotchkin and Kaya, 1984). Although the human eye allows excellent determination of qualitative differences between gel tracks, comparisons are still subjective and not quantifiable. The introduction of computer assisted analysis of electrophoretograms has made PAGE protein analysis a much more valuable tool to the taxonomist. Although previously developed for eukaryotic systems, the first successful application of computer assisted PAGE analysis to bacterial classification was that of Kersters and De Ley (1975). Using cylinder gels, amido black stained cell free protein profiles were scanned in a densitometer to yield absorbance curves (scans) that can be expressed numerically. Internal protein standards allowed each electrophoretogram scan to be 'normalised' to a standard length (see 3.8.1). Normalised scans were divided manually into 90 equal parts, the extinction (min. height) of each position recorded and the values used to compute similarities betwen scans of different strains. Subsequent developments have allowed the entire electrophoretogram analysis procedures and their grouping by clustering techniques to be computer assisted (Jackman, 1982; Jackman, 1983; Jackman *et al.*, 1982). By storage of large numbers of electrophoretograms on computer, new and unknown isolates can be identified by comparison of their electrophoretic patterns with known patterns of reference organisms.

This study includes a computer assisted analysis of PPFM electrophoretograms using an original series of programs written in BASIC for use with the Acorn BBC B microcomputer.

It has been established that the primary limitation of the PAGE protein technique is reproducibility which must be established by a rigorously controlled standard procedure. Studies on the factors influencing reproduciblity have been made (Kersters and De Ley, 1975; Swings *et al.*, 1976; Jackman, 1983).

Recently, the use of high resolution two-dimensional (2-D) gel electrophoresis has been applied to bacterial classification (Swenson *et al.*, 1983; Jellum *et al.*, 1984). Conventional one-dimension PAGE bands often contain several unique proteins of identical mobilities. Two-D systems allow the resolution of these mixtures and enable greater discrimination between strains. As yet, the 2-D protein patterns can be compared only visually but computer assisted techniques are being developed (Jackman, personal communication).

PAGE-protein analysis is analagous with measurement of DNA:DNA homologies in that it is best applied to the measurement of comparatively close taxonomic relationships. The results show good correlation where both sets of data are available (Biavati *et al.*, 1982; Owen and Jackman, 1982; Watabe *et al.*, 1983; Kersters *et al.*, 1984). These techniques are, in general, restricted to intra-generic comparisons but, on occasions, close inter-generic relationships can be detected. The former is particularly true when considering more ancient genera.

## 2.5 Other Taxonomic Methods

Several other methods have been applied successfully to the identification, comparison or classification of bacteria at the generic level. These include: analysis of lipids, lipopolysaccharide or other cell wall fatty acid profiles by gas-liquid chromatography (Bousfield *et al.*, 1983; Weckesser *et al.*, 1983); serological analysis using cross reaction of antibodies raised against either whole cell extracts (Ridell and Williams, 1983), specific enzymes or, more recently monoclonal antibodies raised against whole cells (Macario and Macario, 1983); mass spectrometric analysis of whole cells.

## 2.6 Phylogenetic Techniques

All of the aforementioned techniques are of little use in establishing relationships much beyond the boundary of any particular genus. The ultimate goal of systematics is the classification of organisms according to their genuine evolutionary relationships, i.e. their phylogeny. Improved genetical and biochemical techniques now enable the investigation of distantly, as well as closely, related bacteria and the size of the phylogenetic tree is already impressive (Stackebrandt and Woese, 1984). Phylogenetic studies on prokaryotes must be based on the direct or indirect comparison of conserved genetic sequences as substitute fossil records.

### 2.6.1 Protein Studies

An indirect comparison of nucleotide sequences can be made by examining

the transcribed and translated protein products of conserved genes. This generally involves either protein sequencing or immunological techniques. The electron transport chain protein cytochrome c is a well conserved molecule and has proved useful in recognising phylogenetic relationships amongst eukaryotes, the results correlating well with available paleontological data. The sequencing of cytochrome c has also now been applied successfully to bacterial systems, especially to the investigation of the purple non-sulphur genera (*Rhodospirillaceae*) (Ambler and Meyer, 1979; Dickerson, 1980).

### 2.6.2 Ribosomal RNA Studies

Among the various phylogenetic approaches, emphasis is placed on the use of ribosomal ribonucleic acid (rRNA) molecules. Such molecules have been shown to be well conserved during evolution, largely because of severe functional constraints on their structure. Their functioning relies on the secondary structure of the molecule, making any mutational base change deleterious and effectively selected against. rRNA genes are also universally distributed and are not subject to appreciable lateral transfer (Stackebrandt and Woese, 1981). Two approaches, namely DNA-rRNA hybridisation and rRNA oligonucleotide sequencing are employed for phylogenetic studies.

#### 2.6.2a DNA-rRNA Hybridisation

The principles and practical requirements of DNA:rRNA hybridisation studies are very similar to those for DNA:DNA hybridisation. Of the three major rRNA molecules, the 23S RNA (- 2700 nucleotides) is used preferentially but also of use can be the 16S RNA molecule (- 1540

nucleotides). Comparative studies have shown that the results are virtually identical when these two species are used, due to their equivalent degree of conservatism (M. Gillis, personal communication). The 5S rRNA molecule (- 120 nucleotides) is considered too small for meaningful hybridisation. Any mutational base alterations are correspondingly more significant in this smaller molecule and their effects can be accentuated by its saltatory behaviour.

The technique of DNA:rRNA hybridisation, as applied for taxonomic purposes, has been developed largely by De Ley and co-workers. In brief it involves fixation of high molecular weight DNA on filters and incubation with labelled 23S rRNA (or 16S rRNA) in 2 x SSC, 20% formamide at 50°C for 16 hrs. Filters are washed, treated with RNAase to remove unbound or unpaired segments of rRNA from the hybrids (Gillespie and Spiegelman, 1965; McConaughy *et al.*, 1969; De Ley and De Smedt, 1975) then the degree of rRNA binding determined. In contrast with DNA:DNA homology studies, the percentage of rRNA binding to the DNA is not a reliable measure of RNA homology. This value depends not only on the actual sequence similarity but also upon the genome size, the state of replication of the genome and also the number of rRNA cistrons per genome. Of more use is determination of the thermal stability ($Tm_{(e)}$) of the DNA:rRNA duplexes by step-wise thermal denaturation (De Ley and De Smedt, 1975). Recently an alternative parameter to $Tm_{(e)}$ for comparison has been proposed, namely % RIT, the percentage of labelled RNA released at the starting temperature of thermal denaturation (Van Landschoot *et al.*, 1984). The main problem associated with RNA hybridisation studies is the isolation of sufficient quantities of pure

rRNA of the required specific activity. Studies of factors influencing the hybridisation of rRNA with DNA have been made (Bishop *et al.*, 1969; Bishop, 1970; Gillespie and Gillespie, 1971; De Ley and De Smedt, 1975).

Recently, attempts to identify and clone specific rRNA genes and to carry out specific DNA:rDNA hybridisation studies to determine relatedness have been made (Schleifer *et al.*, 1985). This technique is of potentially greater use than the direct rRNA method.

2.6.2b <u>rRNA Oligonucleotide Sequencing</u>

The sequencing of rRNA molecules is potentially more powerful than the DNA:rRNA hybridisation approach, because lower levels of relationship can be detected (mismatchings in DNA:rRNA hybrids causing the $Tm_{(e)}$ to fall below about 65°C form the lower limit for significant taxonomic conclusions). rRNA sequencing allows a direct comparison of more distantly related genera.

The molecule of choice for sequencing studies is 16S rRNA, being easier to sequence than 23S rRNA and not subject to the saltatory problems of the 5S rRNA molecule. When studies began the 16S rRNA molecule was too large to sequence directly but the then current nucleic acid sequencing technology made it possible to sequence large enough fragments for comparative analyses. In summary, the method involves digesting $^{32}P$ radiolabelled 16S rRNA from a given species with T1 nuclease and resolving the oligonucleotides (15-20 residues) by 2-dimensional paper electrophoresis. The individual oligonucleotides are then sequenced

using a combination of endonuclease procedures to produce, finally, a catalogue of sequences, characteristic for each organism. Oligonucleotide catalogues from different organisms are then compared (which is tantamount to comparing the primary rRNA sequence) pairwise to produce a similarity coefficient - SAB . Standard clustering procedures on SAB values allow the merging of groups to produce a phylogenetic tree or dendrogram. The dendrogram so produced is a reasonable approximation to true phylogenetic relationships provided that the mutational clock in all organisms is isochronic (Stackebrandt and Woese, 1981). Over 400 bacterial species have been characterised by this method to produce by far the most comprehensive insight into the evolutionary relatedness of bacteria (Balch *et al.*, 1979; Woese *et al.*, 1980; Fox *et al.*, 1980; Stackebrandt and Woese, 1981, 1984). Results show good agreement with cytochrome c data and reasonable agreement with DNA-rRNA hybridisation studies (Fox *et al.*, 1980; Stackebrandt and Woese, 1984).

Modern sequencing techniques now make it much more feasible to determine directly the nucleotide sequence of RNA molecules. Already, 5S rRNA molecules have been used for complete sequencing studies (Hori and Osawa, 1979; Dekio *et al.*, 1984) and most results appear to be phylogenetically valid showing good correlation with 16S rRNA studies (Fox *et al.*, 1980).

# Materials & Methods

## 3.1 Strains and Culture Conditions

### 3.1.1 Strains

The strains used in this study are listed in Tables 3 and 4.

### 3.1.2 Culture and Maintenance of Strains

Batch cultures of PPFMs were routinely grown in the mineral salts (MS) medium of Colby and Zatman (1973) with 0.2% (v/v) methanol (GPR) as the carbon source. Cells were grown at 30°C on a rotary shaker, in conical flasks fitted with cotton wool plugs.

Alternatively, the PPFMs were grown using glycerol-peptone (GP) medium (1% (w/v) Difco peptone, 1% (v/v) glycerol, pH 7.0) or nutrient broth (NB) (Oxoid CM67,pH 7.0).

Members of the *Rhodospirillaceae* were grown anaerobically in pyruvate-malate-yeast extract (PMY) medium containing 1.0 g sodium hydrogen malate, 1.0 g sodium pyruvate, 0.5 g $NH_4Cl$, 0.4 g $MgSO_4.7H_2O$, 0.05 g $CaCl_2.2H_2O$, 0.4 g NaCl, 0.2 g yeast extract per litre of double distilled water, pH 6.8. 50 ml of 0.1 M phosphate buffer (pH 6.8) were added to each litre of medium after sterilisation.

Cells were grown at 30°C in flasks capped with rubber 'suba-seals' (William Freeman & Co., England) allowing them to be flushed with $O_2$ free $N_2$ after inoculation. Illumination was provided by tungsten lamps.

Members of the *Enterobacteriaceae* were grown aerobically in flasks

Table 3 Designation and Source of PPFM Strains

| Strain Name | Strain No. | Reference |
|---|---|---|
| *Methylobacterium organophilum* | NCIB 11278 (Strain XX) XY[a] | Patt *et al*. (1976) |
| *Protaminobacter ruber* | NCIB 2879 | den Dooren de Jong (1927) |
| *Pseudomonas rhodos* | NCIB 9421 | Heumann (1962) |
| *Pseudomonas radiora* | NCIB 10815 | |
| *Pseudomonas extorquens* | NCIB 9399 | Janata, Bassalik & Pedcyk (1961) |
| *Pseudomonas mesophilica* | NCIB 11561 (strain A47) | Austin & Goodfellow (1979) |
| *Corynebcterium rubrum* | (CR) | Graf & Bauer (1973) |
| *Vibrio extorquens* | 385B, 602 | Chandra & Shethna (1975) |
| *Mycoplana rubra* | R14 | |
| *Pseudomonas* spp | NCIB 10598, 10601, 10604, 10609, 10610, 10611 | |
| *Pseudomonas* spp. | 1, 135 | Rock *et al*. (1976) |
| *Pseudomonas* sp. | NCIB 9133 (Strain AM1) | Peel & Quayle (1961) |
| *Pseudomonas* sp | NCIB 9686 (Strain M27) | Anthony & Zatman (1964) |
| *Pseudomonas* sp. | NCIB 9141, 9142, 9145 | Hayward (1960) |
| Unnamed strains | CS51 | |
| | B46 | |
| | 45, 317, 434 539, 617, 790 M159-1, N-2, N-6, N-12, 0-31 0-46 | Kuono & Ozaki (1975) |
| | 2CR | |
| | ICI Strain 269 | |
| | D5, D12, D21 | |

[a] supplied by R. Hanson, Gray Freshwater Biology Institute, Minnesota, U.S.A.

All other strains supplied by P. Green, National Collection of Industrial and Marine Bacteria, Aberdeen, U.K.

Table 4 Designation and Source of non-PPFM Strains

| Strain Name | Strain No. | Reference |
|---|---|---|
| *Rhodomicrobium vannielii* | RM5 (Warwick) | Whittenbury & Dow (1977) |
| *Rhodopseudomonas acidophila* | ATCC 25092B | |
| *Rhodopseudomonas blastica* | NCIB 11567 | Eckersley & Dow (1980) |
| *Rhodopseudomonas capsulata* | NCIB 8254 | |
| *Rhodopseudomonas gelatinosa* | NCIB 8210 | |
| *Rhodopseudomonas palustris* | NCIB 8288 | |
| *Rhodopseudomonas sphaeroides* | NCIB 8253 | |
| *Rhodopseudomonas sphaeroides* | subsp. cordata | Gest *et al.* (1983) |
| *Rhodopseudomonas viridis* | ATCC 19567 | |
| *Rhodospirillum rubrum* | NCIB 8255 | |
| *Rhodospirillum tenue* | ATCC 25093 | |
| *Escherichia coli* | C | |
| *Escherichia coli* | K12 | |
| *Escherichia coli* | DH1 | |
| *Escherichia coli* | JC6310 | |
| *Salmonella typhimurium* | LT2 | |
| *Serratia marcescens* | | |
| *Methylococcus capsulatus* | Bath | Whittenbury et al. (1970) |

containing either nutrient broth or Luria-Bertani (LB) medium (10 g Bacto tryptone, 10 g NaCl, 5 g yeast extract per litre of double distilled $H_2O$, pH 7.0) at 37°C on a rotary shaker.

Stock cultures of PPFM strains (except strain CS51) were grown at 30°C and maintained at 4°C on glycerol peptone (GP) agar((GP) broth + 1.5% (w/v) Bacto agar, pH 7.0). Strains were subcultured once every 4 weeks. Strain CS51 stock cultures were grown and maintained on nutrient agar (NA) medium (Oxoid, pH 7.0).

Stock cultures of enterobacterial strains were maintained on nutrient agar at 4°C.

Stock cultures of members of the *Rhodospirillaceae* were grown and maintained at room temperature in stabs of PMY agar (PMY broth + 1.5% (w/v) Bacto agar, pH 6.8). Strains were subcultured once every 2-3 months.

Cells were harvested from batch culture in an MSE Hi-spin 21 centrifuge (10,000 g) 4°C for 15 mins.

## 3.2 Cell Lysis and DNA Isolation

### 3.2.1 Cell Lysis

Cells pelleted from liquid culture were washed once in one-twentieth volume of TES buffer containing

100 mM Tris

10 mM EDTA

100 mM NaCl

taken to pH 8 with conc. HCl

Washed cell pellets were resuspended in 25 ml of TES, pH 8.0 per 2 g wet weight of cells.

A variety of strategies were adopted in an attempt to obtain gentle lysis of PPFM cells. These included

(a) Incubation with lysozyme (1 mg $ml^{-1}$, Sigma Grade 1) at 37°C.

(b) Incubation with lysozyme at 37°C followed by addition of sodium lauryl sulphate (SLS) (2% (w/v) final concn.) at 37°C or 60°C.

(c) Passage of cells through the French Press at $1.4 \times 10^7$ Pa followed by lysozyme and SLS treatment.

(d) Incubation with lysozyme and penicillin (Glaxo, 100 $\mu g\ ml^{-1}$) then SLS treatment.

(e) As in (d) above with French Press pretreatment.

(f) High Pressure French Press Treatment $1.4 \times 10^8$ Pa.

(g) Incubation with lysozyme followed by the addition of alternative detergents such as amino salicylic acid at 1% (w/v) final conc. (Sigma), sodium dodecyl sulphate at 1% (w/v) final concn. (BDH, Electrophoresis grade), sarkosyl at 2% (w/v) final concn. (Koch-Light Laboratories), Brij 58 at 2% (w/v) final concn. (Sigma), deoxycholate at 1% (w/v) final conc. (Sigma) and Triton X100 at 1% (v/v) final concn.

(h) Incubation with snail digestive juice, 100 $\mu g\ ml^{-1}$ (Sigma Type H-2)

(i) Incubation with lysozyme and other catabolic enzymes such as amylase, 500 $\mu g\ ml^{-1}$; chitinase, 1 $mg\ ml^{-1}$; phospholipase, 20 $\mu g\ ml^{-1}$; cellulase, 1.8 $mg\ ml^{-1}$; protease, 1.7 $mg\ ml^{-1}$; followed by treatment with SLS.

(j) Heat (60°C) and cold (4°C) shock of cells followed by lysozyme and SLS treatment.

(k) Repeated freezing and thawing of cells before lysozyme and SLS treatment.

(l) Osmotic shock before lysozyme and SLS treatment.

(m) Growth of cells in low concn. of antibiotics such as D-cycloserine, 100 $\mu g\ ml^{-1}$ (Sigma); ampicillin, 100 $\mu g\ ml^{-1}$ (Sigma); cloxacillin, 100 $\mu g\ ml^{-1}$ (Beechams); penicillin G, 100 $\mu g\ ml^{-1}$ (Glaxo, Na salt); bacitracin, 100 $\mu g\ ml^{-1}$ (Sigma, zinc salt); vancomycin, 100 $\mu g\ ml^{-1}$ (Sigma, hydrochloride); erythromycin, 100 $\mu g\ ml^{-1}$ (Sigma); novobiocin, 100 $\mu g\ ml^{-1}$ (BDH, Na salt); kanamycin, 100 $\mu g\ ml^{-1}$ (Sigma, sulphate); polymixin B, 100 $\mu g\ ml^{-1}$ (Sigma, sulphate) before treatment with lysozyme and SLS.

Batches of cells were also grown up then incubated, after harvesting, with lysozyme and the above antibiotics at 37°C.

(n) Incubation with lysozyme and glycine, or lysozyme and glycerol, followed by SLS treatment.

(o) Incubation with lysozyme after growing the cells on alternative carbon and nitrogen sources.

(p) Treatment with lysozyme and SLS after growing the cells in low concentrations of SLS (0.1%, w/v), sarkosyl (0.1 %, w/v) or DMSO (0.1% v/v).

(q) Pretreatment of cells with strong alkali (1 M NaOH or KOH) followed by lysozyme and SLS treatment.

(r) Incubation with lysozyme and proteinase K, 100 $\mu g\ ml^{-1}$ (Sigma), or protease, 100 $\mu g\ ml^{-1}$ (Sigma) before SLS treatment.

(s) Prewashing the cells with organic solvent such as phenol, acetone, butanol and chloroform before lysozyme and SLS treatment.

The most reproducible and effective method of cell lysis was as follows: An equal volume of either acetone or butanol was added to the cell suspension and the two phases mixed and left at room temperature for 5 mins. The cells were then pelleted in an MSE Chilspin centrifuge by spinning at 5,000 g for 10 mins at $4^{o}C$. Cell pellets were then washed twice with TES (pH 8.0) before being resuspended in the original volume of TES buffer. Lysozyme (20 mg $ml^{-1}$ in double distilled $H_2O$) was added to a final concentation of 1 mg $ml^{-1}$ and the suspension incubated at $37^{o}C$ for 1-2 hrs, or until lysis was apparent by an increase in viscosity of the suspension. Lysis was completed by adding SLS to a final concentration of 2% (w/v) and incubating at $60^{o}C$ for 10 mins.

3.2.2 DNA Isolation

DNA was isolated and purified by the methods of Marmur (1961) and Kirby *et al.* (1966) with the following modifications:

3.2.2.1 Method 1

The cooled cell lysate was made 1 M with sodium perchlorate (5 M) and mixed vigorously to give a foamy suspension. This aids the dissociation of proteins from the nucleic acids. The mixture was shaken gently with an equal volume of chloroform: isoamyl alcohol (24:1 v/v) in a stoppered flask for 20 mins. Phases were separated by spinning at 5,000 g for 15 mins at $4^oC$ and the upper aqueous phase was carefully removed with a large bore pasteur pipette. This extraction step was repeated a further two times. The aqueous phase was made 0.25 M in NaCl by the addition of 5 M NaCl and was then gently overlaid with 2 volumes of ice-cold ethanol (95%). Gentle mixing of these layers with a glass stirring rod allowed the nucleic acids condensed at the interface to be spooled as a threadlike precipitate. Excess fluid was removed by pressing the spooled nucleic acid against the vessel and allowing it to air dry for 5 mins. The precipitate was removed from the rod by gently swirling in 10-20 ml of 0.1 x SSC (0.015 M NaCl, 0.0015 M Na citrate, pH 7.0) and was resuspended rotating gently overnight at $4^oC$. The solution was made up to 1 x SSC by the addition of one-tenth volume of 10 x SSC and ribonuclease (2 mg $ml^{-1}$ ribonuclease A, Sigma) in 0.1 x SSC (preheated at $90^oC$ for 15 mins to destroy any DNAase activity) was added to a final concentration of 100 $\mu g\ ml^{-1}$. After incubation at $37^oC$ for 30 mins, the preparation was extracted, as above, with a chloroform:isoamyl alcohol mixture. This deproteinisation step was repeated until no denatured protein was visible at the interface. The DNA in the aqueous phase was reprecipitated with NaCl and 2 vols ethanol and resuspended in a minimal volume of 0.1 x SSC (2-3 ml) overnight at $4^oC$.

The following were of importance:

(a) If the cell lysate remained cloudy after extraction with chloroform:isoamyl alcohol, a mixture of phenol:chloroform:isoamyl alcohol (25:24:1, crystallised phenol (BDH) saturated with, TES, pH 8.0) was substituted. This is a more powerful deproteinising mixture but if used it must be followed by two chloroform:isoamyl alcohol extractions to remove all traces of phenol. Some DNAs are, however, very sensitive to fracture by phenol.

(b) Routinely, the DNAs were subject to further purification by centrifugation through a caesium chloride (CsCl) gradient. 19 g of CsCl (Fisons) in a 25 ml polycarbonate centrifuge tube were dissolved in 12 ml of 0.1 x SSC. The DNA solution was gently layered on top and the whole contents made up to 34 g with 0.1 x SSC (final density of CsCl solution = 1.71 g $cc^{-1}$). Tubes were completely filled with liquid paraffin and spun at 80,000 g (MSE Superspeed 65) for 42 hrs at 4°C. Gradients were harvested from the bottom. Viscous fractions were pooled and dialysed extensively against 0.1 x SSC at 4°C.

(c) Some DNAs, especially those contaminated with polysaccharide material, were further purified by substituting the final ethanol precipitation with an isopropanol precipitation step. One-tenth volume of acetate-EDTA (3 M Na acetate, 0.001 M EDTA, pH 7.0) was added to the aqueous phase and 0.54 volumes of isopropanol were carefully layered on top. The precipitated DNA was spooled then washed by dipping in increasing concentrations of ethanol at 70, 80, 90 and 95%. The isolated DNA was again resuspended in a

minimal volume of 0.1 x SSC. This precipitation step was not used routinely with PPFM DNA as it often gave poor yields with high loss of precipitable material.

3.2.2.2 Method 2

12 ml of a cell lysate were added to 19 g of CsCl in a 25 ml polycarbonate tube (MSE 34411-147). This was made up to 34 g with TES buffer (pH 8.0) and rotated slowly to dissolve the CsCl. Tubes were completely filled with liquid paraffin and spun at 80,000 g (MSE Superspeed 65) for 42 hrs at 4°C. Gradients were harvested from the bottom and viscous fractions pooled and dialysed extensively against 0.1 x SSC at 4°C. Suspensions were then precipated with ethanol as described in Method 1.

Resuspended DNAs were treated with RNAse then extracted with chloroform:isoamyl alcohol (24:1) until no proteinaceous material was found at the interface. DNAs were resuspended in minimal volumes of 0.1 x SSC following a futher ethanol precipitation step.

3.2.3 Storage of DNA

Aliquots of each DNA were kept at 4°C in small plastic vials over a drop of chloroform. The remainder of each DNA was carefully labelled and stored at -70°C.

3.2.4 Large Scale Plasmid Preparation

The washed pellet from a litre batch culture of cells was resuspended in

7.5 ml of ice-cold 10% (w/v) sucrose in 50 mM Tris HCl (pH 8), left on ice for 10 mins then lysozyme (10 mg $ml^{-1}$) added to a concentration of 1.5 mg $ml^{-1}$. After a further 10 mins on ice, 3 ml of 0.25 M EDTA (pH 8) were added, mixed gently, left on ice for 10 mins then Brij 58 (Sigma) and Na deoxycholate were added to final concentrations of 1% (w/v) and 0.4% (w/v) respectively and mixed gently until lysis was complete. The lysate was centrifuged at 65,000 g in an MSE Superspeed for 30 mins at 4°C to remove cellular debris and much of the chromosomal DNA.

Plasmid DNA in the supernatant was precipitated by the addition of one-third volume of 4 M NaCl, two-thirds volume of 25% (w/v) PEG 6000 at 4°C for 6 hrs then pelleted at 5,000 g for 5 mins. Pellets were gently resuspended in 1 ml of TE buffer (pH 8) at 4°C. To each 1.1 ml of resuspension was added 4 ml of a stock solution of 80 g CsCl + 8 ml ethidium bromide (5 mg $ml^{-1}$ in TE) + 52 ml TE (pH 8). Mixtures were placed in centrifuge tubes (MSE 34411-119) and CsCl/EtBr solution syringed below it to the level of the bottom of the thread before being balanced with liquid paraffin. On banding at 75,000 g in an MSE Superspeed at 18°C for 60 hrs the lower plasmid band was visualised under long wave length UV light and removed with a syringe and needle. After extensive dialysis in TE buffer the ethidium bromide was removed by repeated extractions against water saturated butanol.

### 3.3 Purity of DNAs

Due to the difficulty in purifying many of the PPFM DNAs, 3 separate methods were used to determine the concentration and purity of each DNA isolated.

### 3.3.1 Spectrophotometric determination

DNA aliquots were diluted 20 fold in 0.1 x SSC, placed in 1.0 ml volume quartz cuvettes and the absorbance measurements at 230, 260, 280 and 320 nm read on an LKB Ultrospec spectrophotometer. The concentration of the DNA in the sample was calculated using the formula

$$(14) \quad [DNA]\ \mu g\ ml^{-1} = OD_{260} \times 50 \times \text{dilution}$$

based on the assumption that DNA at a concentration of 50 $\mu g\ ml^{-1}$, with a 1 cm light path has an absorbance value of 1.0. Ratios of $A\frac{230}{260}$, $A\frac{280}{260}$ and $\%A\frac{320}{260}$ were used as an estimate of DNA purity (Mordarski *et al*., 1976).

### 3.3.2 Burton Diphenylamine Assay

A modification of the method of Burton (1968) was used to chemically determine the amount of DNA present.

A stock solution of herring sperm DNA, 2.0 mg $ml^{-1}$ (Sigma) was diluted with 0.1 x SSC into chromic acid washed, silicated glass vials to give a range of 1, 2.5, 5, 10, 25 and 50 $\mu g$ of DNA in 0.5 ml volumes. Duplicates of each of these concentrations were used to construct the calibration curve for the assay. 0.5 ml lots of 0.1 x SSC, with no DNA added, were included as blanks. Duplicate 50 $\mu l$ aliquots of each unknown DNA were added to 0.45 ml lots of 0.1 x SSC. 0.25 ml of 10% (v/v) perchloric acid was added to each vial and vortexed. 2 ml of diphenylamine reagent (BDH, 2 g in 100 ml of glacial acetic acid + 2 ml of concentrated $H_2SO_4$ and acetaldehyde at 0.02% v/v) were added and each

vial revortexed. After overnight incubation at 30°C, the degree of blue colouration produced by the reaction was determined in an LKB Ultrospec at 595 nm.

By constructing a calibration curve of standard herring sperm DNA concentrations against $OD_{595}$ values, the concentrations of the unknowns could be calculated after correcting for the dilution factor.

The concentrations of all unknown DNAs were determined either in duplicate or triplicate.

### 3.3.3 Agarose Gel Electrophoresis

To determine the approximate molecular weight, and relative purity of the DNAs isolated, equivalent samples were run on agarose gels. 0.5 or 1 μg lots of DNAs were made up to 20 μl with 0.1 x SSC and mixed with 5 μl of 5 x gel loading buffer (GLB) (see below) then run on 0.7% (w/v) agarose (Sigma, Type I) gels using either TBE or TAE buffer (see below).

Ethidium bromide (1 μg $ml^{-1}$) was included in the gel and running buffer enabling the DNA bands to be observed directly, using an ultra-violet transilluminator, after electrophoresis. Gels were photographed under UV light using a Polaroid camera.

| TAE Gel Buffer 20x | TBE Gel Buffer 10x |
|---|---|
| 96.9 g Trizma base | 108 g Trizma base |
| 32.8 g Na acetate | 55 g Boric acid |
| 15.2 g EDTA | 9.5 g EDTA |
| pH to 8.3 with glacial acetic acid and made up to 1 l. with double distilled water | made to 1 l. with double distilled water, pH 8.3 |

| TAE Gel Loading Buffer 5x | TBE Gel Loading Buffer 5x |
|---|---|
| 50% (v/v) Glycerol | 50% (v/v) Glycerol |
| 0.5% (w/v) Bromophenol blue | 0.5% (w/v) Bromophenol blue |
| in 5 x TAE buffer | in 5 x TBE buffer |

### 3.4 Determination of % G + C Content of DNA

The moles per cent guanine plus cytosine (% G + C) contents of the PPFM DNAs were determined by the buoyant density method of Mandel and Marmur (1968) in a Beckman Model E Analytical Ultracentrifuge. *Micrococcus lysodeikticus* (Sigma, ATCC 4968) and *Escherichia coli* K12 DNAs (densities of 1.731 and 1.710 g $cm^{-3}$ respectively) were used as internal standards.

#### 3.4.1 Sample Preparation

960 μg of CsCl (BDH, analytical grade) were wetted with TES, pH 7.0. 1 μg of each of the unknown and standard DNAs were added and made up to 1.710 g with TES buffer. The sample was mixed by gentle rolling and placed in the Model E sample chamber with a syringe. Samples were run

for a minimum of 18 hrs at 20°C before being photographed using the ultra violet optics.

### 3.4.2 Calculation of % G + C Content

Photographic negatives from the Model E were analysed on a Joyce 'Chromoscan' and the moles % G + C content determined as described by Mandel and Marmur (1968). Duplicate runs were made for the majority of PPFM DNAs.

## 3.5 Preparation of Soluble Protein Cell-Free Extracts

### 3.5.1 Cell Culture and Lysis

PPFM strains were grown to an $OD_{600}$ of 0.2-0.4 (early-mid log phase) in MS medium + 0.2% (v/v) methanol at 30°C. One litre cultures were pelleted, washed twice in TES buffer and the cells resuspended in 3-5 ml of TES (pH 8.0). Samples of all cultures were steaked to single colonies on GP agar plates before harvesting, to check purity.

One ml aliquots of each cell suspension were disrupted by ultra-sonication in an MSE 12/76 Mark 2 sonicator at a peak to peak maximum of 18 microns. Samples were cooled in an ice-methanol bath and sonication was restricted to 20 sec periods with 60 sec gaps to prevent over-heating of the sonicates.

Lysis was monitored with a light microscope. 8-10 20 sec bursts were required, on average to give over 80% lysis of the PPFM suspensions.

After sonication, the suspensions were spun at 12,000 g in a microfuge (Eppendorf 5412) for 25 mins to pellet the majority of the cellular debris. Cell-free extracts were dialysed extensively at $4^{o}C$ against gel sample buffer (see section 3.6) and stored at $-20^{o}C$.

Soluble protein extracts were similarly prepared from members of the *Rhodospirillaceae* grown to an equivalent OD on PMY medium and from selected PPFMs grown on a variety of alternative carbon sources.

3.5.2 Determination of Protein Concentration in Cell-Free Extracts

Protein concentations of the cell-free lysates were determined using the Bio-Rad protein assay system and bovine serum albumin (BSA) as a standard. Briefly, the method was as follows:

1. Several dilutions of BSA containing 0.2 to 1.4 mg $ml^{-1}$ were prepared for the standard curve.
2. 0.1 ml of standards and appropriately diluted samples were placed in clean test tubes. 0.1 ml of sample buffer was placed in the 'blank' test tube.
3. 5 ml of diluted dye reagent (1 in 5, Bio-Rad concentrate) was added to each tube and gently vortexed.
4. The $OD_{595}$ was measured after 10 mins against the reagent blank. $OD_{595}$ plotted against concentration of standards gave a standard curve from which the unknown values could be determined.

## 3.6 Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS-PAGE) was performed using a discontinuous buffer system including the ionic detergent sodium dodecyl sulphate (SDS) to dissociate proteins into their individual polypeptide subunits.

### 3.6.1 Solutions

Gels were cast between glass plates (small gels 20 x 25 cm, large gels 45 x 25 cm) with a stacking gel polymerised on top of the resolving gel (Laemmli, 1970). Glass plates were separated by lucite spacers with the edges sealed by polypropylene tubing smeared with vaseline, and held together by bulldog clips.

The following gel solutions were used:

| 30% (w/v) Acrylamide Mixture | 20ml |
|---|---|
| 60% (w/v) low bisacrylamide | 10 ml |
| 75% (v/v) glycerol | 7.3 ml |
| Lower gel buffer | 2.5 ml |
| 10% (w/v) SDS | 0.2 ml |

Degassed for 5 mins in a vacuum dessicator before 4 $\mu$l of TEMED (NNNN'-Tetra-methylethylenediamine) and 40 $\mu$l of freshly prepared ammonium persulphate (10% w/v) added.

| 10% (w/v) Acrylamide Mixture | 50 ml | 60 ml | 150 ml |
|---|---|---|---|
| 60% (w/v) high bisacrylamide | 8.3 ml | 9.9 ml | 24.9 ml |
| $H_2O$ | 34.9 ml | 41.9 ml | 104.7 ml |
| Lower gel bufer | 6.25 ml | 7.5 ml | 18.75 ml |
| 10% (w/v) SDS | 0.5 ml | 0.6 ml | 1.5 ml |

Degassed for 5 mins before 10, 12 or 30 μl of TEMED and 100, 120 or 300 μl of freshly prepared ammonium persulphate added to the 50, 60 or 150 ml mixtures respectively.

| Stacking Gel Mixture | 10 ml |
|---|---|
| Stacking gel acrylamide | 3.0 ml |
| $H_2O$ | 4.4 ml |
| Stacking gel buffer | 2.4 ml |
| 10 % (w/v) SDS | 0.1 ml |

Degassed for 5 mins before 5 μl of TEMED and 100 μl of freshly prepared ammonium persulphate added.

60% (w/v) Acrylamide Stocks

Low % gel (high bis)

60 g Acryalmide (Kodak, Eastman)

1.6 g Bisacrylamide (Kodak, Eastman)

High % gel (low bis)

60 g Acrylamide

0.3 g Bisacrylamide

Both made up to 100 ml with double distilled water.

Stacking gel acrylamide

| | | |
|---|---|---|
| 10 g | Acrylamide | made up to 100 ml with |
| 0.5 g | Bisacrylamide | double distilled water |

Buffers

Lower gel buffer

| | |
|---|---|
| 36.33 g Trizma base (3M) | made up to 100 ml with |
| Conc. HCl to give pH 8.8 | double distilled water (pH 8.8) |

Stacking gel buffer

| | |
|---|---|
| 5.98 g Trizma base (0.05 M) | made up to 100 ml with |
| conc. HCl to give pH 6.8 | double distilled water (pH 6.8) |

Running Buffer (5 x stock)

| | | |
|---|---|---|
| 30.2 g | Trizma base | in 1 litre of double |
| 144 g | Glycine | distilled water |

Running buffer = 200 ml of stock + 10 ml (w/v) SDS per litre

Sample Buffer

| | |
|---|---|
| $H_2O$ | 525 ml |
| Stacking gel buffer | 125 ml |
| Glycerol | 100 ml |

3.6.2 Gel Formation

Two gel types were used to resolve the soluble proteins, a 10-30% (w/v) concave gradient gel and more often a 10% (w/v) linear gel.

The pouring of a concave 10-30% (w/v) gradient gel of 50 ml total volume required 20 ml of 30% (w/v) acrylamide mixture placed in a mixing chamber. 50 ml of 10% (w/v) acrylamide was pumped into the mixing chamber using a peristaltic pump and thoroughly mixed with the 30% (w/v) acrylamide solution. The volume of liquid in the mixing chamber remained constant and the gel was poured with the 30% (w/v) acrylamide constantly being diluted by the 10% (w/v) solution. 10% (w/v) linear gels were prepared by pumping a single 10 % (w/v) acrylamide solution of 60 ml (small gels) or 150 ml (large gels) volume through the mixing chamber and between the gel plates.

After pouring the gel mixture was overlaid with water saturated butanol and allowed to polymerise. When set, the butanol was washed off with double distilled water and a stacking gel poured above the resolving gel with a lucite slot former pushed into the top of the stacking gel. After setting the polypropylene tubing and slot former were removed from the gel before being placed in an electrophoresis tank (as described by Studier, 1973) for loading.

Care was taken that on addition of the running buffer to the tank, no air bubbles were trapped between the plates below the gel. Any such air bubbles were removed with a 10 ml syringe and hooked needle filled with running buffer.

3.6.3 Sample Preparation

The amount of soluble protein samples loaded varied depending on the dimensions of the gel and the gel staining method used. For 20 x 25 cm gels stained by Coomassie blue, 50-100 $\mu$g of soluble protein were loaded per gel slot and for gels stained by the silver method (Wray *et al.*, 1979) 10 $\mu$g onto 20 x 25 cm gels and 5 $\mu$g onto larger 45 x 25 cm gels, with narrower slots. If internal standards were to be used, these were added (0.5 $\mu$g of each of thyroglobin and cytochrome C) to the soluble protein samples before denaturation.

Denaturation of the samples (5-50 $\mu$l) was by boiling with 10 $\mu$l of SDS (10% w/v), 10 $\mu$l of 70% (v/v) glycerol, 10$\mu$l of $\beta$-mercaptoethanol and 5 $\mu$l of 0.1% (w/v) bromophenol blue dye in small glass vials for 5 mins. Individual gel slots were loaded with denatured samples using a Hamilton 0.1 ml glass syringe, extensively washed between samples with double distilled water.

3.6.4 Electrophoresis

Electrophoresis was carried out overnight at $4^{o}$C in a cold room at 12 mA (constant current) for 10-30% (w/v) gradient gels and 10 mA (constant current) for 10% linear gels, until the marker dye reached the bottom of the gel. Gels were then removed from between the plates, fixed and stained as required.

3.6.5 Gel Staining

a) Coomassie Blue stain

Gels were immersed in the staining solution (45% v/v methanol, 10% v/v glacial acetic acid and 0.1 % w/v Coomassie blue R250) for 4-5 hrs, then destained in 45% v/v methanol, 10% (v/v) glacial acetic acid to produce an acceptable background colouration.

b) Silver stain

The method of silver staining used was that of Wray *et al*. (1981). This simple and rapid method can detect levels of protein 10-20 times less than those detected by Coomassie blue stain. Briefly the method was as follows:

i) Gel soaked in 50% (v/v) methanol for at least 6 hrs with 2 changes of solution.

ii) Stain solution:

Solution A - 0.8 g of silver nitrate dissolved in 4 ml of distilled water

Solution B - 21 ml of 0.36% (w/v) sodium hydroxide solution mixed with 1.4 ml of 14.8 M ammonium hydroxide solution. Solution A was added drop-wise into solution B with constant mixing and finally made up to 100 ml with distilled water.

iii) The gel was soaked in the stain solution for 20 mins.

iv) The gel was then washed in deionised water for two 5 mins periods.

v) Developing Solution: 2.5 ml of 1% (w/v) citric acid and 0.15 ml of 37% (w/v) formaldehyde solution made up to 500 ml with deionised water. The gel was soaked in developer solution until bands appeared - usually 10-20 mins.

vi) Development was stopped by washing the gel briefly with deionised water then placing in 50% (v/v) methanol, 5% (v/v) acetic acid.

Gels were photographed using Kodak Panatomic X 35 mm film (32 ASA) on a fluorescent light transilluminator.

3.7 Densitometry of Stained Gels

Stained gels were inserted between 2 glass plates and scanned with a Joyce Loebl Mk. III CS microdensitometer interfaced to an Acorn BBC B microcomputer system. Large (45 x 25 cm) gels were cut into two equal parts with a scalpel before being scanned. The optimal settings of the densitometer were as follows: objective, x10; final aperture width, 60 $\mu m$; aperture height, 10; 133 wedge; forward scan speed 2.5; pen damping, 7.5.

A chain driven potentiometer translates the movement of the pen recorder into electrical impulses which were fed via an interfacing series of variable resistors to the microcomputer. This analogue output was then digitised and converted by an analogue to digital converter (ADC, 10 bit resolution) in the computer back to a profile to be displayed on a computer monitor (BBC microcomputer systems) and stored on floppy disc

(Memorex 80 track disc), as 1000 point traces. 31 files could be stored on each disc side. Protein profiles were obtained as direct pen recordings from the densitometer. The computer program designated 'Joyce-L' allowed for the storage and recall of all scanned profiles. Further copies of the traces either singly or superimposed could be obtained on a Radio Shak TRS-80 graphics plotter linked to the computer. To allow more rapid analysis the computer files of each gel trace were formatted from a 32-bit to 16 bit array format and stored separately on disc using the Convert program.

### 3.8 Analysis of Densitometer Traces

Analysis of the protein patterns from early 10-30% (w/v) gradient and 10% (v/v) gels relied simply on scanning each gel track and making direct visual or computer assisted comparisons. The mobility of protein preparations differed slightly across a single gel even under the most reproducible conditions. However, the inclusion of the internal reference proteins thyroglobulin (33K, Pharmacia) and cytochrome C (12K, Sigma) allowed 'normalisation' (Kersters and De Ley, 1975) and subsequent comparison of different scans across the same gel or between different gels.

#### 3.8.1 Normalisation

The major analysis programs written included a compact or expand ('compand') step on the stored 1000 point electrophoretic traces. Computer analysis of the scans determined the positions of the reference thyroglobulin and cytochrome C peaks at either end of the trace and

arbitrarily assigned these to positions 0 and 899 on a new 900 point trace. The intervening points from the original trace were then 'companded' by a scaling factor to take up positions in the same relative order in the new 'normalised' 900 point trace. This allowed the direct comparison of all scanned traces irrespective of the length of the original gel track. Two strategies were adopted in an attempt to analyse the information in the normalised gel traces. All programs listed are original, written in Basic for the BBC B microcomputer by M. Whiteside and D. Hood (University of Warwick).

Direct visual observations and grouping of apparently similar protein profiles supplemented the computer assisted analyses.

### 3.8.2 Peak Search and Position Analysis - COMP 4, AUTOC C(A)

Program COMP (compand) 4 allowed the comparison of any pair of files selected from either side of 2 floppy discs (using Acorn BBC double-disc drive) by a peak search and relative position analysis. Each selected trace was scanned at 256 horizontal levels from the maximum to the minimum trace height. Peaks were detected by analysis of the proximity of trace points at successive levels under the influence of selectable search parameters (input in program menu) allowing for a minimum peak height and width threshold rejection for each trace. This gave some constancy to the number of peaks detected for each electrophoretic trace and enabled the rejection of any spurious background peaks found at finer discrimination levels. Each gel trace was plotted as a reproduction of the original companded gel track but only with the peaks (= bands) detected under the search parameters. A third track was also

plotted marking the peaks coincident in both tracks. Even after normalisation of traces, small variations in the positions of homologous bands existed due to technical imperfections of the system used. Correspondingly, COMP 4 included a 'margin factor' (normally $\pm$ 0.3% trace length) allowing misaligned peaks to be marked as coincident if they fall within narow defined limits on either side of the absolute position for coincidence.

There was often a considerable background absorbance trend in gel traces due to overlapping peaks (Feltham and Sneath, 1979; Jackman, 1983) which was more pronounced when the Coomassie blue stain was used. COMP 4 had the capacity to remove this trend by constructing a grossly averaged trace of just ten points and subtracting 0.5 of this from the original trace. This allowed increased discrimination, especially between traces where different amounts of protein had been loaded and differing numbers of peaks detected.

AUTOC-CA

AUTOC-CA (autocomparison) was a program derived from COMP4 which allowed the automatic peak search analysis of all pairs of traces obtained from any designated gel. The printout from AUTOC-CA simply listed the numbers of peaks selected in each trace and the number of coincident peaks as well as two coefficient values calculated for each comparison.

$$\text{Coefficient 1 (CF1) (Dice coefficient)} = \frac{2 \times \text{number of coincident peaks selected}}{\text{number of peaks selected in Trace A + number of peaks selected in Trace B}}$$

$$\text{Coefficient 2 (CF2) (Jaccard coefficient)} = \frac{2 \times \text{number of coincident peaks selected}}{2 \times \text{number of peaks selected in Trace A or B (whichever was lowest)}}$$

Either coefficient could be selected to be stored on disc as an unsorted similarity matrix for subsequent analysis.

### 3.8.3 Correlation Coefficient Analysis - AUTO EF

An alternative mode of analysing the electrophoretic traces was to compare complete trace profiles rather than to rely simply on peak positions for reference. The programme AUTO EF again allowed the automatic comparison of all combinations of pairs of traces from a single gel. Each trace was again normalised by companding and a comparison of corresponding points made using the correlation coefficient r.

$$(17) \quad rab = \frac{\sum^n (Xa_1 - Xa)(Xb_1 - Xb)}{([\sum^n (Xa_1 - Xa)^2][\sum^n (Xb_1 - Xb)^2])}$$

$Xa_1$, $Xb_1$ - height values for traces a and b at any corresponding sample point along it.

Xa, Xb - average height values for traces a and b.

Most analyses used average traces each of 100 points (9 points averaged) but finer discrimination analysis using traces of 300 points (3 points averaged) or 450 points (2 points averaged) could also be accommodated. A margin factor was also included in the program to compensate for any

slight misalignment of traces (Feltham and Sneath, 1979; Jackman, 1983). Best fit was obtained by recalculating the similarity of each corresponding set of points at increasing points of misalignment on the adjacent portions of the traces to a maximum of 1% of total trace length. Results from AUTO EF analyses were displayed as lists of individual comparisons and their corresponding r values. R values were stored in memory to be saved on disc as an unsorted matrix file at the end of each completed computer run.

On occasions the protein standard markers if included in the comparisons did increase the apparent similarities between two unrelated profiles. Auto EFP was a program derived from Auto EF which allowed for comparison ignoring the internal markers by excluding from the calculation a small percentage of the trace internal from the selected marker positions.

Program Auto EFP had no integral background subtraction capacity but, if required, this could be removed prior to analysis using a program commercially available for the LKB Ultrascan Densitometer (Warwick Control Systems, Warwick). This software was compatible with the files generated by Joyce-L and allowed background subtractions ranging from straight-line through partial background trend to complete background trend removal.

3.8.4 Matrix Sorting - DENDO

DENDO was a program written to sort the similarity matrices generated by AUTO CA and AUTO EF(P) using UPGMA clustering. The instructions for drawing a dendogram were output to a printer enabling the dendograms to

be drawn by hand.

## 3.9 Preparation of Labelled DNA

### 3.9.1 In vivo labelling

Cells were grown in 100 ml of MS medium, 0.2% (v/v) methanol supplemented with 1 $\mu$Ci $ml^{-1}$ of 2,5',8-$^{3}$H adenosine (Amersham). DNA was extracted as previously described in section 3.2 and stored at 4$^{o}$C.

### 3.9.2 In vitro labelling - Nick Translation

DNA probes of a higher specific activity than could be obtained by *in vivo* labelling were prepared by a variation of the method of Rigby *et al.* (1977). All solutions were prepared using sterile double distilled $H_2O$ and transfers were carried out on ice.

Reaction mixture

6 $\mu$l of 5 x nick translation buffer (see below)

1 $\mu$l of 1 mM deoxyadenosine triphosphate (dATP)

1 $\mu$l of 1 mM deoxythymidine triphosphate (dTTP)

1 $\mu$l of 1 mM deoxyguanosine triphosphate (dGTP)

1 $\mu$l (10 $\mu$Ci) of [$\alpha^{32}$P] deoxycytidine triphosphate (dCTP, Amersham) (10 mCi $ml^{-1}$, 3000 Ci $mmol^{-1}$)

1 $\mu$l of DNAase 1 (Sigma, 50 mg $ml^{-1}$)

2 $\mu$l of DNA polymerase 1 (Boehringer, 5U $\mu l^{-1}$)

Incubated at 15$^{o}$C for 3-4 hrs.

On occasions, the radiolabelled nucleotides dGTP or dATP were used in place of dCTP with the equivalent volume of cold dCTP added. 5 $\mu$Ci (5 $\mu$l) of [ $^{3}$H] dGTP were also used occasionally with the volume of distilled water added being altered accordingly. The reaction was stopped by the addition of 2.5 $\mu$l of 0.5 M EDTA, pH 7.5 and 2.5 $\mu$l of 20% (w/v) SDS, 100 $\mu$l of 10 mM Tris-HCl, pH 7.5 and 100 $\mu$l of phenol:chloroform:isoamyl alcohol (25:24:1) were added and vortexed for 1 min . Phases were separated by spinning in a microfuge for 5 mins. The aqueous phase was removed and a further 100 $\mu$l of Tris HCl was re-extracted with the phenol mixture. The second aqueous phase was pooled with the first after spinning and passed down a Sephadex G50 column equilibrated with gel filtration buffer (GFB). 0.4-0.5 ml fractions were collected. The initial labelled DNA peak fractions (4-6) and unincorporated label peak fractions (10-12) were detected using a hand held monitor (Mini Instruments, g.m. monitor).

Radiolabelled DNA fractions were precipitated with 30 $\mu$g ml$^{-1}$ carrier DNA or tRNA, 0.3 M sodium acetate and 2 volumes of 95% ethanol at -20$^{o}$C overnight. Precipitated fractions were spun for 25 mins in a microfuge, pellets were dried under a gentle stream of nitrogen and resuspended in the desired volume of 0.1 x SSC. Probes with a specific activity of >10$^{7}$ cpm $\mu$g$^{-1}$ of DNA were obtained using this method.

Solutions - Nick Translation Buffer 5x

0.25 M Tris-HCl pH 7.9

0.05 M $MgCl_2$

0.05 M Dithiothreitol

250 $\mu g\ ml^{-1}$ bovine serum albumin

Gel Filtration Buffer

10 mM Tris-HCl pH 7.5

1 mM EDTA

50 mM NaCl

3.10 DNA-DNA Hybridisation

DNA hybridisation studies were performed with the DNA extracted from 35 strains of PPFMs and 8 reference strains. The methods used were a modification of the membrane filter technique as first described by Gillespie and Spiegelman (1965) as modified by Denhardt (1966), Legault-Demare *et al*. (1967) and De Ley and Tijtgat (1970). DNA was denatured by heat or alkali treatment and filtered through a nitrocellulose filter, the matrix retained the single stranded DNA. This binding was made irreversible by heat treatment and enabled interaction with added radiolabelled single strand probe DNA in solution. After stringent washing to displace all unhybridised probe DNA the percentage homology was calculated from ratios of the amounts of radioactive probe remaining on the filters. Two methods of hybridisation were employed.

### 3.10.1 Multi-blot Filter Method

Schleicher and Schuell BA85 25 mm gridded nitrocellulose membrane filters (diameter 82 mm, 0.45 $\mu$m pore size) were used.

#### a) Preparation of Filter

5 $\mu$g aliquots of each DNA sample were pipetted into Eppendorf tubes (0.6 ml volume) and made up to 60 $\mu$l final volume with 0.1 x SSC. Samples were denatured by boiling for 10 mins then immediately cooled on ice. 0.4 volumes of ice-cold 20 x SSC were added to make the contents approximately 6 x SSC. Samples were then spotted under vacuum using a Gilson micropipette onto nitrocellulose filters prepared by soaking in distilled water for 10 mins then 6 x SSC for 10 mins. Filters were placed over pads wetted with 6 x SSC in Buchner-type funnels connected to a water vacuum pump under low pressure. Due to the hydrophobic nature of the grid lines, the multiple spots of each sample could be confined accurately to the area of a particular grid square. The position of each sample was unambiguously defined by numbering each row and column on the filter. Filters were rinsed with 20 ml of 6 x SSC after all samples had been sucked dry, then dried at 42$^{o}$C for 30 mins before being baked at 90$^{o}$C under vacuum for 2 hrs to bind the DNA. A maximum of 25 DNA samples were spotted onto each filter leaving sufficient area around each test square for easy dissection. All DNA samples were applied in triplicate (or at least duplicate) for each hybridisation experiment. At least 2 unrelated control DNAs were included to detect non-specific binding of the probe DNA.

b) Hybridisation

Baked filters were soaked for 60 mins in a solution of 2 x SSC, 30% (v/v) dimethylsulphoxide (DMSO) at room temperature (Legault-Demare *et al.*, 1967; De Ley and Tijtgat, 1970). The labelled DNA probe was mixed with 75 $\mu$g of cold homologous carrier DNA, sonicated for 20 sec then an appropriate volume (0.1-0.3 $\mu$g of probe per filter) was heat denatured at 100$^{o}$C for 12 mins. Each soaked filter was placed in a polythene bag (Sterilin, Bibby U.K.) with 6 ml of 2 x SSC, 30% (w/v) DMSO and the denatured probe then heat sealed with the exclusion of all air bubbles. Sealed filters were incubted at 60 $^{o}$C in a water bath fitted with a circulator pump for 24 hr.

c) Washing of Filters

After hybridisation the filters were washed extensively in 2 x SSC at room temperature for 20 mins with gentle shaking to remove unbound probe. This was repeated twice then followed by incubation of the filter in 2 x SSC at 60$^{o}$C for 30 mins. Finally, the filter was dried over blotting paper at 42$^{o}$C for 60 mins.

d) Detection of Hybridisation

Both visual and numerical records of hybridisation were obtained as follows:

i) Visual Record - The filter was placed on a clean sheet of Whatman 3 MM filter paper and covered with cling film. It was then placed in an autoradiograph cassette (Harmer) and overlaid with a sheet of Fuji RX X-ray film and a cronex intensifying screen (Du Pont). Sealed cassettes

were stored at -20°C overnight. After exposure the film was developed (Kodak LX24, 1 in 6 dilution) and fixed (Kodak FX40, 1 in 5 dilution) and the position of black spots recorded. Spot intensity was taken as an approximation of the amount of probe DNA bound and therefore the degree of homology of the filter bound DNA.

ii) Numerical Record - Individual sample spots were removed from the filter by cutting around the sample square and one surrounding square in all directions with scissors. The filter pieces were placed into 3 ml amounts of Beckman EP scintillation fluid and the radioactive content of each determined by counting in an LKB mini-beta scintillation counter for 10 mins. Since the amount of DNA loaded on each square was constant, then, by comparing the counts on test squares with the amount found on squares with the homologous DNA, an accurate value for relative % binding (% homology) could be calculated.

$$\text{Relative \% binding (\% homology)} = \frac{\text{Amount of Probe bound to heterologous DNA}}{\text{Amount of Probe bound to homologous DNA}} \times 100$$

3.10.2 Individual Filter Method

a) Filter Preparation

80 μg lots of DNA were heat denatured by boiling for 10 mins, cooled on ice, made up to 6 x SSC with 20 x SSC and passed through Millipore-type GSWP nitrocellulose filters (25 mm diam.; presoaked with water for 10 mins then 6 x SSC for 10 mins) fixed in a sintered glass filtration apparatus at moderate speed. Filters were washed with 10 ml of 6 x SSC, dried at 42°C then baked in a vacuum oven at 90°C for 2 hrs. 25 mm

diam. filters had an effective binding area of 254 $mm^2$ (18 mm diam) and were saturated by 80 μg of DNA (De Ley and Tijtgat, 1970) (based on the assumption that 1 $cm^2$ of filter is saturated by 32 μg of DNA). The amount of DNA fixed was determined spectrophotometrically by measuring the absorbance at 260 nm of the solution before and after filtration. Filters of 5 mm diam. containing approximately 7.9 μg of DNA were punched from the larger filter and stored at 4°C before use.

b) Hybridisation Conditions

Individual 5 mm diam. filters were soaked for 60 mins in 400 μl of 2 x SSC, 30% (w/v) DMSO in flat bottomed Eppendorf tubes (3 ml). A volume of probe (approxm. 0.5 μg) mixed with homologous carrier DNA was sonicated (20 sec) and denatured by boiling for 12 mins. This was cooled on ice and mixed with an appropriate volume of 2 x SSC, 30% (v/v) DMSO. Filters were transferred to 400 μl of the fresh probe solution and incubated at 60°C for 24 hrs. All filters were included in triplicate.

c) Washing of Filters

After hybridisation, filters were washed with 500 μl lots of solutions as described in Method 1.

d) Quantitation of Hybridisation

The amount of hybridised probe was quantified by liquid scintillation counting of the dried filters and 40 μl aliquots removed from each hybridisation mixture. Relative % binding of heterologous DNAs were determined as in method 1. This individual filter method was also used

with dissected filters prepared as in Method 1.

### 3.10.3 Use of Formamide

Some hybridisations were repeated by methods 1 and 2 but using Denhardts solution (Denhardt, 1966) and formamide (McConaughy *et al.*, 1969) in place of DMSO. The prehybridisation soak in 2 x SSC, 30% (w/v) DMSO was replaced by incubation of the filter(s) with 2 x SSC, 25% (v/v) formamide, 1 x Denhardts solution (see below) at 57°C for 3 hrs. Hybridisation was at 57°C for 24 hrs in the same solution as above, but with the denatured probe added. Washing was as described in Method 1, but was followed by a final wash in 0.5 x SSC at 57°C for 20 mins. Autoradiography and scintillation counting were as before.

Denhardts Solution (100x concentration)

2% (w/v) Ficoll 400

2% (w/v) Polyvinylpyrrolidine (PVP)

2% (w/v) Bovine serum albumen (BSA)

Formamide

Before use formamide was brought to pH 7 and deionised by stirring with 1-2 g of amberlite monobed resin 3B (Sigma) per 100 ml. When the correct pH was reached the resin was removed by filtration under vacuum and the formamide stored at 4°C in the dark.

### 3.10.4 Determination of the Thermal Stability of DNA:DNA Hybrids

After liquid scintillation counting, each filter was removed and rinsed twice in 10 ml of toluene at room temperature to remove all traces of

scintillation fluid. Toluene was then removed by evaporation under slight vacuum. Filters were incubated separately for 20 mins each in a series of glass scintillation vials containing 1.5 ml of 1.5 x SSC, 25% (v/v) formamide in shaking water baths of increasing temperature. Temperatures ranged from 50$^{o}$C to 90$^{o}$C in incremental 5$^{o}$C steps. Filters were transferred with blunt forceps from one vial into the next already equilibrated to the subsequent temperature. The amount of radioactivity in each vial was measured after mixing the contents with 7 ml of scintillation fluid and counting in an LKB mini beta liquid scintillation counter for 5 mins. Filters after the DNA dissociations were rinsed briefly in 0.1 x SSC, air dried then counted in the scintillation counter.

The melting temperature (TM(e)) was calculated as the temperature at which 50% of the bound DNA probe was eluted from the filter. Results were expressed as the depression of the hybrid melting temperature (-$\Delta$Tm(e)), when compared to that of the homologous duplex.

## 3.11 Analysis of DNA Restriction Patterns on Agarose Gels

Restriction endonuclease digests of chromosomal DNA gave distinctive reproducible patterns of fragments when separated by agarose gel electrophoresis. Restriction endonucleases were obtained from Amersham and used with the following buffers as recommended by the manufacturer: EcoRI - 100 mM NaCl, 50 mM Tris, pH 7.5, 10 mM $MgCl_2$, 1 mM dithiothreitol. Hind III, Pst I - 50 mM NaCl, 10 mM Tris, pH 7.5, 10 mM $MgCl_2$, 1 mM dithiothreitol.

Each digestion mixture contained 3 μg of DNA, 2 μl of the appropriate 10 x restriction buffer and 5-15 U of restriction endonuclease made up to 20 μl final volume with sterile double distilled water. Restrictions were carried out at 37°C overnight. The entire volume of each digestion was mixed with 5 μl of gel loading buffer and loaded into a slot on a 0.7% (w/v) TAE agarose gel with no ethidium bromide. After electrophoresis at 30V (constant voltage) for 16 hrs the gels were stained in 500 ml of TAE buffer with 2 μg $ml^{-1}$ ethidium bromide for 20 mins. Gels were photographed on a UV transilluminator with a Polaroid camera.

## 3.12 Detection of Plasmids

Plasmids are autonomous, self replicating DNA structures indigeneous to many bacterial populations. Such extrachromosomal elements can constitute over 10% of the total cellular DNA and their presence can have a profound effect on hybridisation analysis. There is a vast range of methods both for screening bacterial populations for the presence of plasmids and for isolating any plasmids which may exist. In an attempt to screen the PPFM strains used in this study two differing strategies were adopted.

### 3.12.1 In - Gel Lysis Method

This was a variation on the method of Eckhardt (1978) as modified by Simon (1984).

3 ml aliquots of late exponential cultures of PPFMs grown on MS + methanol were pelleted, washed, and resuspended in 500 $\mu l$ of saline EDTA (0.15 M NaCl, 0.01 M disodium EDTA, pH 7.0). 200 $\mu l$ of acetone were added with mixing and left at room temperature for 5 mins. Cells were then pelleted and washed three times in saline EDTA before resuspension in 40 $\mu l$ of lysis buffer (10% (w/v) Ficoll 400, 0.2% (w/v) bromophenol blue, RNAase (50 $\mu g\ ml^{-1}$), lysozyme (2 mg $ml^{-1}$)). After incubation at 37°C for 1 hr, the cell suspensions were loaded into the wells of a 0.7% (w/v) agarose gel in TBE. Immediately behind these, on the anode side, a series of adjacent wells had previously been filled with a solution of 1.5% (w/v) agarose, 2% (w/v) SDS in TBE. Complete lysis was achieved by elecrophoretic transfer of SDS into the wells containing the bacterial cells at 2 mA for 60 mins. DNA bands were separated by further electrophoresis at 100 V for 3 hrs. The gel was stained and visualised under UV light as described previously.

When *E. coli* strains were to be used the initial acetone treatment and washing was omitted. Cell pellets were resuspended in lysis solution and loaded straight into the wells to be left for 1 hr at room temperature before electrophoresis.

3.12.2 Alkali Lysis Method

This was the method of Birnboim and Doly (1979), adapted by Ish-Horowitz (Maniatis *et al.*, 1982) and modified to include the solvent extraction and washing steps as described above. DNA bands were separated by electrophoresis on a 0.7% (w/v) agarose mini gel in TBE (60 mA for 90 mins) and visualised under UV light.

*E. coli* strains were treated as described by Maniatis *et al.* (1982).

3.12.3 Conjugative Transfer of Plasmids into PPFM strains

The problems encountered when lysing the cells of PPFM strains may greatly affect the efficiency of the above two techniques in detecting indigeneous plasmids.

To check on the suitability of these screening methods for the PPFMs, well characterised plasmids were transferred by conjugative mating from an *E. coli* donor into strains *M. organophilum* NCIB 11278 and *Ps. extorquens* NCIB 9399 with compatible antibiotic markers. The donor strain used was *E. coli* K12 JC6310 containing pUB 307 (RP4 derivative, $Tc_{10}^{r}$, $Km_{50}^{r}$) and pKT210 (RSF1010 derivative, $Sm_{100}^{r}$, $Cm_{50}^{r}$). pKT210 was non-conjugative but can be transferred at high frequency by pUB307.

One ml lots of mid log phase *E. coli* JC6310 and stationary phase PPFMs were pelleted, washed, resuspended in 1 ml of $PO_4$/20 buffer (35 mM $K_2HPO_4$, 22 mM $KH_2PO_4$) and filtered onto the surface of 25 mm Millipore GSWP mixed nitrocellulose/acetate filters. Filters were transferred to the surface of a nutrient agar plate, incubated at 30°C for 7 hrs then transferred to be incubated at 30°C in 100 ml lots of MS medium + 0.2% (v/v) methanol + kanamycin (40 $\mu$g $ml^{-1}$). This selects for PPFM transconjugants (kanamycin) and counterselects against *E. coli* donors (carbon source). After a further plate purification step, single colonies of both NCIB 11278, and NCIB 9399 containing pUB307 were obtained. pKT210 was either not transferred or failed to be expressed in the PPFM strains used. These transconjugant strains were included as

lysis controls in the aformentioned methods for plasmid identification.

### 3.13 Isolation of Ribosomal RNA (rRNA)

Unlike DNA isolation procedures, which required gentle lysis of the bacterial cell, RNAs could be obtained using more physical methods for cell breakage.

#### 3.13.1 Bulk RNA Isolation

All glassware and utensils used in the following method had been pretreated with chromic acid (sulphuric acid saturated dichromate solution) to eliminate any contaminating RNAase activity. 2.5 g (wet weight) of PPFM cells were resuspended in 20 ml of a solution of 0.01 M Tris-HCl, pH 7.5; 0.05 M NaCl; 0.5% (w/v) sodium naphthalene 1,5-disulphonate and broken by passage through the French Press at 1.3 x $10^8$ Pa. An equal volume of phenol:water:8-hydroxyquinoline:m-creosol (50:10:0.05:7) was added and stirred gently for 20 mins at 4°C. Phases were separated by centrifugation at 5000 g for 25 mins in an MSE Chilspin. The upper aqueous phase was removed, made 1% (w/v) with Tri-isopropylnaphthalene sulphonate and 6% (w/v) with 4-aminosalicylate and stirred with an equal volume of the phenol mixture at 4°C for 20 mins. After centrifugation, the aqueous phase was made 0.5 M with NaCl, stirred with an equal volume of the phenol mixture at 4°C for 20 mins and the phases separated by centrifugation. RNA was precipitated from the aqueous phse by addition of 2.2 volumes of 95% ethanol, 5% 4M sodium acetate and incubation overnight at -20°C. Pelleted RNA was dried under vacuum and stored at -20°C.

3.13.2 Separation of rRNAs - Formamide Agarose Gels

Individual RNA species were separated by their differing electrophoretic mobilities under denaturing conditions. The following method used formamide as the denaturing agent.

RNA samples were run on a 1.5% (w/v) flat bed agarose gel in E buffer/formamide.

| 5 x E buffer | | |
|---|---|---|
| | 43.5 g | Trizma base |
| | 60.84 g | $NaH_2PO_4.2H_2O$ |
| | 3.7 g | disodium EDTA |
| | all per litre of double distilled water | |

Formamide was deionised as described in 3.10.2. The gel and running buffer were made from 600 ml of formamide + 576 ml of double distilled water + 24 ml of 5 x E buffer. Gels were prerun for 15 mins at 20 mA (constant current). Samples were loaded with 2 volumes of gel loading buffer - formamide and electrophoresed at 40 mA (constant current) for 3-4 hrs.

| Gel Loading Buffer | | |
|---|---|---|
| | 1.2 ml | 5 x E Buffer |
| | 3.8 ml | $H_2O$ |
| | 0.9 g | Ficoll 400 |
| | 0.02% (w/v) | bromophenol blue |

Prior to use this was mixed with 15 ml of formamide. After electrophoresis, gels were soaked in water for 15 mins to remove the

formamide then stained in 0.5 M NaCl, 5 $\mu g\ ml^{-1}$ ethidium bromide for 20 mins. Background ethidium bromide was removed by soaking in fresh 0.5 M NaCl for 15 mins and then gels were fixed in 10% (w/v) acetic acid for 20 mins. Upon visualisation under long wave UV light, the relevant bands were dissected from the gel using a sharp scalpel blade.

#### 3.13.3 Electroelution

Gel slices containing nucleic acid were dissected from agarose gels with a sharp scalpel blade and sealed with a minimal volume of 0.5 x TBE buffer in dialysis tubing (Scientific Industries International Ltd.). These were electrophoresed at 100V (constant current) for 3 hrs in a shallow layer of 0.5 x TBE buffer in an electrophoresis tank. The polarity of the current was reversed for 2 mins to release the nucleic acid from the wall of the dialysis bag. The buffer from around the gel slice was removed, the bag washed with 500 $\mu l$ of fresh 0.5 x TBE and the lot extracted once with phenol then twice with chloroform:isoamyl alcohol (24:1). Nucleic acids were removed by ethanol precipitation (see 3.2.2.1) then resuspended in an appropriate volume of TE (10 mM Trizma base, 1 mM $Na_2$ EDTA, pH 8.0) buffer.

### 3.14 Blotting and Hyridisation of Restricted DNA Fragments

#### 3.14.1 Transfer of DNA from Agarose Gels to Nitrocellulose

The resolution of restricted DNA fragments on agarose gels, staining, DNA denaturation and blotting onto nitrocellulose were carried out as described by Southern (1975).

### 3.14.2 Hybridisation of Filters

Blotted filters were hybridised as described previously but under the conditions with formamide as the denaturing organic solvent. Hybridisations were in 10 ml of 2 x SSC, 25% (v/v) formamide, 1 x Denhardts solution, 50 $\mu g\ ml^{-1}$ denatured herring sperm DNA with probe DNA added at 57°C for 36 hrs. Washing and autoradiography of the filters were as described previously.

## 3.15 DNA:rRNA Hybridisation

DNA:rRNA hybridisations were performed by the method of De Ley and De Smedt (1975) but using 5 $\mu g$ of filter bound DNA and 1 $\mu g$ of rRNA incubated in 200 $\mu l$ of hybridisation solution in small sealed tubes. When necessary, filters were counted for 20 mins in 3 ml aliquots of scintillation fluid (Beckman EP) in an LKB mini-beta scintillation counter.

# Results and Discussion

## 4. DNA Isolation and Base Composition

### 4.1 Cell Lysis

Batch cultures of PPFMs grown on MS medium + 0.2% methanol yielded 1.0-2.0 g wet weight of cells $l^{-1}$ when harvested at late log phase after 2-5 days growth at 30°C. Only rarely was the yield better after growth on NB or GP medium.

A major problem when working with DNA from the PPFMs is its purification compounded by problems in obtaining gentle lysis of the cells to minimise shearing. Of the strains tested only strain B-46 could be lysed fully (as determined by phase contrast microscopy and visual checks on the turbidity and viscosity of the cell suspension) by incubating with lysozyme at 37°C. Progressively harsher and more involved lysis procedures enabled an increasing number of strains to be broken but other strains showed little appreciable cell lysis under most conditions. It must be noted that not all strains were tested under each condition. The procurement of an effective lysis method was concentrated on selected strains, namely NCIB 11561, 11278, 9421 and 9399. Strains NCIB 9399 and 9421 proved to be resistant to lysis by most of the treatments listed in section 3.2.1 of the Methods. Method (q) (alkali pretreatment) proved effective in breaking the cells but gave concomittant degradation of the DNA. In general antibiotics, catabolic enzymes and physical shock treatments had no effect on cellular integrity and ease of lysis. The only lysis method found to be applicable to all strains tested was that of lysozyme and detergent treatment following a solvent pre-treatment step. Chloroform proved to

be an effective solvent but was difficult to remove by washing. Acetone and butanol were equally effective and easily removed by washing. This treatment gave 70-100% lysis of many of the strains and 20% lysis of even the most refractory PPFMs. All DNAs extracted for hybridisation studies were obtained using this same lysis method. Limited analysis of cell wall components (Sato, 1978; Urakami and Komagata, 1979; Patt and Hanson, 1978; Nishimura *et al.*, 1983) and electron microscopic analysis of cellular thin sections (Patt *et al.*, 1974; P. Green, personal communication) have been carried out with representatives of the PPFMs. No unusual extracellular layers constituting a barrier to lysis have been identified. Resistance to the action of common lytic agents could be due partly to the high proportion of lipoproteins present in the cell wall (Patt and Hanson, 1978; Ito and Iizuka, 1971), these being partially denatured by the brief solvent pretreatment step.

4.2 DNA Purification

Method (1) (3.2.2.1) proved the most useful for DNA preparation and routinely gave pure DNA of a high molecular weight. With Method (2) DNAs prepared from some strains were prone to contamination. Absorbance ratios for the purified DNAs were generally within the limits of A280/A260 0.50-0.54, A230/A260 0.45-0.49 and % A320/A260 ≤2%. DNAs with abberant ratios were further purified but often appeared pure when subject to agarose gel electrophoresis or diphenylamine determination. The reason for consistent atypical ratios (especially A230/260) with some DNAs remains unanswered.

PPFMs which could be lysed efficiently yielded 1-2 mg of DNA per 2 g wet weight of cells. Strains showing less efficient lysis gave as little as 100 $\mu$g of DNA per 2 g wet weight of cells. A few strains could be lysed efficiently but regularly gave low DNA yields, perhaps due to abnormally fragile DNA or to potent nuclease activity within the cells. Lysis, but not relative DNA yields, was found to be dependent to a minor extent on growth conditions and storage of the cell pellets once harvested.

As a general rule strains in cluster 11 under the classification of Green and Bousfield (1982) were more resistant to lysis and gave lower DNA yields than strains in their cluster 10. Cluster 11 DNAs proved to be more sensitive to shearing after treatment with phenol.

*In vivo* labelling with $^3$H adenosine gave DNAs of specific activities ranging from $5 \times 10^3$ to $6 \times 10^4$ cpm $\mu g^{-1}$. DNAs labelled similarly with $^3$H guanosine had specific activities only 10-20% of these values. DNA nick translated *in vitro* had specific activities of $10^6$-$5 \times 10^6$ cpm $\mu g^{-1}$ with $^3$H dATP radiolabel and $10^7$-$4 \times 10^7$ cpm $\mu g^{-1}$ using $^{32}$P dCTP radiolabel.

### 4.3 DNA Base Composition

The base composition (expressed as % (G+C) content) of DNAs isolated from representatives of the PPFMs and calculated from their buoyant density values by the method of Mandel *et al*. (1968), are listed in Table 5. Values for the melting temperatures (Tm) were obtained from

the equation

19) $Tm = 69.3 + (0.41 \times \% (G+C))$ (Mandel *et al.*, 1968)

Some values were further estimated by the thermal denaturation method in the laboratory of De Ley and coworkers. The values of NCIB 11278 71.1% (70.5%); strain 317 71.1% (71.4%); NCIB 2879 68.8% (69.1%) obtained show excellent correlation with those found by the buoyant density method (in brackets), in accord with the findings of De Ley (1970c). Table 5 lists any previously published values for the base composition of PPFM DNAs as determined by the thermal denaturation method. In general these are lower than the values obtained in this study. The study of Green and Bousfield (1982) indicated a range of % (G+C) contents from 60.5-67.3% and 64.2-70.5% for their clusters 10 and 11 respectively. Most likely this reflects some of the difficulties experienced in obtaining pure DNA preparations from these organisms. DNA contaminants can lower the apparent Tm, and hence the % (G+C) content, moreso when determined by the thermal denaturation method.

The DNA base composition for all the PPFM strains tested fell within a narrow 4% range of 68.4-72.4 mols %. Across the bacterial spectrum the range extends from about 24-25% (for certain mycoplasmas) to 76-77% (for certain streptomyces). The PPFMs come close to the top of this range but as yet the evolutionary significance and any advantage conferred by a high % (G+C) content are uncertain.

A narrow range of base composition for the PPFMs is encouraging as it is

mols %. The converse is not true in that coincident base composition values are not necessarily indicative of relatedness (Schleifer and Stackebrandt, 1983).

The question of whether fixed boundaries should be set for the range of % (G+C) contents for taxa of various levels is much debated. Some genera have very narrow, and others very wide, ranges of DNA base composition. In general, ranges of 5 and 10% are accepted as the limits for strains included as members of a common species and genus respectively (Bradley, 1980). The latter is also regarded as the useful limit for DNA homology studies. Base composition values are now an obligatory component of the minimum list of data required for the description of a new species.

The only conclusion that can be drawn from the base composition results in isolation is that the PPFMs are not necessarily totally unrelated.

Table 5 Buoyant Density, % (G+C) Content and Melting Temperatures for DNAs Isolated from PPFMs

| Strain | Buoyant Density in CsCl 20°C (g cm$^{-3}$) | (G+C) Content (mols %) | Melting Temperature (°C) (Tm) |
|---|---|---|---|
| 317 | 1.7310 | 72.4[a] | 99.0 |
| 0.31 | 1.7303 | 71.8 | 98.7 |
| 10610 | 1.7280 | 69.4 | 97.8 |
| 10815 | 1.7309 | 72.3 (65.0)[b] | 98.9 |
| N-12 | 1.7289 | 70.3 | 98.1 |
| 11278 | 1.7292 | 70.5 (66.0)[c] | 98.2 |
| 2CR | 1.7291 | 70.5 | 98.2 |
| 617 | 1.7294 | 70.8 | 98.3 |
| R14 | 1.7305 | 71.9 | 98.8 |
| 434 | 1.7278 | 69.2 | 97.7 |
| 9399 | 1.7280 | 69.4 (65.6)[d] | 97.8 |
| 9421 | 1.7304 | 71.8 | 98.7 |
| 10598 | 1.7284 | 69.8 | 97.9 |
| 2879 | 1.7277 | 69.1 (65.6)[d] | 97.6 |
| 539 | 1.7272 | 68.6 | 97.4 |
| B46 | 1.7271 | 68.4 | 97.3 |
| M159-1 | 1.7282 | 70.6 | 97.8 |
| 10609 | 1.7287 | 70.1 | 98.0 |
| CS51 | 1.7298 | 71.2 | 98.5 |
| 11561 | 1.7285 | 69.9 (65.8)[e] | 97.9 |
| 135 | 1.7289 | 70.3 (66.1)[d] | 98.1 |
| 10611 | 1.7294 | 70.9 | 98.4 |
| D12 | 1.7286 | 69.9 | 98.0 |
| *R. sphaeroides* | 1.7277 | 69.1 | 97.6 |
| *M. lysodeikticus* | 1.7312 | 72.6 | 99.1 |

[a] values ± 0.5%

[b] Ito and Iizuka, 1980

[c] Patt *et al.*, 1974

[d] Urakami and Komagata, 1984

[e] Austin and Goodfellow, 1979

Values measured against *E. coli* K12 DNA - buoyant density = 1.710 g cc$^{-3}$

Values in brackets are those previously published for the organisms concerned.

## 5. PAGE Analysis of Soluble Proteins

### 5.1 Preparation of Soluble Protein Extracts

All members of the PPFMs and the *Rhodospirillaceae* included in this analysis could be broken by repeated bursts of sonication. The average protein content of the cell free extracts was in the range of 1.3-7 mg $ml^{-1}$. Aliquots of 0.6-3.5 $\mu l$ were required to load each individual gel slot with 10 $\mu g$ of protein.

### 5.2 Gel System

#### 5.2.1 Staining Methods

Duplicate gels run under identical conditions were stained one with coomassie blue and one by the silver stain method. The results indicated that the silver stained gels gave sharper, intenser and more numerous bands better suited for scanning by the microdensitometer. Such gels require less protein and have relatively less general background trend than is found with coomassie stained gels (Kersters and De Ley, 1980; Jackman, 1982).

#### 5.2.2 Gradient Gels

Initial trials with 10-30% gradient polyacrylamide gels proved to give good resolution of the total protein banding pattern, especially over the lower end of the molecular weight range. Fifty to sixty-five distinct bands were obtained for each electrophoretogram. A problem, however, with gradient gels was that on staining the regions of high

acrylamide concentration rehydrated to a greater degree than the lower % regions resulting in a characteristic bowing of the gel matrix. The microdensitometer can scan only linearly and such bowing precludes accurate scanning of all but the centremost gel lanes. Similar problems have been encountered by other workers (Jackman, 1982) and gradient gels are invariably used for only qualitative comparisons of protein profiles.

### 5.2.3 Linear Gels

The use of linear 10% polyacrylamide gels avoided the problems mentioned above with little loss in the number of detectable protein bands. All gel lanes could be run in parallel. Occasionally the outermost tracks did still exhibit some bowing but these were ignored during subsequent analysis. Acrylamide at a concentration of 10% is an average of that used by other workers. Moore *et al.* (1980) reported that over a range of 7-12% there was no single concentration of acrylamide that produced the most distinctive patterns between different strains.

The denaturing SDS-PAGE system has been used relatively little in taxonomic studies (Kersters and De Ley, 1980) but was found to produce a much more uniform banding pattern along the entire gel profile than did a non-denaturing system. This allowed greater discrimination between some strains.

### 5.2.4 Protein Molecular Weight Markers

Ten to thirty % gradient and later, 10% linear gels were run with a mixture of 6 proteins of known molecular weights (Pharmacia) in the

outermost slots. These can act as molecular weight reference markers for each electrophoretogram run on any gel. Because of the problems mentioned above, this is most useful in quantitative analysis with a linear gel system, but even with linear gels, the effective track length can vary from the middle to the outermost tracks. Ideally, multiple lots of standard proteins should be run at intervals across each gel but this limits the number of test tracks available. Molecular weight standards are of paramount importance as a reference for companding and later gels circumvented these problems by a direct addition of high and low molecular weight protein markers to each soluble protein extract prior to electrophoresis. Being at either extreme of the range of molecular weights effectively resolved by the gel system, these markers are ideal for the companding of individual electrophoretograms. A similar system has been used by Kersters and De Ley (1975) but running each protein mixture twice, once with and once without reference protein markers. By superimposing the densitometric traces the position of the standards were marked on the tracing of the latter prior to analysis. The method of Jackman (1983) incorporated the soluble extract from a known reference strain in the central slot of each gel and relied on uniform electrophoresis within each gel for accurate comparison. The method of Feltham and Sneath (1979) used no marker proteins and relied on the stacking/resolving gel interface and the lower most band from each gel lane as markers for length normalisation. Studies where comparisons of electrophoretograms are only qualitative also often include a known reference strain on each gel (Moore *et al.*, 1980; Biavati *et al.*, 1982).

## 5.3 Reproducibility

Kersters and De Ley (1980) and Jackman (1983) indicated that the most critical factor influencing the potential application of the PAGE-protein system to bacterial taxonomy is its reproducibility. The importance of reproducibility for reliable quantitative or even qualitative analysis cannot be over emphasised. All aspects of the procedure from sample preparation, electrophoresis, microdensitometry through to electrophoretogram analysis must be rigorously controlled to allow meaningful comparisons to be made. Unlike the system of Kersters and De Ley (1975) where initial electrophoretogram analysis was manual, this study used a system where all computations were computer controlled, effectively eliminating any human error at this stage. The emphasis, therefore was placed on obtaining electrophoretograms by a highly ordered and standardised procedure with minimal experimental fluctuations.

### 5.3.1 Scanning

The reproducibility of gel scanning was investigated by scanning 3 unique lanes from individual gels ten times each along the same scan path and ten times each along individual scan paths. With the AUTO CA method of peak search and position analysis repeat scans along the same path had average similarity values of 1.0, 1.0 and 0.98. Using the AUTO EF method of correlation coefficient analysis, repeat scans along the same path had average r values of 0.99, 0.99 and 0.98 at low resolution (100 points) and mean values of 0.99, 0.98 and 0.97 at high resolution (450 points). Repeat scans along individual paths had mean r values of 0.99, 0.98 and 0.98 at low resolution and 0.99, 0.97 and 0.97 at high

resolution. Such values compare favourably with those reported previously (Jackman, 1983).

### 5.3.2 Preparation of Soluble Proteins

Correlation coefficient analysis on 3 preparations of soluble proteins from a common batch of cells run in unison on 3 gels gave a mean r value of 0.97. This is in agreement with the findings of Kersters and De Ley (1975), that any particular preparation method can be reproducible when carefully controlled. Moore *et al.* (1980) reported that sonication was not a reproducible method of sample preparation but use of an ice-methanol bath alleviates the problems with protein denaturation that they encountered. Kersters and De Ley (1975) have reported that comparisons between proteins prepared by different methods are not so valid.

One-half of a soluble protein preparation was subject to repeated freeze-thawing (at -20$^{o}$C) for 20 cycles. When run with the untreated half on 3 separate gels, correlation coefficient analysis gave a mean value for comparison of r = 0.99. Thus, freeze-thawing of samples had a negligible effect on the resultant electrophoretograms.

### 5.3.3 Culture Medium

Two strains (NCIB 9399 and R14) were grown up to an equivalent OD in a variety of growth media and the resultant protein extracts run together on 3 separate gels, scanned and analysed by AUTO EFP correlation coefficient analysis (Table 6).

Table 6 AUTO EFP Analysis of Electrophoretograms from Cells Cultured on a variety of Growth Substrates

(low resolution analysis)

| Strain | Substrate(s)* | Mean r values |
|---|---|---|
| 9399 | MA-GP | 0.91 |
| (*Ps. extorquens*) | MA-NB | 0.89 |
| | MA-PM | 0.87 |
| | MA-MS | 0.85 |
| | GP-NB | 0.96 |
| | GP-PM | 0.90 |
| | GP-MS | 0.89 |
| | NB-PM | 0.94 |
| | NB-MS | 0.89 |
| | PM-MS | 0.85 |
| R14 | NB-GP | 0.96 |
| | NB-MS | 0.89 |
| | GP-MS | 0.92 |

* NB - Nutrient broth; PM - pyruvate malate; MA - mineral salts + glycerol; GP - glycerol peptone; MS - mineral salts + methanol.

In general the complex media tended to give more similar electrophoretograms and the defined media yielded more diverse profiles. As can be seen from Figure 9, there were very few major differences between the profiles in terms of unique bands. The most significant differences were in the intensities of some fairly major bands, these being emphasised by the correlation coefficient analysis. Differences were less with peak search and position analysis but this indicated the importance of culturing the PPFM strains in an identical medium to minimise variability. Hanna *et al.*, 1983 reported no differences in the resultant electrophoretograms when strains were cultured on different media. In their study, however, there was only qualitative analysis of the protein profiles and therefore, not affected by changes in band intensities. Moore *et al.* (1980) reported that the level of background varies with the growth medium in agreement with the results in this study. Kersters and De Ley (1975) found no significant differences between cells grown in liquid or solid media of the same composition.

5.3.4 Culture Age

Another factor which may be critical in obtaining reproducible electrophoretogram comparison is the age of the culture when cells are harvested. Strains NCIB 9399 and M159-1 were grown up to approximate early, mid and late log phases of growth in the same medium and their protein extracts electrophoresed to detect quantitative differences. The results in Table 7 show average correlation coefficients for comparisons between the electrophoretograms. Again, as shown in Figure 10, the main differences were in band intensities with only a few unique bands appearing.

Table 7 AUTO EFP Analysis of Electrophoretograms from Cells Harvested at Stages through the Growth Cycle

| Comparison | r value | Comparison | r value |
|---|---|---|---|
| 9399P - 9399E | 0.98 | M159-1P - M159-1E | 0.97 |
| 9399P - 9399M | 0.96 | M159-1P - M159-1M | 0.93 |
| 9399P - 9399L | 0.89 | M159-1P - M159-1L | 0.92 |
| 9399E - 9399M | 0.94 | M159-1E - M159-1M | 0.92 |
| 9399E - 9399L | 0.83 | M159-1E - M159-1L | 0.90 |
| 9399M - 9399L | 0.91 | M159-1M - M159-1L | 0.85 |

Abbreviations:

Cells harvested at approximate: E - early; M - mid; L - late log phase of growth cycle.

P denotes early-mid log phase extract normally used in analysis.

9399 - *Ps. extorquens* NCIB 9399.

Figure 9 Protein profiles for strains NCIB 9399 and R14 cultured on various growth substrates

Growth media - MS, mineral salts + methanol; NB, nutrient broth; GP, glycerol peptone; PM, pyruvate malate;
Tracks, 1 - strain 9399 on MS; 2 - 9399 on NB; 3 - 9399 on GP; 4 - 9399 on PM; 5 - R14 on MS; 6 - R14 on NB; 7 - R14 on GP.

9399 - *Ps. extorquens* NCIB 9399

Figure 10 Protein profiles for strains NCIB 9399 and M159-1 cultured on mineral salts medium with methanol and harvested at various culture turbidities

Cells harvested at: E, early; M, mid; L, late log phase of growth. P denotes early-mid log phase extract normally used in analysis.
Tracks, A - 9399 P; B - 9399 M; C - 9399 E; D - 9399 L; E - M159-1 P; F - M159-1 E; G -M159-1 M; H - M159-1 L.

FIG 9

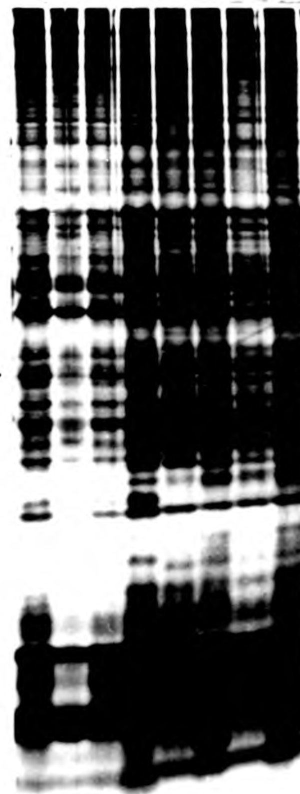



FIG 10

FIG 9

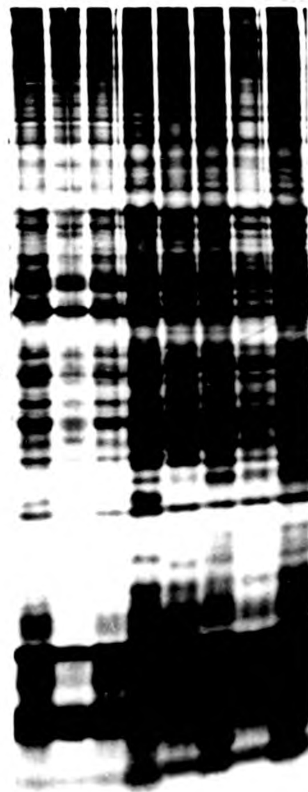



FIG 10

The results again emphasized the bias of the correlation coefficient mode of analysis to large broad peaks. The spread of r values was greater than that found after altering the growth medium (Table 6) and indicated that organisms must be harvested at as near an equivalent stage of growth as possible. Cells harvested at late log phase appear to give the most disparate comparison values. Hanna *et al.* (1983) and Kersters and De Ley (1975) reported only small differences in some minor bands from cells harvested at differing phases of growth. The average value of $r = 0.92$ from Table 7 is less than that reported by Kersters and De Ley (1975). The relative effect of the stage of growth will vary for different groups of organisms but for the PPFMs it appears to be an important influence on reproducible electrophoretogram comparisons.

### 5.3.5 Gels

Further factors with an influence on reproducibility are gel composition and electrophoretic conditions. Ionic impurities (Kersters and De Ley, 1975), β-mercaptoethanol (Marshall and Williams, 1984) and protein contamination from an external source (Ochs, 1983) have all been shown to influence gel patterns but can be easily controlled by using rigorously clean glassware, pure materials and a bit of care. More difficult to control are the precise electrophoretic conditions, homogeneity of the gel matrix and the reproducibility of staining. The first of these proved not too great a problem when the same powerpacks and electrophoresis tanks were used to run gels at a constant temperature. The second and third factors proved to be the most problematic and the ones having the most pronounced effect on reproducibility in this study.

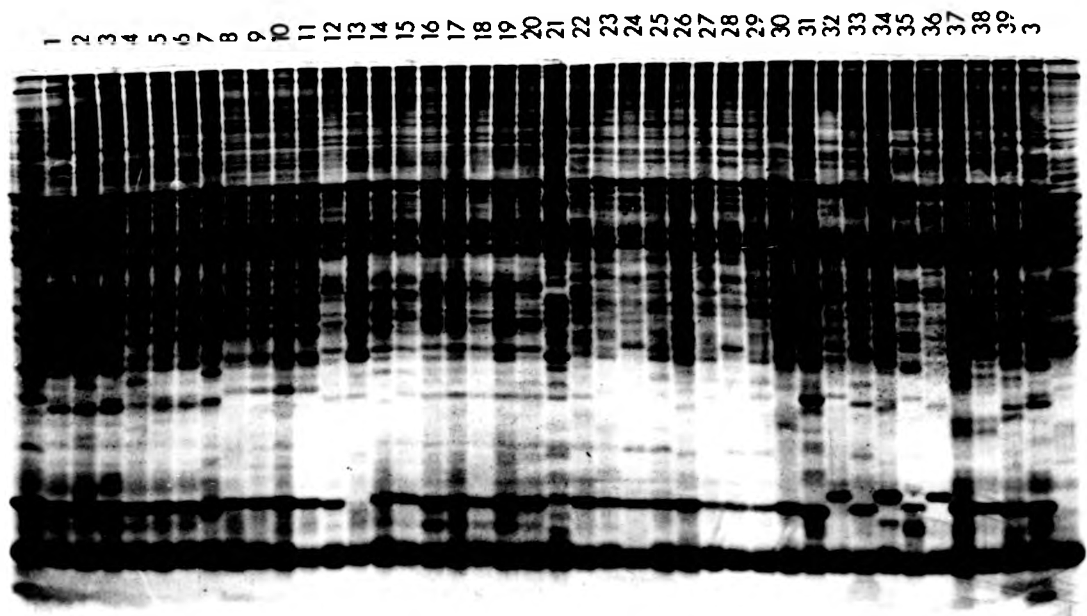
Concave gradient gels were of little use for quantitative analysis relying on relative distances because of localised heterogeneity in acrylamide gradients between consecutive gels. Linear 10% gels proved to be much more reproducible. The protein extract from strain NCIB 10609 was run on 2 randomly chosen lanes (to prevent proximity bias clustering) across 8 different gels. The mean similarities for comparison within each gel using the AUTO EFP and AUTO CA programs were r = 0.99 and r = 0.96 respectively. However, the mean similarity for inter-gel comparisons by AUTO EFP was r = 0.86 over a range of values from 0.73 to 0.97. This value is above that of r = 0.80 given as the lower limit of reproducibility for statistical significance (Sneath and Johnson, 1972; Jackman, 1983) but is below that found by Kersters and De Ley (1975). Moore *et al* (1980) reported that the most critical factor influencing reproducibility of gel composition was the sensitivity of acrylamide to $O_2$. In this study gels were prepared from common stock solutions but poured individually and thus it was difficult to have absolute control over the $[O_2]$ and polymerisation conditions. Ideally a cassette-type gel system as described by Jackman (1983) should be used if available. In this study the problem was circumvented by developing gels with increasing numbers of lanes and thus limiting the number of inter-gel comparisons required to compare all strains under study. Early 12 lane gels were modified to run with 21 lanes and were subsequently replaced by a larger gel system with 42 lanes. These gels were sufficiently large to accommodate all of the strains under study and allowed them to be run under identical electrophoretic conditions (Figure 11).

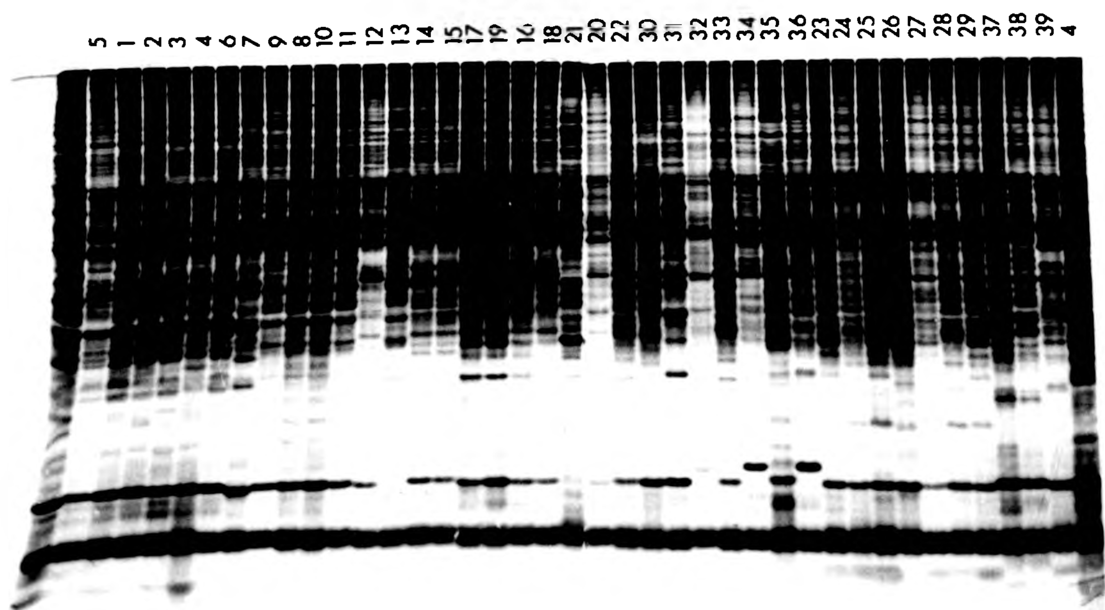
Figure 11 Soluble protein extracts from representatives of the PPFMs cultured under equivalent conditions

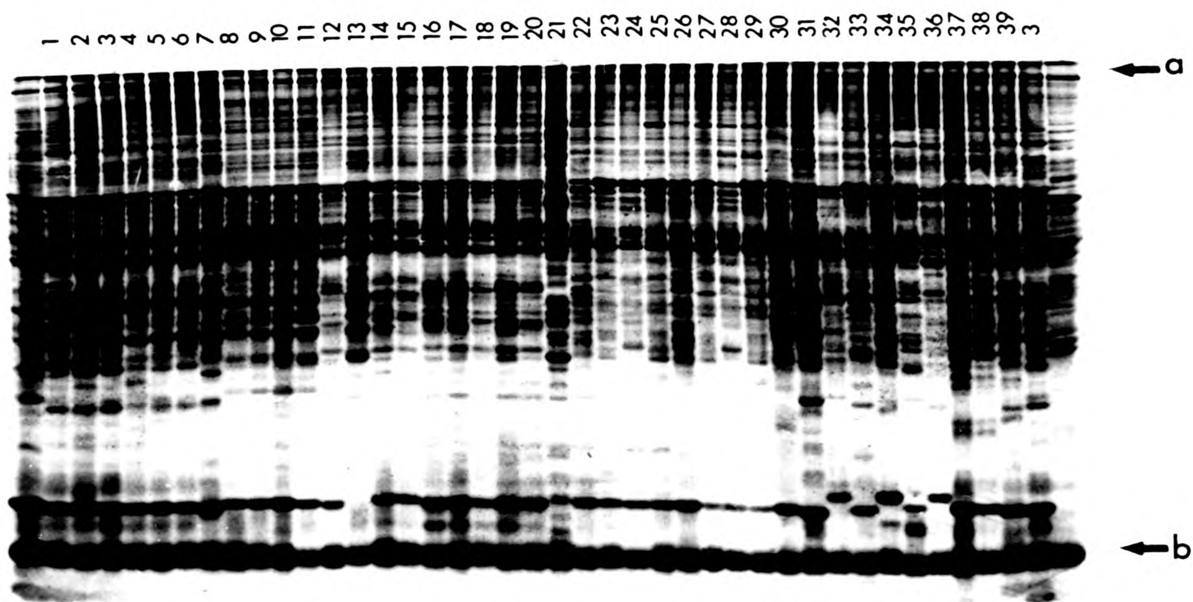
42 lane 10% linear gels. Silver stained.

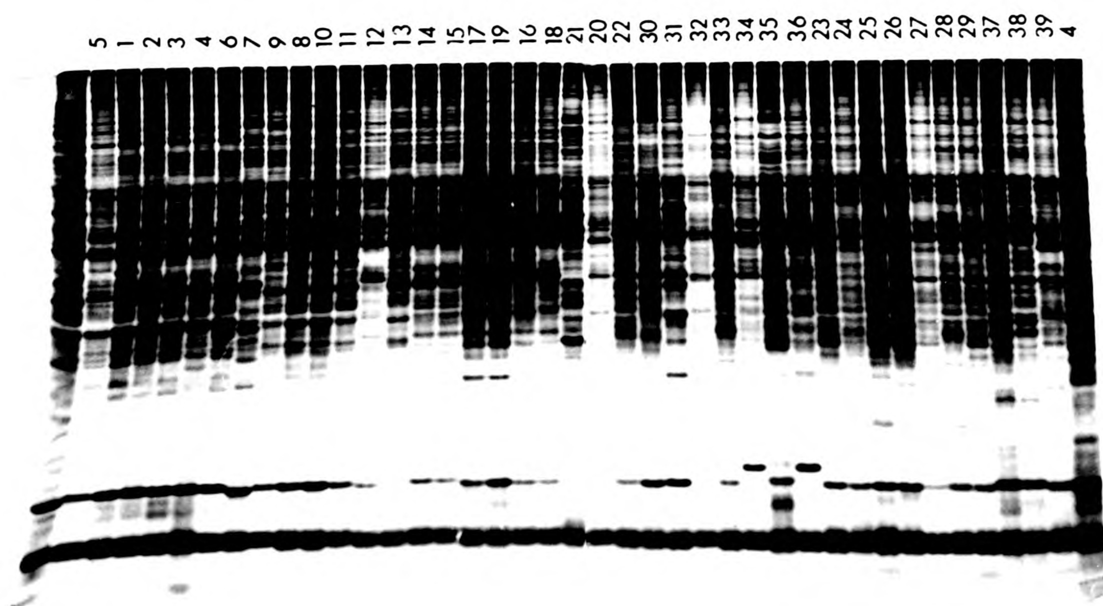
Strains: 1 - NCIB 2879; 2 - N-12; 3 - NCIB 9399; 4 - 539; 5 - NCIB 9133; 6 - NCIB 9686; 7 - NCIB 11278; 8 - NCIB 10609; 9 - 135; 10 - NCIB 10610; 11 - NCIB 10604; 12 - CR; 13 - N-6; 14 - NCIB 9141; 15 - NCIB 9145; 16 - NCIB 10611; 17 - NCIB 10598; 18 - CS51; 19 - M159-1; 20 - NCIB 10601; 21 - *E. coli* C; 22 - NCIB 10815; 23 - 0-31; 24 - 2CR; 25 - 617; 26 - NCIB 11561; 27 - 269; 28 - NCIB 9142; 29 - N-2; 30 - R14; 31 - NCIB 9421; 32 - B46; 33 - 434; 34 - D12; 35 - D5; 36 - D21; 37 - 317; 38 - 602; 39 - 790.

Protein bands indicated, a - thyroglobulin; b - cytochrome c.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 3
a
b


5 1 2 3 4 6 7 9 8 10 11 12 13 14 15 17 19 16 18 21 20 22 30 31 32 33 34 35 36 23 24 25 26 27 28 29 37 38 39 4


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 3
a
b


5 1 2 3 4 6 7 9 8 10 11 12 13 14 15 17 19 16 18 21 20 22 30 31 32 33 34 35 36 23 24 25 26 27 28 29 37 38 39 4

### 5.3.6 Staining

After electrophoresis gels were stained following a strictly adhered to protocol. Despite this, differences were found in the times required for development of the banding patterns indicating some variability inherent within the silver stain method used. Although most suitable for microdensitometry, final analysis revealed that repeat gels stained by the silver method contained many quantifiable differences in the intensities of both major and minor bands. This has no effect on analysis when all strains are run on a common gel matrix but it is thought that the variability in relative staining intensities was by far the largest single influence contributing to the disparate similarity values found for inter-gel comparisons. Consequently, it is recommended that for any similar quantitative study the staining of gels by the silver method be viewed with caution.

## 5.4 Peak Search and Position Analysis

Shown in Figure 12 is a tracing reproduced by computer after being stored on disc by program Joyce-L. Figure 12(b) shows an example of peak search on this trace under the parameters given.

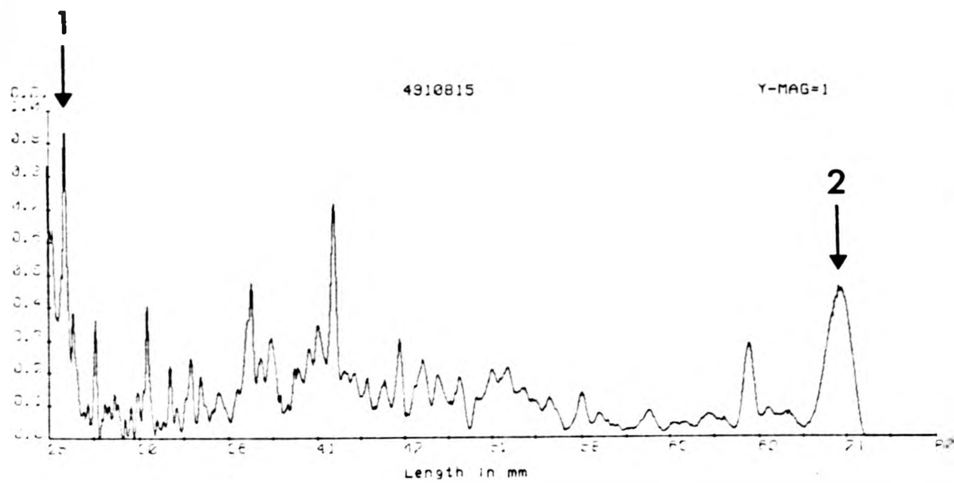
### 5.4.1 COMP 4

Figure 13 shows a plot from the comparison of a pair of companded electrophoretograms using program COMP 4. In this instance peak search parameters (of minimum height and width thresholds input in menu) were selected to include the major 40-50 peaks from each electrophoretogram for analysis but finer discriminatory analysis could be achieved with up

Figure 12 Peak search and positioning on an electrophoretogram of Ps. radiora strain NCIB 10815 soluble proteins

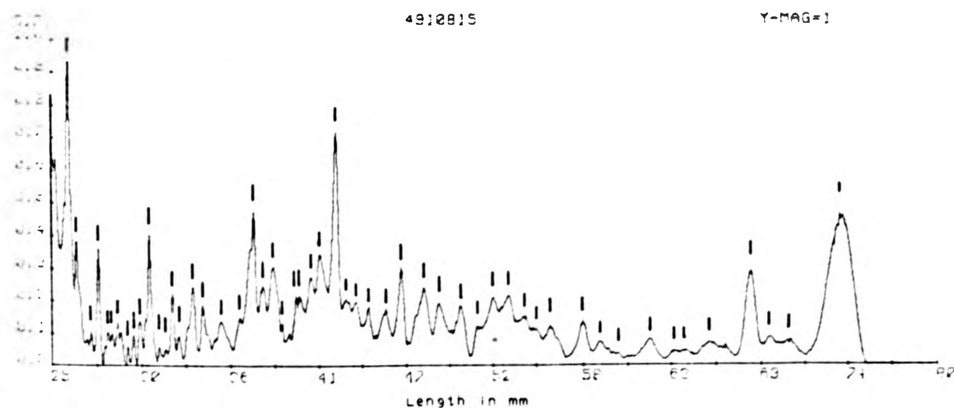
Trace has background trend removed. 51 peaks detected (ignoring standard protein peaks) under search parameters of height 20 and width 5.

Protein peaks indicated. 1 - thyroglobulin; 2 - cytochrome c.




| | |
|---|---|
| FILENAME | 4910815 |
| DATE | 1:6:85 |
| SCAN RATE | 100 |
| START SCAN AT | 25 |
| END SCAN AT | 80 |
| TRACK NUMBER | 1 |
| ABS. RANGE | 1 |
| COMMENTS | comment |

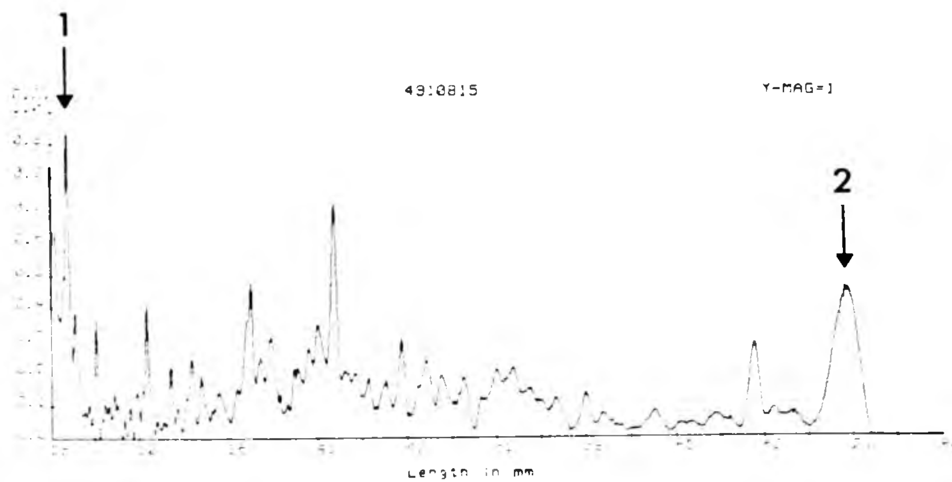
a



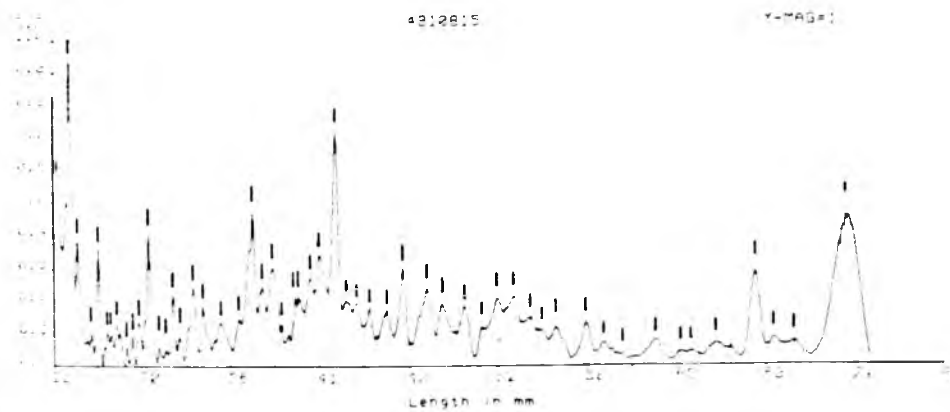


| | |
|---|---|
| FILENAME | 4910815 |
| DATE | 1:6:85 |
| SCAN RATE | 100 |
| START SCAN AT | 25 |
| END SCAN AT | 80 |
| TRACK NUMBER | 1 |
| ABS. RANGE | 1 |
| COMMENTS | comment |


b




a




b

Figure 13 Pairwise comparison of PPFM soluble proteins by peak search and position analysis using program COMP 4

Peak search parameters of minimum height 40 and width 5 used. Tracks indicated as C show protein bands marked as coincident for the pairwise comparison below.

Figure 13 Pairwise comparison of PPFM soluble proteins by peak search and position analysis using program COMP 4

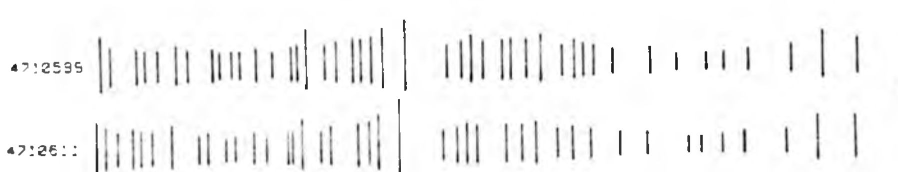
Peak search parameters of minimum height 40 and width 5 used. Tracks indicated as C show protein bands marked as coincident for the pairwise comparison below.

eins by peak

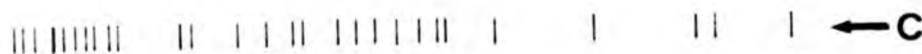
am COMP 4

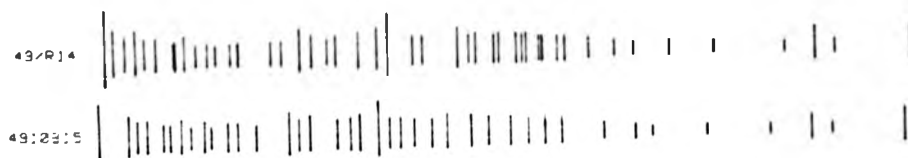
used. Tracks

or the pairwise




[a]



[b]

to 60 peaks. At this higher level, however, the peak dimensions were often such that they overlap with any spurious peaks and background noise levels from the microdensitometer. Gel analysis by COMP 4 requires the manual construction of a similarity matrix, its sorting and tracing of a dendogram making it very labour intensive. Consequently, COMP 4 was used primarily as a manual check for the correct functioning of the marker search, peak search, background and margin parameters and companding procedures on electrophoretograms from any gel prior to the use of the automated analysis programs.

5.4.2 DENDO

The program DENDO allowed the complete sorting of similarity matrices produced by either AUTO-CA or AUTO EF(P) and stored on disc, by UPGMA clustering. To function, the number of electrophoretograms in each matrix must be input at line 510 as N%. The program allowed the printing of the input matrix but this can be removed by inserting REM at lines 40 and 50. This program is original and on cross-checking with an established program used routinely at the National Collection of Industrial and Marine Bacteria, Torry Research Station, Aberdeen was found to be in absolute agreement.

DENDO took approximately 7 minutes to completely sort a 40 x 40 matrix.

5.4.3 AUTO CA

AUTO CA allowed the complete comparison of the data from a 42 lane gel in about 12 hrs. Either of the two calculated coefficients (CFI-Dice,

CF2-Jaccard) could be selected to form a similarity matrix for subsequent analysis by DENDO. Figure 14 shows the UPGMA clustering of the results combined from three 42 lane gels using the Jaccard coefficient. The option of background subtraction incorporated into this program had negligible effect on the resultant comparisons because of the method of multi-level peak search used.

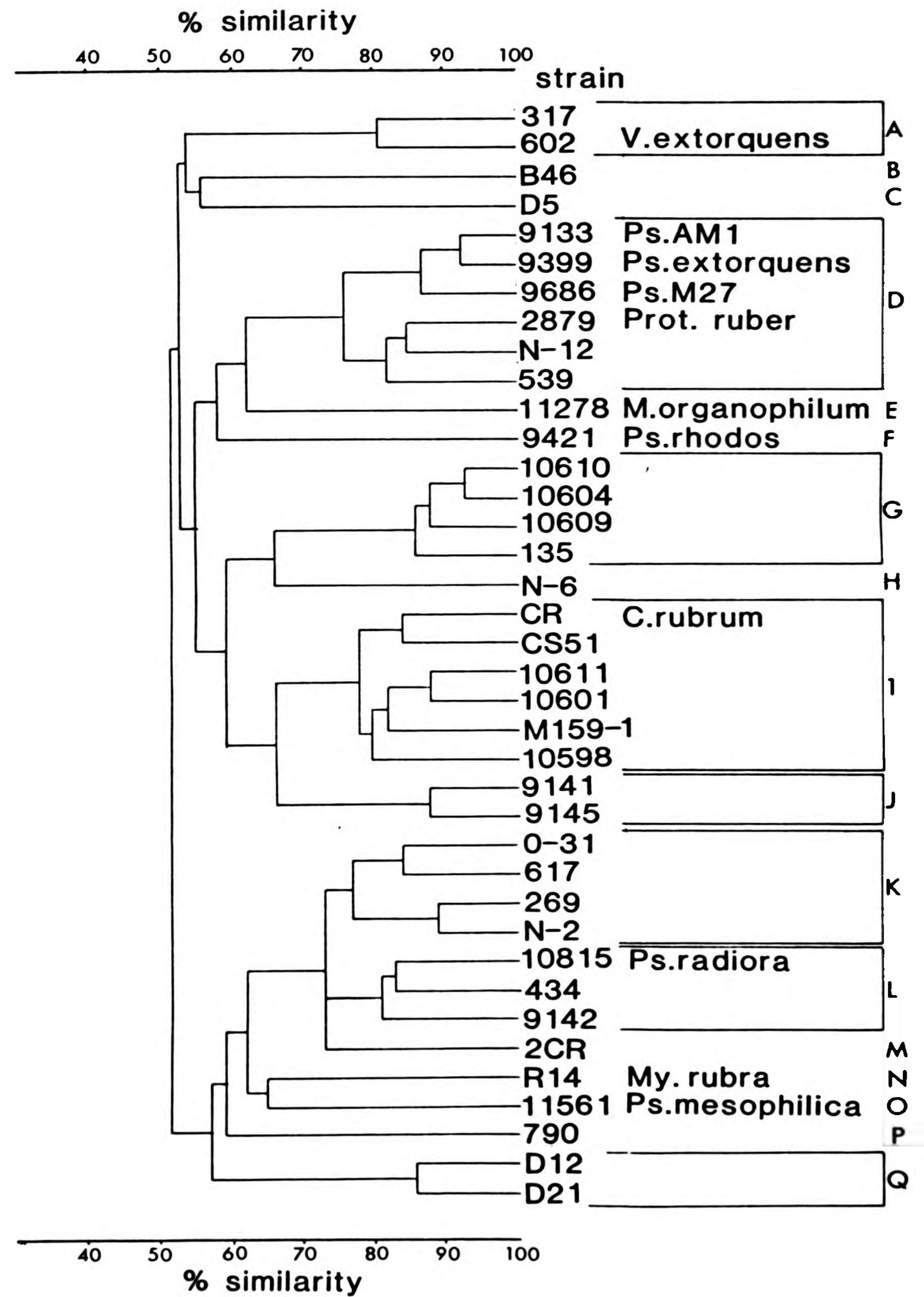
At normal discriminatory levels, any background noise peaks from the microdensitometer, were omitted from calculation by the peak threshold values selected. This negated the need for trace averaging to reduce noise levels as described by Jackman (1983) and was used only to reduce subsequent computing time. AUTO CA's variable margin factor allowed a degree of latitude as regards gel reproducibility but worked best on traces when no more than 60 peaks were selected. Above this number, in areas of high band frequency, the errors from proximity matching of non-homologous bands became unacceptable and cancelled any advantage conferred by the use of the margin factor.

Kersters and De Ley (1975) reported that peak position and distance analysis was unsatisfactory as it was sensitive to experimental differences in the amounts of protein applied to each gel lane. Without doubt this is true when protein concentrations vary by a factor of two or three but the variable search parameters used in this study largely avoided such problems.

The dendrogram in Figure 14 is derived after an averaging of the results from 3 gels, examples of which are given in Figure 11. Results from

Figure 14 Dendrogram showing the clusters obtained by UPGMA linkage after peak search and position analysis with program AUTO CA (Jaccard coefficient) on electrophoresed PPFM soluble protein extracts

Results are an average from three 42 track gels.



individual gels were very similar with differences only in the internal positions of closely related clusters and the positions of a few of the single member strains. Any aberrant results when cross-checked by direct examination of the gel or gel photographs were invariably attributed to atypical gel profiles (a slight bowing or bending of some bands). The choice of coefficient (Dice or Jaccard) had little effect on the overall pattern of the dendrogram except that the Dice coefficient resulted in a greater spread of the similarity values along the x axis of the dendrogram and single member clusters were displaced.

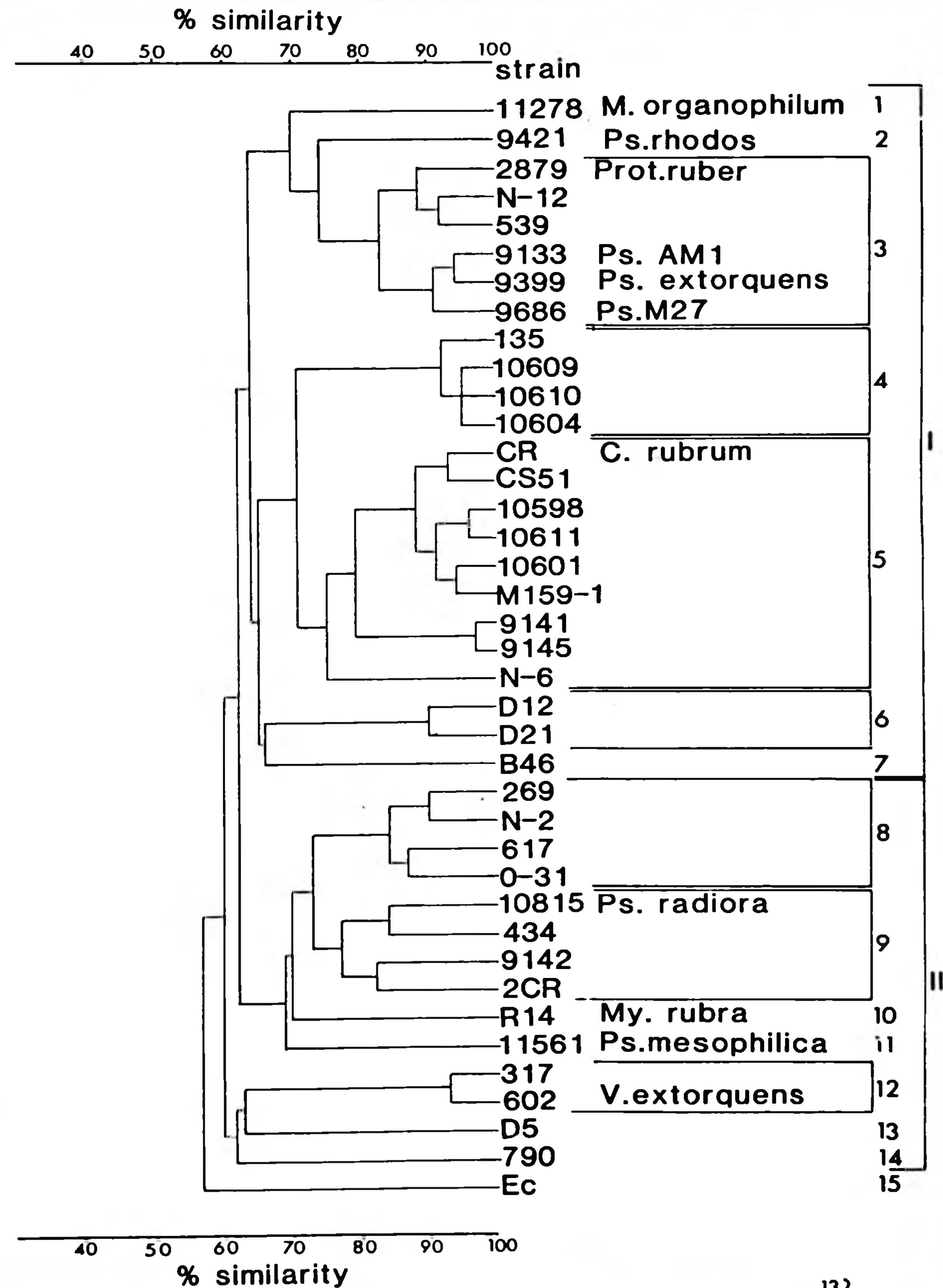
The most significant effect was on any electrophoretogram with an unusually high or low number of detectable bands when compared to the average. Under these circumstances the strain in question could be misplaced when using the Dice coefficient. Kersters and De Ley (1975) have shown that the UPGMA method gives the most satisfactory clustering of electrophoretic data.

### 5.5 Correlation Coefficient Analysis

Figure 15 shows the UPGMA clustering of the results combined from three 42 lane gels analysed by the correlation coefficient method using program AUTO EF. Similar results were obtained using the program AUTO EFP, but with a looser clustering of many of the less related strains (because of the exclusion of standard proteins in the analysis). AUTO EF and AUTO EFP allowed the complete comparison of a 42 lane gel in about 10-11 hrs. The dendrogram in Figure 15 was constructed after low resolution (100 point) analysis of the elctrophoretic traces. This is

Figure 15 Dendrogram showing the clusters obtained by UPGMA linkage after correlation coefficient analysis with program AUTO EF on electrophoresed PPFM soluble protein extracts

Results are an average from three 42 track gels. EC = *Escherichia coli* C.



the minimum resolution reported by Feltham and Sneath (1979) below which comparisons would be affected by trace averaging. Resolution could be increased to 300 or 450 points for analysis. This led only to minor rearrangements within clusters and to the repositioning of some of the less related single member clusters. Correspondingly there was an increase in the spread of similarity values along the x axis indicating an increased discriminatory potential. This is in agreement with Jackman (1983) who reported that 100 and 500 point analyses gave virtually identical results. Background subtraction, facilitated by the Ultroscan program, had no effect on comparisons when straight line subtraction was used. The Pearsons product moment correlation coefficient is such that removal of a uniform background over the entire trace length from one or both profiles in any pairwise comparison has no effect on the value for similarity obtained. Thus the system is fairly insensitive to possible minor fluctuations in the relative amounts of protein loaded between the gel lanes. Complete trend background subtraction used minimal valley values to remove the entire background trend. When used it caused only slight alterations to the resultant clustering. Subtraction increased the spread of linkage values along the x axis, again increasing the potential discrimination between unknowns.

## 5.6 Discussion

As indicated in Figure 15, Section I corresponds with cluster 10 of Green and Bousfield (1982) and Section II to their cluster 11 with the exceptions of strains D5, D12 and D21 not included in their study.

When compared, the results show a high degree of correlation with those obtained by the peak search and position analysis (Figure 14). Choosing an arbitrary level of similarity of 0.75, clusters A, D, G, I, K, L and Q from Figure 14 correspond with clusters 12, 3, 4, 5, 8, 9 and 6 respectively from Figure 15 with few exceptions. Clusters 6 and 12 contain the same strains as Q and A respectively in the same relative order. Clusters 3, 4 and 8 correspond to clusters D, G and K respectively with minor rearrangements of the internal groupings. This is not surprising considering the intrinsic differences in the method of analysis in each case. Cluster 5 corresponds with clusters I and J and cluster 9 to cluster L and strain M. Such fragmentation of clusters with the peak search and position mode of analysis must be due largely to the reduced number of test points on which the calculations are based. Single member clusters 1, 2, 7, 10 and 11 are linked at approximately the same relative positions as E, F, B, N and O respectively. The remaining cluster 6 and single member strains 13 and 14 show linkages different from Q, C and P respectively but their inclusion appears towards the lower end of the discriminatory resolution of the system as shown by the proximity of clustering of an *E. coli* soluble protein profile at a level just below these PPFM single member strains. When included, a protein profile of *Rhodopseudomonas sphaeroides* clustered at a level equal to that of the least associated single member PPFM strains.

Both sets of data agreed well with visual observations made from the gels and gel photographs. No pattern of major bands, common to all strains, could be identified. There were no obvious problems with the

clustering of 'intruders', noted by Kersters and De Ley (1975). Figure 16 shows the clustering of two representative PPFMs strains run on gels alongside representatives of the *Rhodospirillaceae* and analysed using AUTO EF. The linkage values between members of the *Rhodospirillaceae* indicate that their level of inter-relatedness is at the lower end of the resolution of this system and that members exhibit a high degree of heterogeneity.

The novelty of this PAGE-protein analysis system makes it very difficult to relate the results to any firm conclusions as to the taxonomic inter-relationships within the PPFMs. However, it is clear from both sets of data that there exist some tightly clustered groups of strains and that the degree of heterogeneity within the group as a whole is higher than was previously believed (Green and Bousfield, 1983).

The method of correlation coefficient analysis has an advantage over peak search and position analysis in that computation involves data more representative of the complete trace contour. Information on peak size and valley width as well as positions are included in the calculations. Kersters and De Ley (1975) have stated that this method is more universally applicable to a wide range of bacterial strains.

The 'compensation' method of Kersters and De Ley (1975) could not be applied successfully to protein profiles of the PPFMs as it relies on the positions of detectable equivalent protein bands as markers for profile uniformity. Such common bands were not easily found across a range of PPFM protein profiles, this itself being an indication of the

When compared, the results show a high degree of correlation with those obtained by the peak search and position analysis (Figure 14). Choosing an arbitrary level of similarity of 0.75, clusters A, D, G, I, K, L and Q from Figure 14 correspond with clusters 12, 3, 4, 5, 8, 9 and 6 respectively from Figure 15 with few exceptions. Clusters 6 and 12 contain the same strains as Q and A respectively in the same relative order. Clusters 3, 4 and 8 correspond to clusters D, G and K respectively with minor rearrangements of the internal groupings. This is not surprising considering the intrinsic differences in the method of analysis in each case. Cluster 5 corresponds with clusters I and J and cluster 9 to cluster L and strain M. Such fragmentation of clusters with the peak search and position mode of analysis must be due largely to the reduced number of test points on which the calculations are based. Single member clusters 1, 2, 7, 10 and 11 are linked at approximately the same relative positions as E, F, B, N and O respectively. The remaining cluster 6 and single member strains 13 and 14 show linkages different from Q, C and P respectively but their inclusion appears towards the lower end of the discriminatory resolution of the system as shown by the proximity of clustering of an *E. coli* soluble protein profile at a level just below these PPFM single member strains. When included, a protein profile of *Rhodopseudomonas sphaeroides* clustered at a level equal to that of the least associated single member PPFM strains.

Both sets of data agreed well with visual observations made from the gels and gel photographs. No pattern of major bands, common to all strains, could be identified. There were no obvious problems with the

clustering of 'intruders', noted by Kersters and De Ley (1975). Figure 16 shows the clustering of two representative PPFMs strains run on gels alongside representatives of the *Rhodospirillaceae* and analysed using AUTO EF. The linkage values between members of the *Rhodospirillaceae* indicate that their level of inter-relatedness is at the lower end of the resolution of this system and that members exhibit a high degree of heterogeneity.

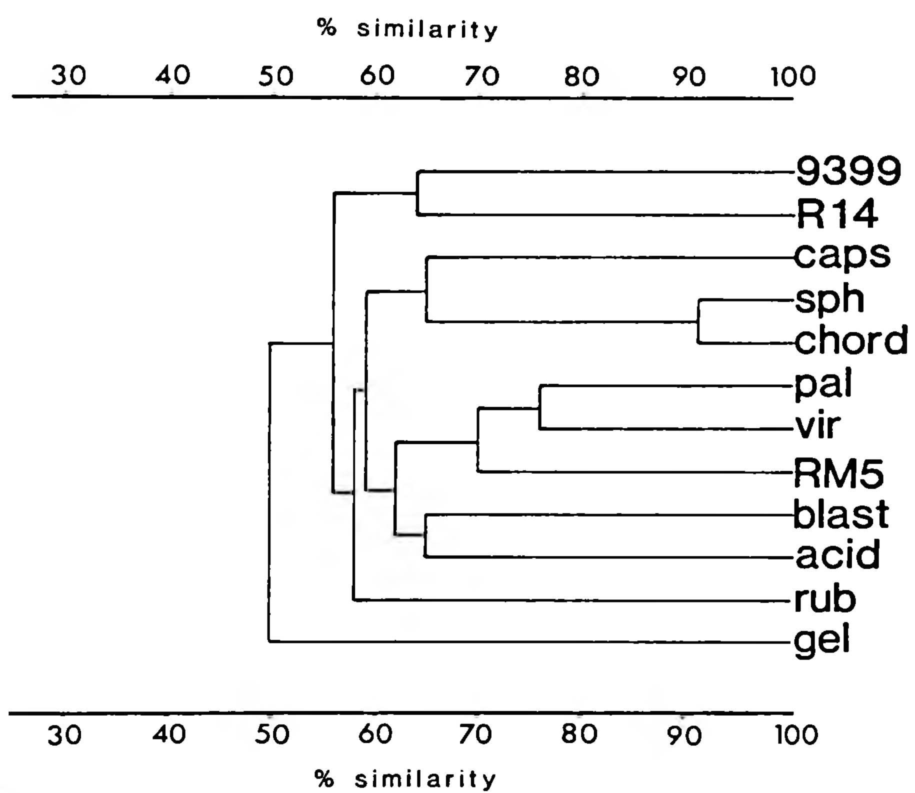
The novelty of this PAGE-protein analysis system makes it very difficult to relate the results to any firm conclusions as to the taxonomic inter-relationships within the PPFMs. However, it is clear from both sets of data that there exist some tightly clustered groups of strains and that the degree of heterogeneity within the group as a whole is higher than was previously believed (Green and Bousfield, 1983).

The method of correlation coefficient analysis has an advantage over peak search and position analysis in that computation involves data more representative of the complete trace contour. Information on peak size and valley width as well as positions are included in the calculations. Kersters and De Ley (1975) have stated that this method is more universally applicable to a wide range of bacterial strains.

The 'compensation' method of Kersters and De Ley (1975) could not be applied successfully to protein profiles of the PPFMs as it relies on the positions of detectable equivalent protein bands as markers for profile uniformity. Such common bands were not easily found across a range of PPFM protein profiles, this itself being an indication of the

Figure 16 **Dendrogram constructed by the UPGMA Method after Correlation Coefficient Analysis on Soluble Protein Extracts from Representatives of the PPFMs and Rhodospirillaceae**

caps - *Rhodopseudomonas capsulata*; pal - *Rhodopseudomonas palustris*; vir - *Rhodopseudomonas viridis*; RM5 - *Rhodomicrobium vannielii*; blast - *Rhodopseudomonas blastica*; acid - *Rhodopseudomonas acidophila*; rub - *Rhodospirillum rubrum*; gel - *Rhodopseudomonas gelatinosa*; 9399 - *Ps. extorquens*; sph - *Rhodopseudomonas sphaeroides*; chord - *Rhodopseudomonas sphaeroides* (chordata).


% similarity
30
40
50
60
70
80
90
100
9399
R14
caps
sph
chord
pal
vir
RM5
blast
acid
rub
gel
30
40
50
60
70
80
90
100
% similarity

heterogeneity within the group. In this study margin factors were the equivalent to their compensation of scans.

Hotchkin and Kaya (1984) have stressed the problems encountered with co-migration of non-homologous bands and suggested that emphasis should be switched to methods such as checking isofunction by staining for specific enzymic activities. However, constant coincidence across the entire electrophoretogram is unlikely.

Many workers have shown that results from PAGE analysis of soluble proteins agreed well with data obtained from DNA:DNA hybridisation studies (Jackman, 1983; Biavati *et al.*, 1982; Kersters and De Ley, 1980; Swings *et al.*, 1976), bio- and sero-typing studies (Fotos *et al.*, 1984) and phenotypic tests (Moore *et al.*, 1980). In general the technique is regarded as having a discriminatory potential equal to that of DNA:DNA hybridisation but with some genera, and a finely tuned analysis system, can go even further. Biavati *et al.* (1982) reported that PAGE analysis allowed discrimination between strains with high and equivalent DNA homology values. PAGE-protein analysis allows determination of the homogeneity of a group of organisms without any real phylogenetic insights.

Much of the early development of the PAGE system was for the rapid identification of large numbers of clinical isolates. Workers have reported being able to use strain specific (Moore *et al.*, 1980), species specific (Kersters and De Ley, 1980) and genus specific (Biavati *et al.*, 1982) banding patterns for the identification of unknown isolates. This

is ideal when it is known that isolates are likely to be of a similar type. PAGE can be routinely used to screen large numbers of strains and identify group representatives to be included in further DNA homology experiments.

Should protein patterns across a range of organisms be very similar and greater discrimination be required, techniques such as 2-dimensional PAGE, iso-functional enzyme studies or peptide mapping could be employed to this end.

## 6. Nucleic Acid Hybridisation Analyses

### 6.1 DNA:DNA Hybridisation Analysis

Table 8 lists the results from the hybridisation of PPFM DNAs and the corresponding melting temperatures of the hybrids.

#### 6.1.1 DNA Purity and Denaturation

Many difficulties were experienced during early DNA:DNA hybridisation experiments with PPFM DNAs. These were largely attributed to problems with cell lysis (see 4.1) and the isolation of DNA of sufficient purity. Monitoring the absorbance at 260 nm ($A_{260}$) of some early DNA preparations indicated an increase in absorbance (hyperchromic shift) of 8-20% on thermal denaturation and filter retention of denatured DNA of only 30-70%. These values were not acceptable for quantitative hybridisation experiments and were distant from the values of 25-35% and 90%+ respectively for pure, high molecular weight DNA.

Improved cell lysis and DNA isolation procedures overcame many of the initial difficulties and values of 23-33% for the hyperchromic shift and 80% plus for filter retention of DNAs were routinely obtained. Such improvements were coordinated with a series of experiments designed to optimise the dissociation and filtration conditions. Repeat DNA dissociation experiments using 20 DNAs over a range of concentrations showed that DNAs in 0.1 x SSC, on average, had a hyperchromic shift 6% greater than when in TE bufer. Although not necessarily indicative of greatly improved DNA dissociation, 0.1 x SSC was the solvent of choice

## Key to Table 8

| | | |
|---|---|---|
| Ps1 | - | *Pseudomonas 1* |
| NCIB 9421 | - | *Ps. radiora* |
| NCIB 2879 | - | *Prot. ruber* |
| NCIB 9133 | - | *Ps. AM1* |
| NCIB 9399 | - | *Ps. extorquens* |
| NCIB 11278 | - | *M. organophilum* |
| NCIB 10815 | - | *Ps. radiora* |
| NCIB 11561 | - | *Ps. mesophilica* |
| R14 | - | *Mycoplana rubra* |
| CR | - | *Corynebacterium rubrum* |

Table 8 DNA hybridisation and $-\Delta Tm_{(e)}$ results

| Filter DNA Source | Radiolabelled DNA Source | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 135 | | NCIB 10609 | | NCIB 10611 | | NCIB 10598 | |
| Strain | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) |
| 135 | 100 | | 101 | 0.2 | 39 | 6.2 | 44 | 4.2 |
| NCIB 10610 | 98 | | 103 | 0 | 34 | | 34 | |
| NCIB 10609 | 102 | | 100 | 0 | 38 | | | |
| M159-1 | 37 | | 37 | | 98 | 0 | 92 | 0.1 |
| NCIB 10611 | 36 | | 33 | 7.3 | 100 | 0 | 93 | 0.1 |
| NCIB 10598 | | | | | 97 | 0 | 100 | 0 |
| CS51 | | | | | 89 | 0.4 | 91 | 1.1 |
| Ps1 | | | | | | | | |
| 385B | | | | | | | | |
| NCIB 9141 | 29 | | 30 | | 35 | 4.3 | 43 | 5.1 |
| N-6 | 13 | | 28 | 7.0 | 45 | | 39 | |
| 317 | | | | | | | | |
| NCIB 9421 | 20 | | 19 | 8.3 | 32 | 8.4 | 25 | |
| NCIB 2879 | 39 | | 37 | 6.6 | 50 | 5.5 | 43 | 5.6 |
| 539 | 36 | | 39 | | 43 | | | |
| NCIB 9133 | | | | | | | | |
| NCIB 9399 | 33 | | 35 | | | | | |
| N-12 | | | | | | | | |
| NCIB 11278 | 26 | | 27 | 7.9 | 29 | 7.4 | 28 | 7.4 |
| D12 | 10 | | | | | | | |
| B46 | 10 | | | | 14 | | 12 | 9.5 |
| 434 | | | 15 | | | | 13 | |
| NCIB 9142 | 14 | | 14 | | 16 | | 17 | |
| NCIB 10815 | 15 | | | | 13 | | 13 | |
| 617 | 10 | | 12 | 9.8 | 20 | 9.3 | 13 | 9.7 |
| 0-31 | 14 | | 13 | 10.2 | 19 | | 15 | 10.0 |
| 269 | 14 | | | | 14 | 9.8 | | |
| 45 | | | | | | | | |
| NCIB 11561 | 10 | | 11 | 9.5 | | | 12.3 | 9.8 |
| 790 | 11 | | 15 | 10.0 | 19 | 10.5 | | |
| N-2 | | | | | | | | |
| R14 | | | 117 | 7.8 | | | 17.1 | |
| 2CR | | | | | | | | |
| CR | | | | | | | | |

(contd/.)

Table 8 (contd.)

| Filter DNA Source | Radiolabelled DNA Source | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | NCIB 11278 | | NCIB 2879 | | NCIB 9399 | | NCIB 9133 | |
| Strain | RPB | -ΔTm(e) (°C) | RPB | -ΔTm(e) (°C) | RPB | -ΔTm(e) (°C) | RPB | -ΔTm(e) (°C) |
| 135 | 27 | 7.8 | 38 | 6.5 | | | | |
| NCIB 10610 | 28 | | 37 | 6.5 | 34 | | | |
| NCIB 10609 | 26 | | 39 | | 38 | 6.5 | 38 | 5.9 |
| M159-1 | 25 | 8.9 | 47 | | | | | |
| NCIB 10611 | 26 | | 45 | 5.2 | 44 | 4.9 | 44 | 4.7 |
| NCIB 10598 | 24 | | | | | | 52 | |
| CS51 | 22 | | | | | | | |
| Ps1 | | | | | | | | |
| 385B | | | | | 37 | | 37 | |
| NCIB 9141 | 21 | | 35 | | | | | |
| N-6 | 17 | | 39 | | | | | |
| 317 | 16 | 9.9 | | | | | | |
| NCIB 9421 | 13 | | 22 | 7.6 | 20 | 7.9 | 19 | 6.8 |
| NCIB 2879 | 24 | | 100 | 0 | 81 | 1.9 | 80 | 1.3 |
| 539 | 20 | | 83 | 1.6 | 78 | 2.0 | 78 | 1.4 |
| NCIB 9133 | 25 | 8.2 | | | | | 100 | 0 |
| NCIB 9399 | 21 | 9.5 | 82 | 1.2 | 100 | 0 | 87 | 0.5 |
| N-12 | | | | | 84 | 1.5 | 80 | 0.8 |
| NCIB 11278 | 100 | 0 | 17 | 9.1 | 18 | 9.2 | 18 | 7.4 |
| D12 | 11 | | 8 | | | | | |
| B46 | 9 | 11.3 | | | | | 11 | |
| 434 | 18 | | 16 | | | | | |
| NCIB 9142 | 17 | | 11 | | | | | |
| NCIB 10815 | 13 | | 12 | 10.0 | | | | |
| 617 | 13 | | 13 | 9.7 | | | | |
| 0-31 | 16 | 10.0 | 16 | 9.9 | 18 | 10.9 | | |
| 269 | 11 | 10.2 | | | 10 | 10.7 | 10 | 9.6 |
| 45 | | | | | | | | |
| NCIB 11561 | 12 | | 9 | 10.0 | | | 9 | 9.2 |
| 790 | 14 | | 13 | | | | | |
| N-2 | | | | | | | | |
| R14 | 10 | | 14 | 9.2 | | | | |
| 2CR | 12 | 10.5 | | | | | | |
| CR | 19 | 8.9 | | | | | | |

(contd/.)

Table 8 (contd.)

| Filter DNA Source | Radiolabelled DNA Source | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Ps1 | | 385B | | NCIB 9141 | | NCIB 9421 | |
| Strain | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) |
| 135 | | | 30 | 7.2 | | | 22 | 6.9 |
| NCIB 10610 | 36 | | 30 | 7.5 | 34 | | 25 | |
| NCIB 10609 | 34 | 6.8 | | | 33 | 6.5 | | |
| M159-1 | 90 | 0.6 | 48 | 4.7 | 47 | 5.6 | 26 | |
| NCIB 10611 | 86 | 0.9 | | | 52 | 5.9 | 27 | 6.9 |
| NCIB 10598 | 94 | 0.7 | 56 | 5.0 | 53 | 4.9 | | |
| CS51 | | | | | 55 | 4.6 | | |
| Ps1 | 100 | 0 | | | | | | |
| 385B | 48 | 5.6 | 100 | 0 | | | | |
| NCIB 9141 | 40 | 5.7 | 42 | 5.5 | 100 | 0 | | |
| N-6 | | | | | 41 | 5.2 | 25 | |
| 317 | | | | | | | | |
| NCIB 9421 | 24 | 7.4 | | | 26 | 8.1 | 100 | 0 |
| NCIB 2879 | 34 | | 32 | 5.6 | 45 | 5.9 | 20 | 7.3 |
| 539 | | | | | 36 | | | |
| NCIB 9133 | 33 | | | | | | | |
| NCIB 9399 | 40 | 5.8 | 36 | 5.8 | 36 | | 19 | |
| N-12 | | | | | | | | |
| NCIB 11278 | 20 | 7.5 | 17 | 8.4 | 25 | 8.1 | 15 | 9.2 |
| D12 | | | | | | | 12 | 9.1 |
| B46 | | | | | 12 | | 9 | 9.1 |
| 434 | | | | | | | | |
| NCIB 9142 | 11 | 9.4 | 12 | 9.7 | 16 | 10.5 | 19 | |
| NCIB 10815 | | | | | | | 18 | 8.6 |
| 617 | | | | | 16 | | 14 | 8.0 |
| 0-31 | | | | | 20 | 11.4 | 17 | |
| 269 | 11 | | | | 10 | 11.5 | | |
| 45 | 14 | | | | | | | |
| NCIB 11561 | 10 | | 14 | 10.2 | | | 13 | 8.7 |
| 790 | | | | | | | 17 | 10.0 |
| N-2 | | | | | | | | |
| R14 | | | | | | | 14 | 8.5 |
| 2CR | | | | | | | | |
| CR | | | | | | | | |

(contd/.)

Table 8 (contd.)

| Filter DNA Source | Radiolabelled DNA Source | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | NCIB 11561 | | NCIB 9142 | | NCIB 10815 | | 617 | |
| Strain | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) | RPB | $-\Delta Tm(e)$ ($^{o}C$) |
| 135 | 9 | 9.1 | | | | | | |
| NCIB 10610 | 8 | 9.2 | 16 | | 12 | | | |
| NCIB 10609 | | | 16 | 9.9 | 13 | 10.5 | 16 | |
| M159-1 | 14 | | 20 | 10.3 | 14 | 10.9 | | |
| NCIB 10611 | 13 | 9.2 | | | | | 18 | 10.6 |
| NCIB 10598 | | | | | | | | |
| CS51 | | | | | | | | |
| Ps1 | | | | | | | | |
| 385B | | | | | | | | |
| NCIB 9141 | | | 14 | | 15 | | | |
| N-6 | 9 | | 9 | | 8 | | | |
| 317 | 10 | | 18 | | 13 | | | |
| NCIB 9421 | 11 | | 24 | | 23 | 10.3 | | |
| NCIB 2819 | 11 | 9.3 | 14 | 10.5 | 14 | 11.2 | | |
| 539 | | | 16 | | 17 | | | |
| NCIB 9133 | | | | | | | | |
| NCIB 9399 | 10 | | | | | | 16 | |
| N-12 | | | | | | | | |
| NCIB 11278 | 10 | | 17 | 9.9 | 16 | 10.1 | 14 | 9.7 |
| D12 | 10 | 8.8 | | | 17 | 10.3 | 14 | |
| B46 | 8 | | 12 | 10.9 | 8 | | 12 | 11.0 |
| 434 | 38 | 6.6 | 68 | 3.2 | 58 | 3.5 | 49 | 4.8 |
| NCIB 9142 | 32 | 6.9 | 100 | 0 | 81 | 0 | 58 | 4.6 |
| NCIB 10815 | 29 | 6.7 | 90 | 0 | 100 | 0 | 52 | 4.8 |
| 617 | 31 | 6.9 | 53 | 5.3 | 56 | 4.9 | 100 | 0 |
| 0-31 | 27 | | 62 | 5.2 | 53 | 5.5 | 95 | 0 |
| 269 | | | 39 | 5.7 | 44 | 5.3 | 88 | 0.4 |
| 45 | | | | | | | | |
| NCIB 11561 | 100 | 0 | 37 | 8.0 | 28 | 8.6 | 37 | 8.6 |
| 790 | 18 | 8.7 | 24 | 10.0 | 24 | 9.1 | 32 | 10.7 |
| N-2 | | | 29 | 7.9 | 27 | 8.1 | 26 | 7.9 |
| R14 | 13 | | 16 | 9.9 | 13 | 10.2 | 16 | 11.3 |
| 2CR | | | 49 | 4.0 | 53 | 4.4 | | |
| CR | | | | | | | 13 | |

(contd/.)

Table 8 (contd.)

| Filter DNA Source | Radiolabelled DNA Source | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 45 | | D12 | | 317 | | CR | |
| Strain | RPB | $-\Delta T_m(e)$ ($^oC$) | RPB | $-\Delta T_m(e)$ ($^oC$) | RPB | $-\Delta T_m(e)$ ($^oC$) | RPB | $-\Delta T_m(e)$ ($^oC$) |
| 135 | 22 | 9.1 | | | 23 | 8.9 | 47 | 5.6 |
| NCIB 10610 | | | | | | | | |
| NCIB 10609 | | | 16 | 10.7 | | | | |
| M159-1 | | | | | 19 | | 101 | 0.3 |
| NCIB 10611 | | | 17 | 10.5 | 21 | 8.9 | 88 | 0.6 |
| NCIB 10598 | | | | | | | | |
| CS51 | | | | | | | | |
| Ps1 | | | | | | | | |
| 385B | | | | | | | 55 | 4.3 |
| NCIB 9141 | | | | | | | 47 | 4.4 |
| N-6 | | | | | | | | |
| 317 | 16 | 9.3 | | | 100 | 0 | | |
| NCIB 9421 | | | | | | | | |
| NCIB 2819 | 18 | 9.1 | | | | | | |
| 539 | | | | | | | | |
| NCIB 9133 | | | | | 16 | 9.7 | | |
| NCIB 9399 | | | 15 | | 17 | | 45 | 4.7 |
| N-12 | | | | | | | | |
| NCIB 11278 | 18 | 9.1 | 13 | 10.6 | 20 | 8.8 | 28 | 8.0 |
| D12 | | | 100 | 0 | | | | |
| B46 | | | 18 | 9.3 | | | | |
| 434 | 41 | 5.1 | 12 | | 17 | 9.2 | | |
| NCIB 9142 | 49 | 4.8 | 17 | 10.3 | 19 | 9.2 | | |
| NCIB 10815 | | | 16 | 10.3 | | | | |
| 617 | 89 | 0.6 | 13 | 10.3 | | | | |
| 0-31 | 94 | 0.7 | 14 | 10.9 | 17 | 9.9 | 17 | 8.1 |
| 269 | 89 | 0.6 | 10 | 10.3 | 16 | 10.1 | | |
| 45 | 100 | 0 | | | | | | |
| NCIB 11561 | 33 | 7.9 | 14 | 10.4 | 10 | 9.9 | 17 | 8.8 |
| 790 | 26 | 9.1 | 18 | 10.7 | 15 | 10.3 | | |
| N-2 | 31 | 7.0 | 9 | | | | | |
| R14 | | | 11 | | | | | |
| 2CR | | | | | | | | |
| CR | | | 8 | | | | 100 | 0 |

(contd/.)

Table 8 (contd.)

| Filter DNA Source | Radiolabelled DNA Source | | | | | |
|---|---|---|---|---|---|---|
| | 2CR | | 602 | | NCIB 9145 | |
| Strain | RPB | $-\Delta Tm(e)$ ($^{o}$C) | RPB | $-\Delta Tm(e)$ ($^{o}$C) | RPB | $-\Delta Tm(e)$ ($^{o}$C) |
| 135 | | | | | 36 | |
| NCIB 10611 | 16 | 8.8 | 22 | | 49 | |
| NCIB 9141 | | | | | 88 | |
| 317 | | | 91 | 0.9 | | |
| 539 | 19 | | 25 | 9.7 | | |
| B46 | 11 | 9.6 | 11 | 11.2 | | |
| 434 | 63 | 3.1 | 19 | 7.9 | | |
| 0-31 | 57 | 4.4 | 20 | 10.0 | | |
| NCIB 11561 | 28 | 6.7 | 17 | 8.7 | | |
| 790 | 17 | | 22 | 10.0 | | |
| 2CR | 100 | 0 | | | 11 | |
| 602 | 15 | 8.8 | 100 | 0 | 17 | |
| NCIB 9145 | | | | | 100 | |

| Filter DNA Source | Radiolabelled DNA Source | |
|---|---|---|
| | NCIB 9399 | |
| Strain | RPB | $-\Delta Tm(e)$ ($^{o}$C) |
| NCIB 9399 | 100 | 0 |
| *E. coli* K12 | 0.4 | 10.9 |
| *Salmonella typhimurium* LT2 | 0.5 | |
| *Rhodopseudomonas blastica* | 1.2 | |
| *Methylococcus capsulatus* (Bath) | 0.9 | 11.5 |
| *Methylomonas trichosporium* | 2.0 | 11.5 |
| *Pseudomonas mesophilica* | 0.3 | 11.1 |
| *Methylomonas albus* BG8 | 0.4 | 11.8 |

RPB - Relative % DNA binding

for the DNAs in all subsequent hybridisation analyses. Monitoring of the hyperchromic shift under controlled conditions was in itself a good indication of DNA purity (De Ley, 1970c).

Further experiments showed that denaturation of DNA by boiling at 100°C, and by the addition of NaOH to 0.2 N (Baker, 1977), were equally effective at dissociating PPFM DNA, as measured by filter retention. Alkali denaturation is often recommended for DNAs with a high % (G+C) content (Johnson, 1981). This can involve simply the addition of alkali and filtration of the denatured DNA at a high pH (Baker, 1977) or can include pH neutralisation prior to filtration (Gillespie and Spiegelman, 1965; Stackebrandt and Kandler, 1979). The high pH filtration method has the advantage in that it can prevent renaturation or fold back of the DNA, is more rapid and is insensitive to the salt concentration. However, it was found during preliminary hybridisation experiments that a high pH (13-14) could cause yellowing of the nitrocellulose and allowed DNAs to spread beyond the area of application due to some partial breakdown of the filter matrix. This is undesirable, especially when using the multi-blot hybridisation system. Attempts to fix alkali denatured DNAs to the filter by infra-red irradiation compounded this problem making filters very brittle on subsequent handling. As heat denaturation is simple for multiple samples, and allowed a level of binding of DNAs to the filter equal to that with the alkali - neutralisation procedure, it became the method of choice in subsequent experiments.

The melting temperature of high % (G+C) DNAs often approaches that of boiling water and has forced some workers to use either a glycerol or oil bath (De Ley and Tijtgat, 1970) or a water bath with added salt to elevate the incubation temperature. Although the melting temperature of the PPFM DNAs fell within the range of 97-99°C in this study there was little difference in the filter retention of DNAs heated to 100°C or 106°C.

### 6.1.2 DNA Filter Retention

The proportion of DNA retained by a filter depends on factors other than the purity of the DNA. Listed in Tables 9 and 10 are some parameters investigated to maximise the binding of DNA to filters. Table 9 shows results for the average amounts of denatured PPFM and *E. coli* DNAs bound to a variety of filter types. The values are comparable with those reported previously (De Ley, 1970a; De Ley and Tijtgat, 1970). Schleicher and Schuell BA85 followed by Millipore GSWP proved to be the most efficient nitrocellulose types. Millipore HAHY filters gave the poorest retention of denatured DNA. At a slow rate of filtration (approximately 0.5 ml $min^{-1}$) no significant differences were noted in the retention values over the range of DNA concentrations (15-85 $\mu g$ $ml^{-1}$) used.

The average fragment size of filtered chromosomal DNA, as shown by agarose gel electrophoresis, was in the range of 40-85 kb ($2.6-5.6 \times 10^7$ mol wt), compared with the minimum size required for efficient filtration and binding of $5 \times 10^6-10^7$ mol wt (De Ley and Tijtgat, 1970). Sonicated DNA (1-2 kb) was retained at a lower level. DNAs degraded

during the isolation and purification procedures, although not always showing greatly reduced filter retention levels, were released preferentially under hybridisation conditions. Table 9 lists the relative amounts of DNA released from the filters after 20 hrs incubation at 55°C. Degraded DNA and DNA bound to Millipore HAHY filters showed the greatest release. Most of this occurred within the first 20 min of incubation at the hybridisation temperature. Values were comparable with those reported by Legault Demare *et al.* (1967). De Ley and Tijtgat (1970) have shown that at incubation temperatures above 70°C, 40-85% of filter bound DNA is released and above 80°C virtually all DNA is released. Thus high temperatures should be avoided. Under defined conditions the amount of filter DNA released was reproducible and no separate preincubation step was necessary to eliminate the released DNA and obtain meaningful results.

Table 10 shows the effect of altering the filtration conditions on the retention of strain NCIB 10609 DNA. The results show an increased retention of DNA with increasing concentration of salts in the solvent, in agreement with Phillips (1969) and Baker (1977). 6 x SSC was the solvent of choice for the filtration and fixation of denatured DNA on filters. No significant difference was noted between filtration at 4°C or room temperature.

DNA prepared from strain 11561 by the two methods as described previously (3.2.1, 3.2.2) was denatured and filtered through Millipore GSWP filters. The % retentions differed by only 2% indicating no significant variation with the DNA preparation method. These DNAs also

Table 9 Retention of Denatured DNA by Nitrocellulose and Nylon Membrane Filters

| Filter Type | Average Retention of denatured DNA (%) | % of Bound DNA Released under Hybridisation Conditions[b] |
|---|---|---|
| Millipore GSWP* | 88[a] | 16 |
| Millipore GSWP* sonicated DNA | 67 | 38 |
| Millipore HAWP* | 85 | - |
| Millipore HAHY* | 76 | 23 |
| Schleicher & Schuell BA85* | 91 | 12 |
| Paul Biodyne Nylon | 94 | - |

* 0.45 μm pore size nitrocellulose

a $^{32}$P NCIB 9399, 135 and *E. coli* C DNAs filtered over range of 12-75 μg ml$^{-1}$ in 6 x SSC, 4$^{\circ}$C. DNAs denatured by heating at 104$^{\circ}$C for 10 mins.

b Hybridisation conditions of 3 x SSC, 1 x Denhardts solution, 30% formamide for 20 hrs at 55$^{\circ}$C.

Table 10 Retention of Denatured DNA by Schleicher and Schuell BA85 Nitrocellulose Filters

| Solvent | % DNA Retained |
|---|---|
| 0.1 x SSC | 37[c] |
| 3 x SSC | 73 |
| 6 x SSC | 88 |
| 6 x SSC, 4°C | 86 |
| 10 x SSC | 89 |

[c] Results averages of 4-6 trials with $^{32}P$ labelled NCIB 10609 DNA heated at 104°C for 10 mins.

gave equivalent levels of homologous DNA binding in subsequent hybridisation experiments

Both methods of DNA hybridisation in this study used levels of filter bound DNA just below those for DNA saturation of the filters (32 $\mu g$ $cm^{-2}$, De Ley and Tijtgat, 1970).

6.1.3 Radiolabelled Probes

Preliminary hybridisation experiments used $^{3}H$-DNA probes prepared by *in vivo* labelling methods but even those with the highest specific activities (>5 x $10^{4}$ cpm $\mu g^{-1}$) resulted in low levels of filter bound counts under predicted optimal hybridisation conditions. The proximity of the values to that for background radiation and control filters made interpretation of the results difficult and reproducibility impossible. DNA probes of higher specific activity (>$10^{6}$ cpm $\mu g^{-1}$), obtained by *in vitro* radiolabelling, improved the situation and gave maximum values of 12-13% binding of the probe to the homologous DNA, well below those of 30-50% reported in the literature (Stackebrandt and Kandler, 1979).

Detailed monitoring of the hybridisation reactions indicated that up to 60% of the input counts were suppressed on binding to the filter and by quenching from the reaction mixture. Experiments using both $^{3}H$- and $^{32}P$-labelled DNA probes simultaneously revealed the significantly lower efficiency of counting of the $^{3}H$ label. Scintillation counting may have been improved by dissolution of the filters with Brays mixture (De Ley and Tijtgat, 1970) but such treatment precludes the melting of the filter bound hybrids. Under identical conditions, the level of

reassociation in the homologous system was increased to between 30-45% of the input probe DNA with $^{32}P$ rather than $^{3}H$ labelled DNA. *In vitro* labelling of DNA with $^{32}P$ also provided probes with a higher specific activity (>$10^7$ cpm $\mu g^{-1}$) and enabled results to be recorded visually by autoradiography. $^{32}P$ DNA also made the results statistically more meaningful, and no problems were experienced with quenching of the counts from filter bound DNA or free DNA in solution (Table 11).

6.1.4 Hybridisation Conditions

Hybridisation conditions were optimised by experimentally varying the parameters of the reaction mixture. The optimal temperature for renaturation (Tor) of PPFM DNAs should be between 72-74°C in 1 x SSC, assuming that renaturation is optimal at 25°C below the melting temperature. This is corrected to an average of 78°C for the higher salt concentration (3 x SSC) present in the reaction mixture which effectively lowers the temperature of incubation (Gillis *et al.*, 1970). Increased ionic strength is known to increase both the rate of reassociation and the stability of the hybrids so formed (Brenner *et al.*, 1969). Homologous hybridisation reactions between strain 135 filter bound and $^{32}P$ labelled probe DNAs incubated for 20 hrs in 3 x SSC at Tor, Tor + 10°C and Tor - 10°C gave levels of binding of 31%, 22% and 42% respectively. Although the lower temperature results in a higher % binding of the homologous DNA, it is said to be of low stringency and the hybrids formed are less stable and not completely matched over their entire lengths. In the heterologous system this results in high degrees of homology which are false (Kingsbury *et al.*, 1969). Thus, Tor was the tempemture of choice for optimal renaturation conditions.

Table 11 **Quenching of Radiolabelled DNA by the Hybridisation Mixture**

| Hybridisation mixture | % Counts Quenched | |
|---|---|---|
| | $^{3}H$ labelled DNA | $^{32}P$ labelled DNA |
| 0.1 x SSC | - | - |
| 3 x SSC | 6.3[a] | 0[a] |
| 3 x SSC, 1 x Denhardts | 10.6 | 0 |
| 2 x SSC, 1 x Denhardts, 20% formamide | 16.6 | 0 |
| 2 x SSC, 30% DMSO | 28.5 | 1.0 |

[a] Measured against labelled DNA in 0.1 x SSC. Counted in Beckman EP scintillation fluid. Results are averages from triplicate samples of two DNAs.

A problem with the PPFMs is that at temperatures above 70°C there is an unacceptable release of the filter bound DNA and chain scission and depurination can also occur (McConaughy *et al.*, 1969). Organic solvents can be included in the reaction mixture to allow incubation at a suitably depressed temperature. When the organic solvent DMSO was used, reactions were carried out at the optimal temperature (Tor.d) for renaturation as described by De Ley and Tijtgat (1970). The efficiency of homologous reannealing was compared between the DMSO system above, the formamide system of McConaughy *et al.* (1969) and renaturation in 3 x SSC alone, each at their optimal renaturation temperature. Table 12 lists the results. The DMSO system allowed the highest levels of homologous reassociation, in agreement with the findings of De Ley and Tijtgat (1970) in that it provided the most suitable medium for hybridisation of DNAs with a high % (G+C) content. The low level of background counts was comparable with that found with Denhardts solution (Johnson and Ordal, 1968) and indicated that input DNA could only renature or hybridise with the reference DNA. No difference in the levels of hybridisation were found after extended incubation in DMSO for 36 or 46 hrs.

Table 13 lists the differing values for homology found when filter bound DNAs were hybridised against $^{32}P$ radiolabelled DNA from strain NCIB 10598 in DMSO at 60°C (Tor.d) and 53°C (Tor.d - 7°C). In all cases homology values were higher at the less stringent tempenture. Representative filters from each temperature, when heated at 75°C for 10 min in 2 x SSC, 25% formamide, invariably showed a greater % release of filter bound counts after hybridisation at 53°C than at 60°C. This

Table 12 Homologous Reassociation of NCIB 10609 DNA under Varying Hybridisation Parameters

| Hybridisation Mixture | $^{3}H$ 10609 Probe DNA | $^{32}P$ 10609 Probe DNA | |
|---|---|---|---|
| | % Binding | % Binding | % Background |
| 2 x SSC, 30% DMSO, 60°C | 19.3[b] | 45.1[b] | 0.3[c] |
| 2 x SSC, 1 x Denhardts, 20% Formamide, 60°C | 11.4 | 26.8 | 0.3 |
| 3 x SSC, 78°C | 8.5 | 19.2 | 4.9 |

[b] Results are averages of 5 separate reactions. 5 µg of filter bound DNA incubated with 0.05 µg of probe DNA for 20 hrs. Schleicher and Schuell BA85 filters.

[c] Values are for the % of input probe DNA bound to surface of blank control filters.

Table 13 Effect of Temperature on the Hybridisation of PPFM DNAs with the DMSO System

| Source of Unlabelled Filter Bound DNA | Relative % Binding of $^{32}P$ Labelled NCIB 10598 DNA 60°C | 53°C |
|---|---|---|
| 10598 | 100[a] | 100[a] |
| 10611 | 94 | 99 |
| 2879 (*Prot. ruber*) | 40 | 57 |
| 11278 (*M. organophilum*) | 28 | 43 |
| 135 | 37 | 49 |
| 10610 | 33 | 50 |
| 9141 | 41 | 61 |
| 434 | 16 | 26 |
| 10815 (*Ps. radiora*) | 14 | 26 |
| 11561 (*Ps. mesophilica*) | 12 | 19 |
| 9421 (*Ps. rhodos*) | 25 | 39 |
| *E. coli* C | 0.5 | 1.1 |

[a] Average of duplicate hybridisations by the multi-blot method in 2 x SSC, 30% DMSO.

supports previous findings that hybrids formed at temperatures below that optimal for renaturation are generally less stable (De Ley and Tijtgat, 1970).

An effective calibration of the DMSO system was performed by carrying out heterologous reassociation reactions between strains with established DNA homology values. DNAs were prepared from members of the *Enterobacteriaceae* and radiolabelled *E. coli* and *S. typhimurium* LT2 DNAs used as probes against themselves and the other filter bound DNAs. The results in Table 14 were in good agreement with those reported previously (Brenner and Cowie, 1968; Brenner *et al.*, 1969), indicating that the hybridisation system is producing results comparable with those found by other workers.

Experiments with the multi-blot system, substituting an equivalent quantity of unrelated herring sperm DNA as carrier for the probe, resulted in about 20% less hybrid formation averaged over the complete filter, but the relative % bindings remained unchanged. Under both conditions, the $Tm_{(e)}$ values were identical.

The only previously published DNA homology values for the PPFMs are those reported by Hanson (1980) using radiolabelled DNA from *M. organophilum* strain XX NCIB 11278 hybridised againt DNAs isolated from a variety of methylotrophs. Strain XX DNA showed 45-48% homology with that from *Pseudomonas* AM1 NCIB 9133. This is almost double the value of 25% found in this study (Table 8) and even greater than the value for the reciprocal reaction. From information available the hybridisation

Table 14 DNA Reassociation and Thermal Stability of Hybrids Formed Amongst Members of the *Enterobacteriaceae*

| Source of Unlabelled Filter Bound DNA | Source of Labelled DNA: $^{32}$P *S. typhimurium* LT2 — Relative % Binding | $^{32}$P *S. typhimurium* LT2 — $-\Delta Tm_{(e)}$ °C | $^{32}$P *E. coli* C — Relative % Binding | $^{32}$P *E. coli* C — $-\Delta Tm_{(e)}$ °C |
|---|---|---|---|---|
| *Salmonella typhimurium* LT2 | 100 | 0 | 22.8 | 9.7 |
| *Serratia marcescens* | 11 | 10.5 | 7 | 12.1 |
| *Escherichia coli* C | 29 | 8.8 | 100 | 0 |
| Strain 135 (PPFM) | 0.6 | - | 0.6 | 13.6 |

experiments of Hanson were carried out at 65°C in 0.14 M phosphate buffer, this being equivalent to approximately 1 x SSC (Kingsbury *et al.*, 1969). Under these conditons, the optimal reassociation temperature should be between 72-74°C and thus the hybridisation reactions of Hanson were carried out at 7-9°C below Tor. From Table 13 it can be seen that similar disparate values were found in this study at the sub-optimal temperature of 53°C (7°C below Tor).

The DNA homology values in Table 8 are those obtained after DNA hybridisation studies by the multi-blot method. The average standard error of the means for repeat hybridisations was 2.6, equivalent to an average variation of 6.4% about the mean values. This compares favourably with values reported previously (De Ley *et al.*, 1973; Christiansen *et al.*, 1981). Hybridisations by the individual incubation method gave equivalent homology values but with some 2-4% lower than those above. Repeat individual incubations had an average standard error of the means of 2.1, equivalent to an average variation of 5.1% about the mean values.

All PPFM DNA-DNA inter-relatedness is greater than the homology values obtained against DNAs from reference strains included in this study (Table 8).

6.1.5 Comparison of Results

Figure 17 shows a dendrogram constructed from the available multi-blot DNA:DNA hybridisation data by incomplete UPGMA linkage. It must be stressed that this dendrogram is not necessarily a completely accurate

Figure 17 Dendrogram constructed by incomplete group average linkage of the homology data available for DNAs isolated from the PPFMs and hybridised by the multi-blot method

– compared with Fig 15 overleaf

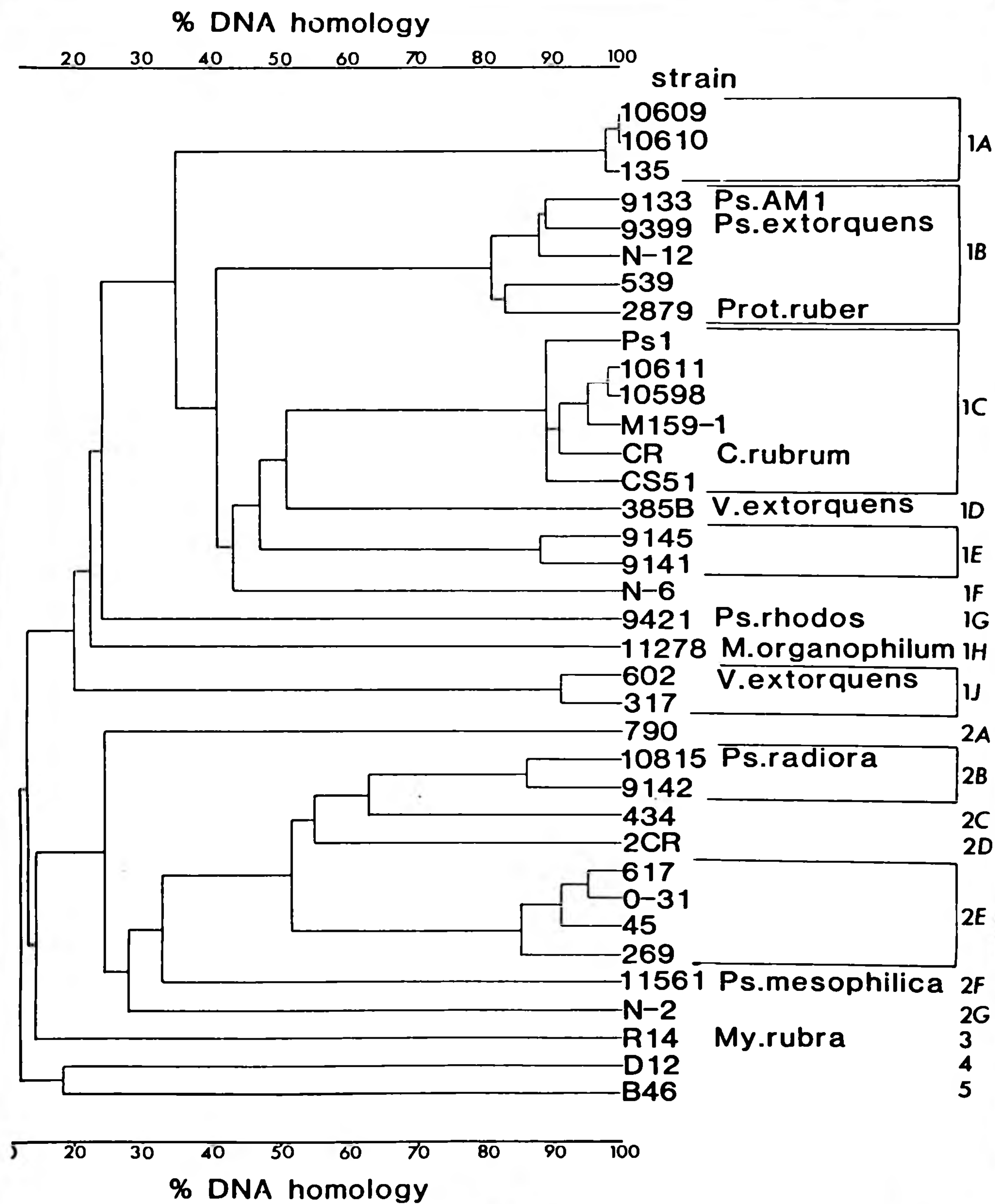

% DNA homology
20
30
40
50
60
70
80
90
100
strain
10609
10610
135
1A
9133 Ps.AM1
9399 Ps.extorquens
N-12
539
2879 Prot.ruber
1B
Ps1
10611
10598
M159-1
CR C.rubrum
CS51
1C
385B V.extorquens
1D
9145
9141
1E
N-6
1F
9421 Ps.rhodos
1G
11278 M.organophilum
1H
602 V.extorquens
317
1J
790
2A
10815 Ps.radiora
9142
2B
434
2C
2CR
2D
617
0-31
45
269
2E
11561 Ps.mesophilica
2F
N-2
2G
R14 My.rubra
3
D12
4
B46
5
% DNA homology

Fig 15

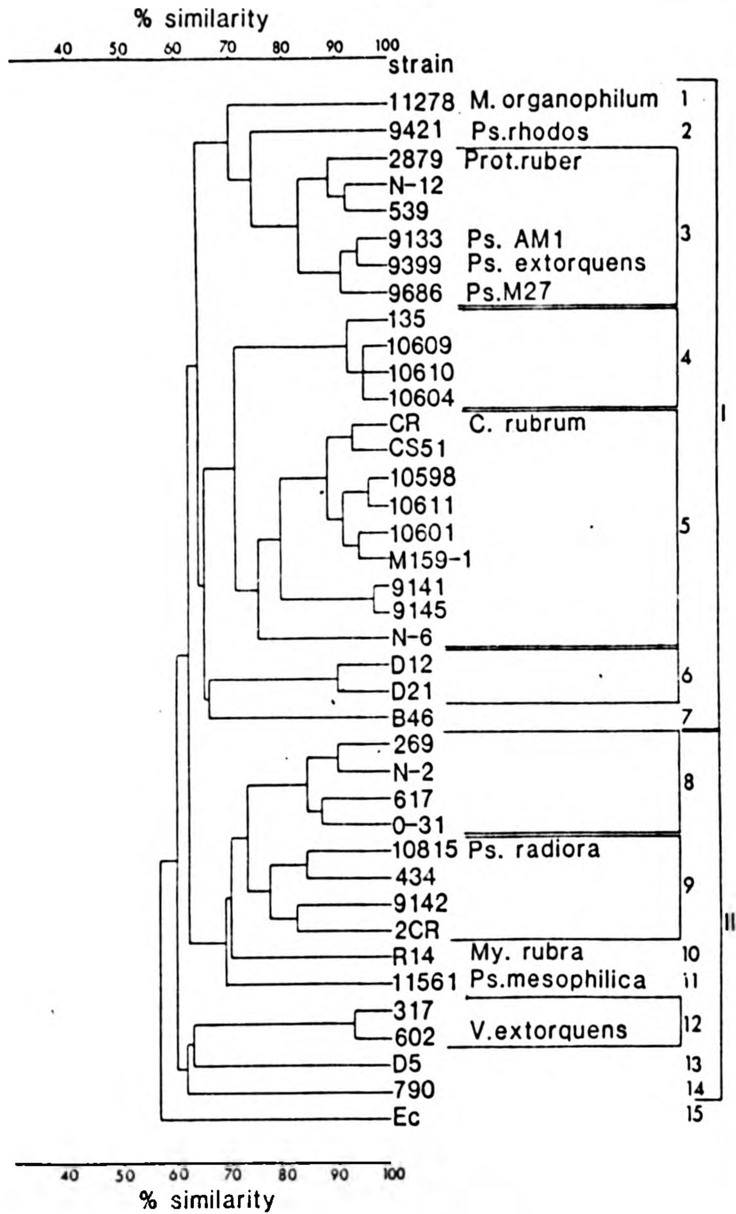
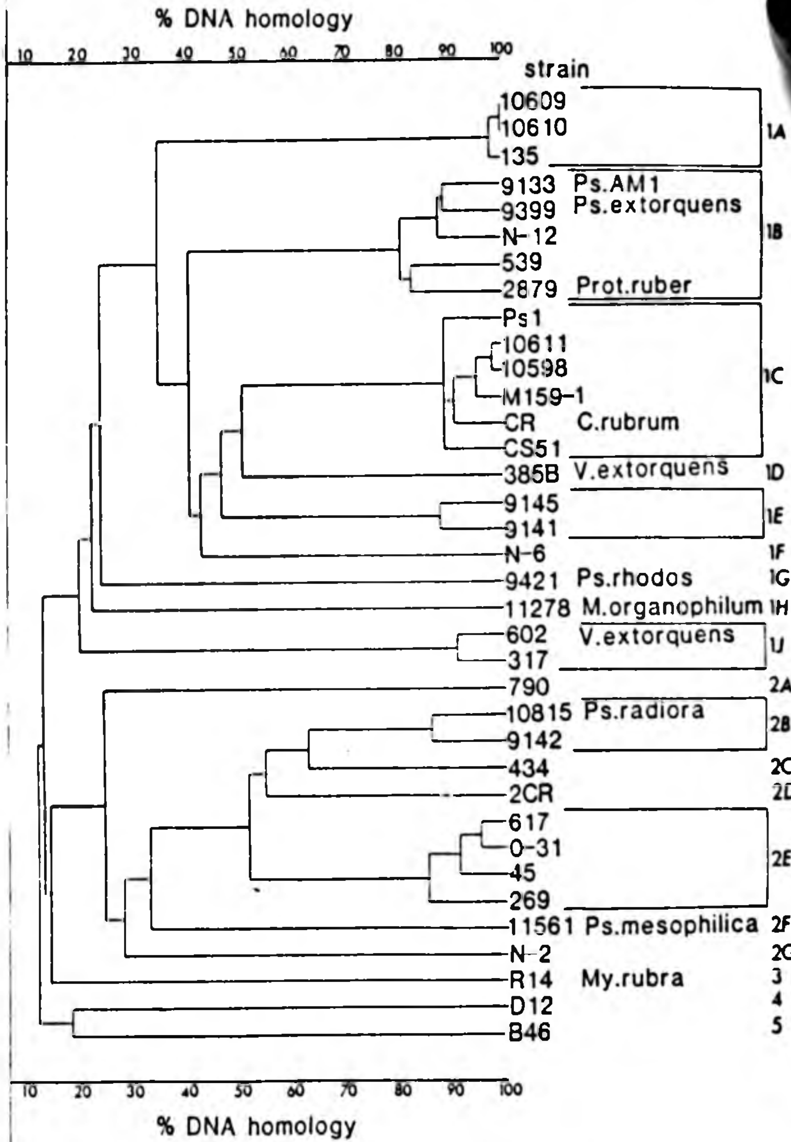



Fig 17



representation of all inter-relatedness as not all possible pairwise combinations of hybridisations were carried out.

The results show a high degree of congruence (r = 0.92) with the results obtained by the PAGE analysis of soluble proteins (Figure 21) similar to that found by Owen and Jackman (1982). At DNA homology levels >20%, major clusters 1 and 2 agree with the major groups in clusters I and II respectively obtained by AUTO EF correlation coefficient analysis of the soluble proteins (Figure 15). Single member clusters 1H, 1J, 1K, 4 and 5 are still valid but show differing linkage values and appear more distantly related by hybridisation studies. Presumably this is because they were approaching the lower limits of resolution of the PAGE system as discussed previously. At DNA homology values >70% there was a high degree of similarity of the corresponding electrophoretograms. Tight clusters 1A, 1B, 1C, 1E, 1J, 2B and 2D agreed well with the clusters found after AUTO EF protein analysis with only a few minor internal rearrangements. Below 70% DNA homology there were obvious deviations from the major banding patterns found in the tight clusters. The results from peak search and position analysis on electrophoretograms similarly showed a significant degree of congruence with the hybridisation values, but exhibited more diverse intra-cluster relatedness. Major clusters are in good agreement between the two.

Strains with zero DNA homology when tested against PPFM DNAs (e.g. *E. coli*) still show over 50% protein pattern similarity. Such artificially high levels are consistent between unrelated strains and must be attributed to the random coincidence of major protein bands and to the

best fit (margin of overlap) method of analysis used where the highest similarity obtained is considered the most appropriate. Forty to 50% similarity is the basal value for discrimination between unrelated strains allowed by the system under the parameters used.

A similar high degree of congruence between DNA hybridisation and PAGE-protein data has consistently been found from previous studies (Owen and Jackman, 1982; Kersters and De Ley, 1975; Biavati *et al.*, 1982). Such agreement is not always so apparent when comparing hybridisation data with results from numerical studies (Johnson, 1973; Owen and Jackman, 1982) or serological studies (Christiansen *et al.*, 1981). When different methods for DNA hybridisation are used on the same organisms, they too generally show a high level of agreement of results (Brenner *et al.*, 1969; Schleifer and Stackebrandt, 1983).

### 6.1.6 DNA Hybrid Dissociation

Listed in Table 8 are values for the melting temperatures of DNA hybrids formed between representative PPFM DNAs. The melting temperature ($Tm_{(e)}$) of a homologous hybrid, formed at Tor, need not always be theoretically or experimentally equivalent to the Tm values for native DNA (Brenner *et al.*, 1969), but the values are usually very close. $Tm_{(e)}$ values calculated for homologous PPFM hybrids were generally 79-81°C which, when corrected for the salt and formamide concentrations, correspond to values of 97-99°C in 1 x SSC. This is in excellent agreement with the theoretical values calculated from the buoyant density of the DNAs (Table 5) and the Tm values observed after thermal denaturation of PPFM DNAs (4.3). Homologous PPFM duplexes formed at

Tor.d have melting profiles essentially equivalent to that of native DNA. With heterologous DNAs, as the level of hybridisation decreases the thermal stability of the resulting heterologous duplexes are less (Johnson, 1973) because of the increased base sequence diversity and base pairing imperfections.

The relationship between % DNA homology and $-\Delta Tm$ (Figure 18) is essentially linear but with a disproportionate increase in $-\Delta Tm_{(e)}$ at lower homology values, below 20%. From the slope of the graph, a decrease in $Tm_{(e)}$ of $1^{o}C$ seems to correspond to a decrease in relative % binding of 9-10%. This value is less than the value of 5% homology $^{o}C^{-1}$ reported by Johnson (1973) and reflects differences in the systems used. A similar value of 8-9% homology $^{o}C^{-1}$ was found with the DMSO system by De Ley *et al.* (1973) and they too showed an increase in $-\Delta Tm_{(e)}$ values at lower homologies. As $Tm_{(e)}$ values are relatively insensitive to fluctuations in the level of filter bound DNA they form an ideal method for cross-checking the spread of values found in heterologous reassociation reactions. $Tm_{(e)}$ values *per se* can therefore be used as a measure of genome diversity amongst bacteria (Johnson and Ordal, 1968). In accord with the DNA homology data, $-\Delta Tm_{(e)}$ values indicated a significant degree of heterogeneity within the PPFMs.

Shown in Figure 19(a) are the melting profiles of homologous and heterologous DNA hybrids alongside the equivalent elution profiles (Figure 19(b)). Both show the proportionate increase in the release of filter bound counts found with heterologous hybrids at lower temperatures. Under sub-optimal conditions ($<$Tor.d) it is often

Figure 18 Correlation between % DNA homology and depression of hybrid melting temperature ($-\Delta Tm_{(e)}$) values for the PPFMs

Results combined from hybridisation experiments with NCIB 11561, 2879 and 135 probe DNAs.

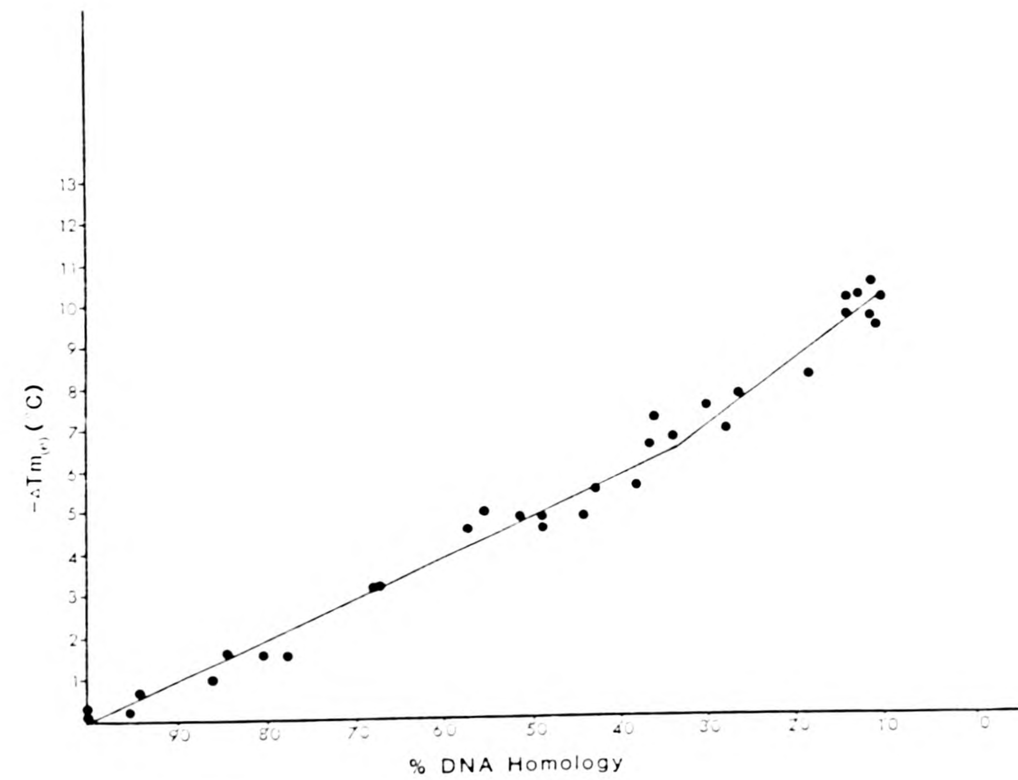


Figure 19a  Thermal Denaturation Profiles for Duplexes Formed between PPFM DNAs

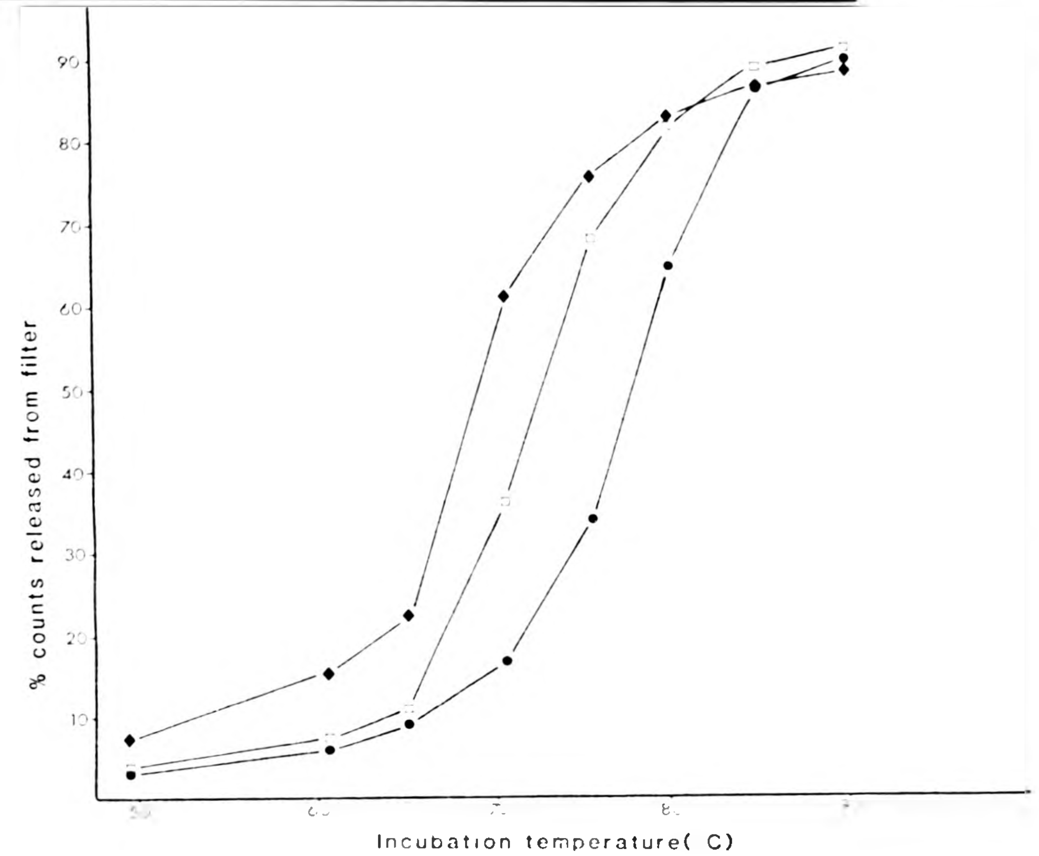
Hybrids formed between PPFM filter bound and NCIB 9133 $^{32}P$ labelled DNAs.

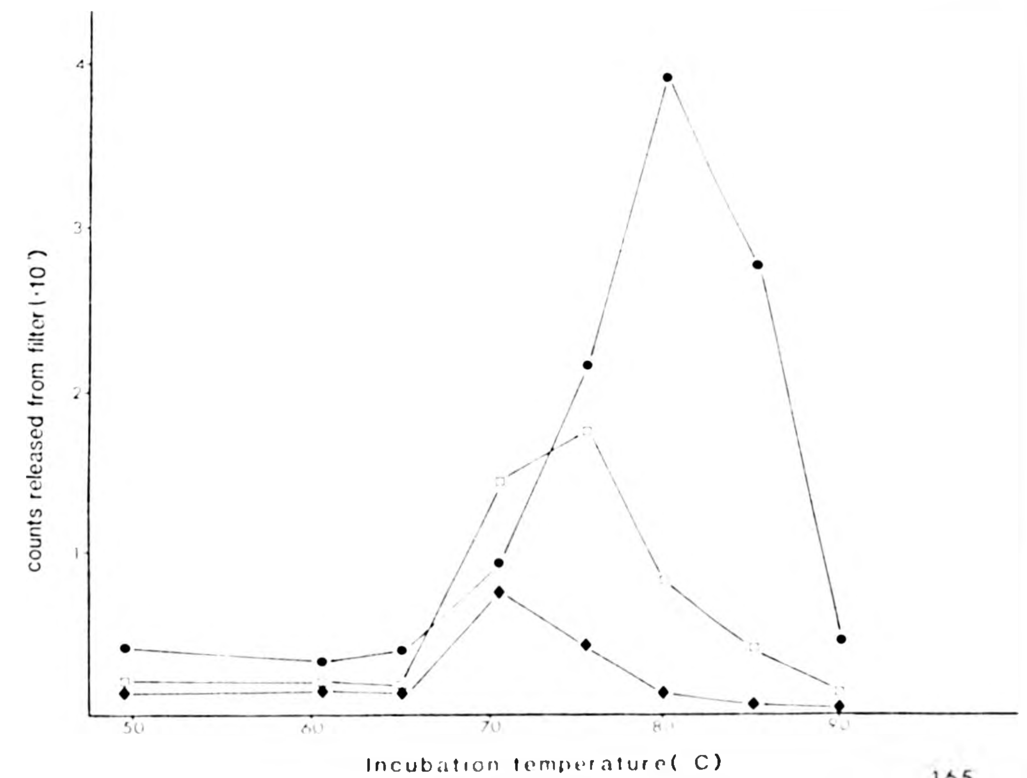
Filter bound DNAs: ●-----● NCIB 9133; □-----□ NCIB 10609; ◆-----◆ NCIB 11561.

Figure 19b  Elution Profiles for Duplexes Formed Between PPFM DNAs

Hybrids formed between PPFM filter bound and NCIB 9133 $^{32}P$ labelled DNAs.

Filter bound DNAs: ●-----● NCIB 9133; □-----□ NCIB 10609; ◆-----◆ NCIB 11561.



possible to distinguish 2 hybrid types, even with homologous reassociation reactions. These are a labile type which is easily denatured, consisting of relatively non-specific duplexes, and the hybrid *sensu stricto* showing true duplexing (Kingsbury *et al.*, 1969). The increase in apparent homology values at sub-optimal temperatures (Table 13) can be explained by an increase in the amount of unstable, less specific hybrid formed. Such hybrids can be dissociated at lower temperatures. At stringent, supra-optimal renaturation temperatures there is a general decrease in the level of heterologous hybridisation as only the most perfect stable duplexes are formed.

Stringency is a term often used with reference to hybridisation studies. Stringent (or exacting) conditions are considered as those allowing detection of only the closest relationships, normally at temperatures <20°C below Tor. Tor is often considered as a temperature of moderate stringency and lower temperatures as non-stringent.

### 6.1.7 Taxonomic Implications

DNA-DNA hybridisation reactions on membrane filters are necessarily based upon averages of multiple reactions occurring between the free denatured DNA, absorbed filter DNA and perhaps some released filter DNA in a common medium (Flavell *et al.*, 1974). It must be presumed that the interactions measured in DNA:DNA hybridisation experiments are a reasonable reflection on the genetic and evolutionary relationships amongst the PPFMs. If DNAs from two strains are able to reassociate only poorly, under optimal conditions, it is evident that these organisms are no longer closely related. Conversely the ability to form

specific duplexes between two strains is indicative of relatedness.

Many attempts have been made to relate DNA homology values to the levels of the taxonomical hierarchy. Historically this has involved investigation of the DNA relatedness between members of well established genera and using these to form standards for unknowns. Different complexities in the established genera make it difficult to propose absolute rules for converting DNA data into meaningful taxonomic groupings. Also, with no standard hybridisation system, it is not always possible to directly compare data from one investigator with that from another. Each result is essentially only correct for the criteria employed in that instance and attempts to equate them with lines of speciation can be somewhat arbitrary. Generalisations, however, can and have been made for data obtained under accepted 'optimal' reaction conditions. Johnson (1973) proposed a DNA homology level of 60% as the lower limit for strains (genosub-species) considered to be members of the same species. This value was later reaffirmed by Schleifer and Stackebrandt (1983) but is not absolute. Some workers have used a value of 70% (Mordarski and Szyba, 1976) and others a value of 55% (Koops and Harms, 1985; Seldin and Dubnau, 1985) as the limit to include organisms as closely related strains within the same species.

Commonly, a DNA homology value of 20% is considered as the lower limit for strains (geno-species) to be taken as closely related members of the same genus (Johnson, 1973; Schleifer and Stackebrandt, 1983). Again with some genera, lower values of 15% (Stackebrandt *et al.*, 1981), 12% (Dopfer *et al.*, 1982), and even 10% (Stackebrandt and Kandler, 1979)

have merited inclusion of organisms in the same genus.

The grouping of organisms by DNA hybridisation data is still rather arbitrary for non-phenotypically defined genera, but when considered with the supplementary phenotypic data available for the PPFMs it provides a powerful tool in measuring their heterogeneity as an aid to the classification of this group of bacteria. The high degree of congruence between the DNA hybridisation data and the PAGE-protein data - Figure 21 - shows that together they provide a powerful force to resolve the taxonomic dilemmas typical of many genera.

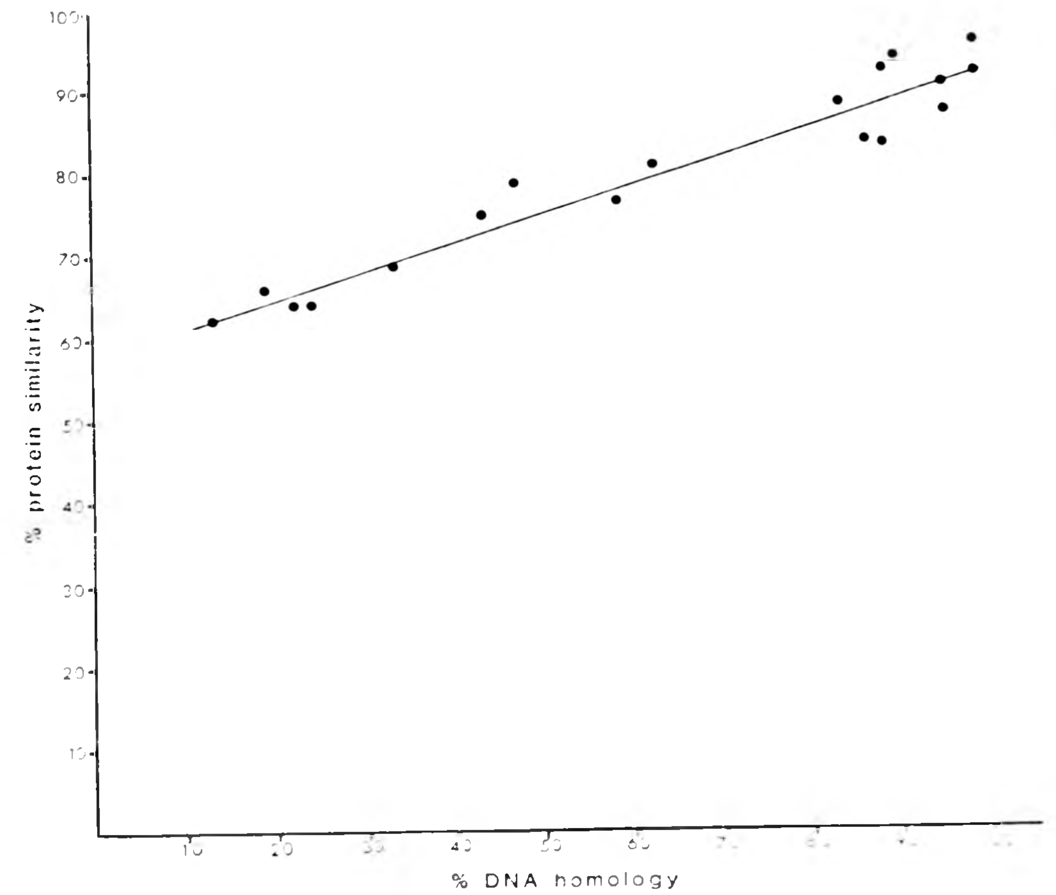
## 6.2 DNA:rRNA Hybridisation

rRNA labelled *in vivo* with $^3$H uracil, was isolated from strain NCIB 11278 at a specific activity of 980 cpm $\mu g^{-1}$ for the combined 16S and 23S moieties. When used in filter hybridisation experiments, very low levels of filter bound counts were obtained. The equivalent to 0.2-0.5% of the input rRNA was bound to each filter giving counts of 5-10 cpm above background level. These results were statistically insignificant and no melting of the hybrids was attempted. Counting of samples for 5 mins each as per the method of De Ley and De Smedt (1975) would not have increased the validity of the results.

No reason can be offered for the low specific activity of rRNA as the label did appear to taken up by by the cells. rRNAs with low specific activities are not unknown and some as low as 960 cpm $\mu g^{-1}$ have been used previously (De Smedt *et al.*, 1980), but in general they must be in

Figure 21 Correlation between DNA homology and PAGE-protein data obtained after studies on the PPFMs

Results combined from experiments with 135, 617 and NCIB 10611 $^{32}P$ labelled DNAs.



the order of >3,000 cpm $\mu g^{-1}$ for efficient counting. Increasing the levels of filter bound DNA and rRNA in individual reactions would have helped but was prohibitive in the amounts of nucleic acids consumed. The main limitation was the availability of rRNAs. For effective hybridisations they would need to be prepared in greater bulk by a method similar to that described by De Ley and De Smedt (1975).

## 6.3 DNA:rDNA Hybridisation

pCOI is a derivative of plasmid pBR325 with an 8.4kb fragment of DNA from *Rhodomicrobium vannielii* RM5 (Warwick), containing 16S rRNA coding sequences, inserted at the Eco RI site within the chloramphenicol resistance gene. The plasmid, maintained in *E. coli* DH1 was kindly supplied by C. Oakley (University of Warwick). pCOI was prepared from a DH1 lysate and the cloned insert isolated then nick translated and used as a probe againt PPFM DNAs, restricted with Eco RI, electrophoresed then blotted onto nitrocellulose. Complementary bands containing putative 16S rRNA coding sequences were identified in all strains tested (Figure 23). Fragments of the corresponding molecular weights were enriched from agar slices dissected after electrophoresis of the digested DNAs. Following purification these were reannealed and ligated with linearised pBR325 vector then transformed into *E. coli* DH1 (Maniatis *et al.*, 1982). Transformants (at a frequency of 0.1-0.3%) were selected for by growth on tetracycline (10 $\mu g\ ml^{-1}$) then growth on chloramphenicol (10 $\mu g\ ml^{-1}$) to look for recombinant plasmids (2% of transformants). Recombinant plasmids were checked for putative 16S rRNA sequences by hybridisation against a probe from the pCOI 16S rRNA

fragment then a more specific end-labelled (Nichols *et al.*, 1982) 16S rRNA probe from *M. organophilum* XX. One putative recombinant plasmid from strain 539 DNA was cut with Eco RI, electrophoresed through agar and the 9kb inserted fragment removed and purified. This was radiolabelled with $^{32}P$ by nick translation then used as a probe against DNA from a variety of PPFM strains.

Results were inconclusive. Individual reactions with 5 $\mu g$ lots of PPFM filter bound and 0.1 $\mu g$ of probe DNAs gave an average binding of the probe of 0.5-0.9%. The levels of binding bore little resemblance to the homology groups determined after DNA:DNA hybridisation studies. The hybrid melting temperature in the homologous system was very low at 70.9°C. One would have expected melting temperatures at around 85°C, equivalent to those obtained with the DNA:DNA system. The spread of $\Delta Tm_{(e)}$ values was only 3.1°C and individual values again bore little relation to the DNA homology groups. The reason for this is not entirely clear but may be due to the fact that the rRNA coding sequences are likely to constitute only a small fraction of the 9kb fragment and the flanking sequences dilute out the specificity of the interaction. Thus the binding would be rather less specific than expected, but does not explain the lowered thermal stability when hybridised to DNA from the source strain 539.

To optimise the system, the rRNA genes must first be subcloned to obtain a probe largely unique for 16S rRNA coding sequences, thus eliminating any spurious effects from flanking DNA sequences. Had time allowed this would have been the next step in the development of the DNA:rDNA system.

Unquestionably, there were several advantages of using cloned rRNA genes as probes in preference to rRNAs. These included: much higher specific activity probes by *in vitro* radiolabelling; hybridisations could be performed with varying degrees of stringency for optimal comparison; rDNA circumvented the problems of secondary structure associated with RNA molecules; DNA was relatively simpler to isolate and work with; homology values *per se* could be more meaningful than with rRNA; rDNA probes were more amenable to the rapid identification of large numbers of unknown organisms by multi-blot analysis.

# 7. Restriction Analysis and Detection of Plasmids in the PPFMs

## 7.1 Restriction Analysis of PPFM DNAs

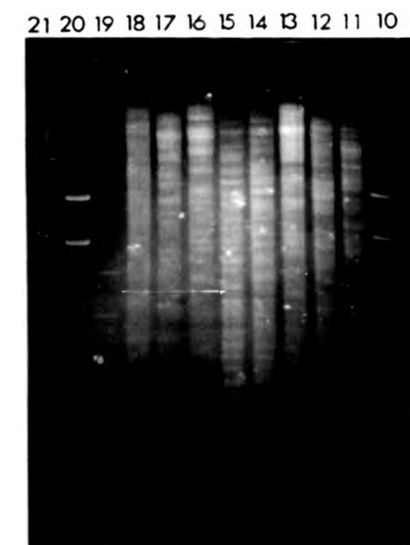
Figure 22 shows the patterns of DNA fragments from selected PPFM DNAs cleaved by restriction endonucleases. Some heavily staining bands are just discernable but in most tracks there is a high level of background due to overlapping of the large number of heterogeneous fragments produced. Restriction enzymes with recognition sites rich in G+C sequence will cut high % (G+C) DNA at many more sites than enzymes which recognise A+T rich sequences, making comparison of patterns more difficult. Of the enzymes used in this study, Eco RI has a hexanucleotide recognition sequence of GAATTC and Bam HI TGGCCA. Thus, one would expect Bam HI to produce more numerous fragments with correspondingly lower molecular weights. This can be seen in Figure 22 where the highest staining intensity (≡DNA fragment concentration) is at a lower molecular weight for Bam HI digests than for Eco RI digests. Both enzymes, however, gave fragments too numerous for good pattern resolution and easy comparison. No difference was noted in the patterns of DNA digested for periods ranging from 30 mins to 7 hrs indicating that the nuclease reactions were effectively complete within the time normally allowed. Had an enzyme with a more A - T dependent restriction site (e.g. Hpa I) been used this may have allowed an increase in the resolution of fragment patterns. From Figure 22 the only digested DNAs with obviously similar patterns were those from strains NCIB 10610 and 135.

Figure 22 Fragment patterns of PPFM DNAs cleaved by restriction endonucleases

Gel (a). Eco RI digestion of DNA from strains, track 1, 135; 2, NCIB 10609; 3, 434; 4, NCIB 9421; 6, 539; 7, *E. coli* C; 8, NCIB 11278. Bam HI digestion of DNAs from strains, track 9, NCIB 10598; 10, NCIB 10610 .

Gel (b). Eco RI digestion of DNA from strains, track 11, 135; 12, NCIB 11278; 13, 539; 14, *Rhodomicrobium vannielii* RM5; 15, *E. coli* C; 16, 317; 17, NCIB 10598; 18, 617; 21, NCIB 11561. Track 20 is an Eco RI digest of pCOI.

(a)

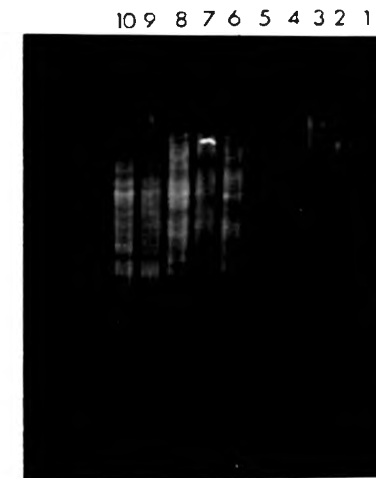



(b)

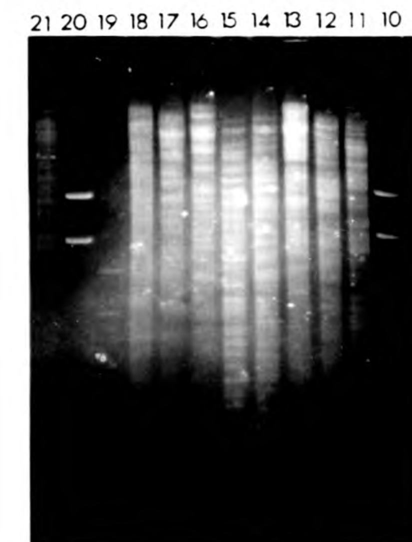
Figure 22 Fragment patterns of PPFM DNAs cleaved by restriction endonucleases

Gel (a). Eco RI digestion of DNA from strains, track 1, 135; 2, NCIB 10609; 3, 434; 4, NCIB 9421; 6, 539; 7, *E. coli* C; 8, NCIB 11278. Bam HI digestion of DNAs from strains, track 9, NCIB 10598; 10, NCIB 10610.

Gel (b). Eco RI digestion of DNA from strains, track 11, 135; 12, NCIB 11278; 13, 539; 14, *Rhodomicrobium vannielii* RM5; 15, *E. coli* C; 16, 317; 17, NCIB 10598; 18, 617; 21, NCIB 11561. Track 20 is an Eco RI digest of pCOI.




(a)




(b)

By choosing a suitable range of restriction enzymes one could estimate the relative % (G+C) content of unknown DNAs. Conversely, if the % (G+C) content and size of the DNA are known one can predict the number of fragments after digestion with any selected enzyme (Razin *et al.*, 1983).

When discrete bands are obtained in restriction patterns it should be possible to compare patterns either, directly, or after scanning the gels or photographic negatives with a microdensitometer. If standard DNAs of known molecular weights are used, as with the protein system, quantitative comparisons can be made (Hookey *et al.*, 1985).

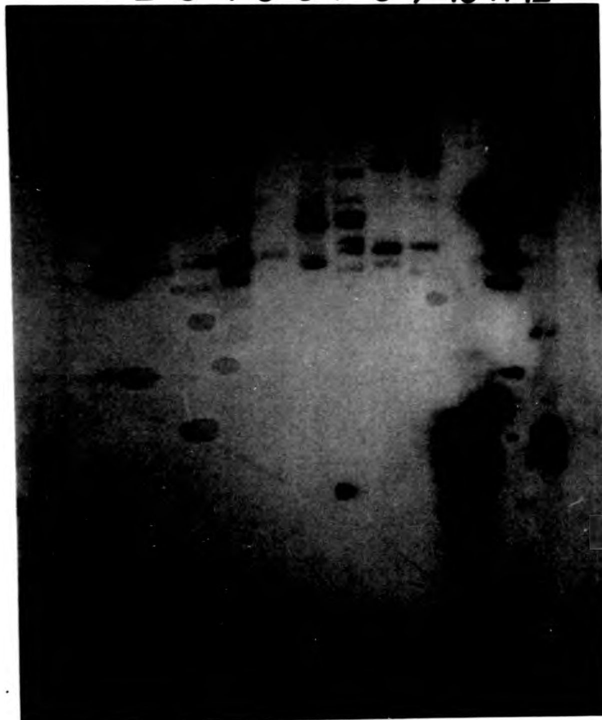
### 7.1.1 Hybridisation of Restricted DNAs

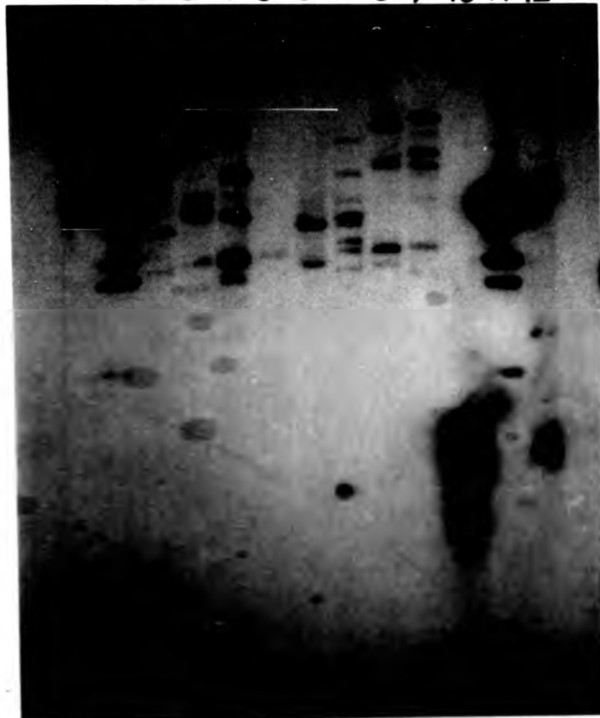
Restriction analysis can be extended to give a direct comparison of genome arrangement. Figure 23 shows an autoradiograph from restricted PPFM DNAs blotted onto nitrocellulose then hybridised against a 16S rDNA probe of 8.4 kb prepared from *Rhodomicrobium vannielii* (RM5) and cloned in pBR325. If the PPFMs were a homogeneous group of organisms one would expect the number and spatial positioning of bands hybridising with the probe to be fairly constant between strains. However, as Figure 23 shows, drastic differences exist between the representatives of the high (>60%) DNA homology clusters indicating a significant degree of heterogeneity within the group.

<u>Figure 23</u> <u>Autoradiograph of restricted PPFM DNAs blotted onto nitrocellulose then hybridised against a probe from the 8.3 kb fragment of pCOI</u>

Eco RI digests of strains, track 2, 135; 3, *M. organophilum* NCIB 11278; 4, 539; 5, 0-31; 6, *R. blastica*; 7, NCIB 10598; 8, 617; 9, *Ps. mesophilica* NCIB 11561; 12, *E. coli* C.

Tracks 1 and 11 - Eco RI digests of pCOI, track 10, Eco RI/Pst I digest of pCOI.


1 2 3 4 5 6 7 8 9 10 11 12


1 2 3 4 5 6 7 8 9 10 11 12

### 7.2 Detection of Plasmids in the PPFMS

#### 7.2.1 In Gel Lysis Method

Screening the PPFMs for plasmid DNA by the in-gel lysis method proved negative for all of the strains tested (Figure 20). This was true both for the method as originally described by Eckhardt (1978) and the flat bed method modified to give improved in-gel cell lysis. With the latter, cell lysis could be observed by a clearing of the cell suspension during electrophoresis of the SDS and by the appearance of chromosomal DNA fragments in the gel matrix. When included, the control strains P11278 and P9399 containing plasmid PUB307 gave similar negative results indicating that cell lysis was not sufficient to allow ready detection of low copy number PPFM plasmids by this method. *E. coli* controls containing pUB307 were readily detected. Alterations to the cell concentration and lysis conditions had no effect. Similar problems have been found by workers using a similar method to screen other methylotrophs for plasmid DNA (Lidstrom and Wopat, 1984).

#### 7.2.2 Alkali Method

The alkali lysis method (Maniatis *et al.*, 1982) again only gave lysis of PPFM strains when the modifications described in 3.12.2 were included. Plasmids in the *E. coli* JC6310 control strain were readily detectd by this method and in P9399 (pUB307) but only using a very high cell density. No plasmid DNA was detected in strain P11278 under any conditions.

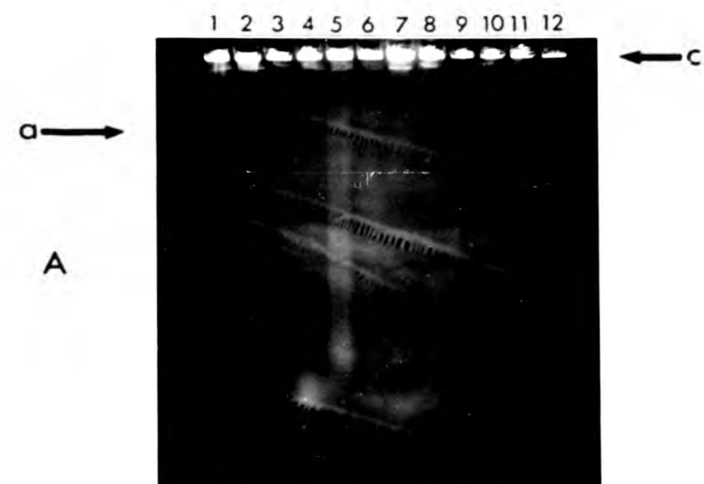
Figure 20 Nucleic Acid from PPFM cells electrophoresed through agarose after disruption by in the in-gel lysis method

Gel A - strains in tracks 12, N-6; 11, 317; 10, M159-1; 9, NCIB 9399; 8, NCIB 9399 (pUB 307); 7, NCIB 10815; 6, NCIB 9421; 5, NCIB 11561; 4, NCIB 11278; 3, NCIB 11278 (pUB 307); 2, 135; 1, *E. coli* K12 (pUB 307).

a = plasmid band detected    c = chromosomal DNA

Gel B - strains in tracks 11, 617; 12, NCIB 10609; 13, NCIB 9133; 14, NCIB 9399; 15, NCIB 9686; 16, M159-1; 17, NCIB 9142; 18, NCIB 9141; 19, 0-31; 20, 269; 21, N-6.

Gel A electrophoresed at 2mA for 60 mins, 50V for 2 hrs then 100V for 2 hrs. Gel B electrophoresed as per methods (3.12.2).



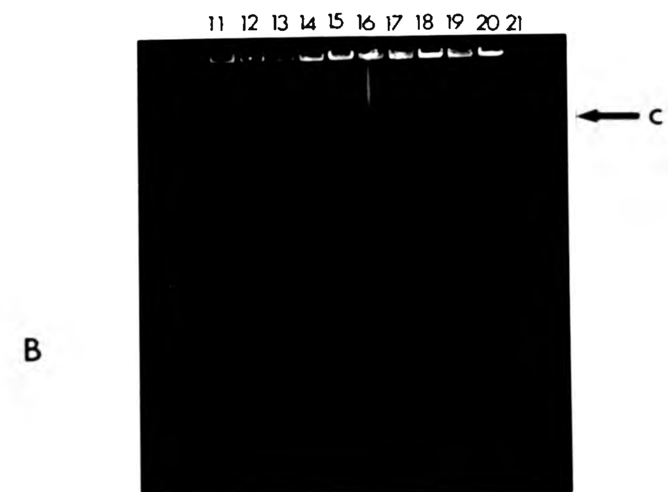



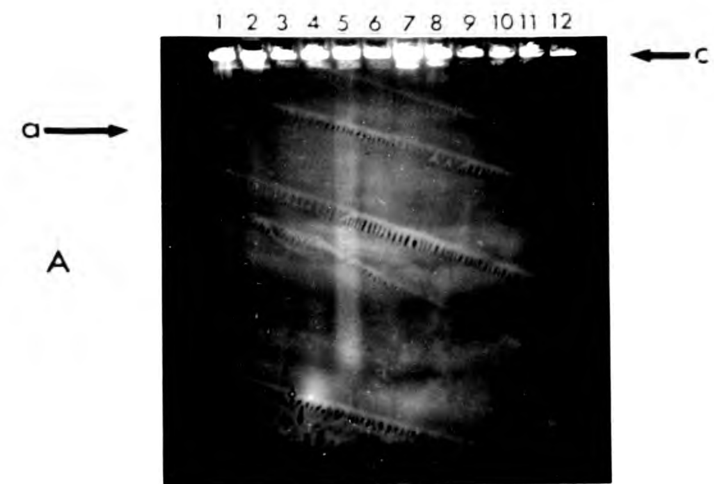
Figure 20 Nucleic Acid from PPFM cells electrophoresed through agarose after disruption by in the in-gel lysis method

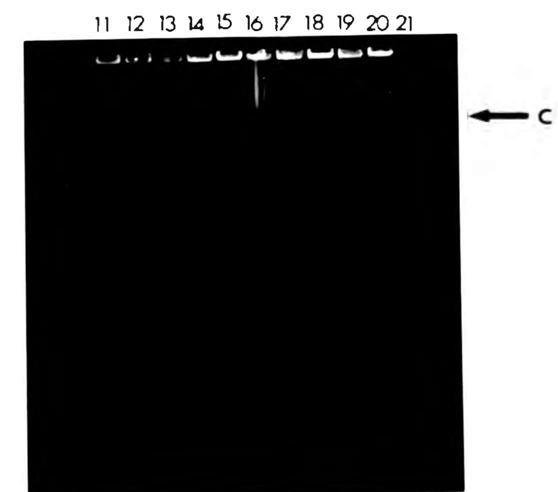
Gel A - strains in tracks 12, N-6; 11, 317; 10, M159-1; 9, NCIB 9399; 8, NCIB 9399 (pUB 307); 7, NCIB 10815; 6, NCIB 9421; 5, NCIB 11561; 4, NCIB 11278; 3, NCIB 11278 (pUB 307); 2, 135; 1, *E. coli* K12 (pUB 307).

a = plasmid band detected c = chromosomal DNA

Gel B - strains in tracks 11, 617; 12, NCIB 10609; 13, NCIB 9133; 14, NCIB 9399; 15, NCIB 9686; 16, M159-1; 17, NCIB 9142; 18, NCIB 91[illegible]; 19, 0-31; 20, 269; 21, N-6.

Gel A electrophoresed at 2mA for 60 mins, 50V for 2 hrs then 100V for [illegible] hrs. Gel B electrophoresed as per methods ([illegible].12.2).



Strains 9399, 9133, 10815, 135, 617 and 9421 tested by this method also gave negative results. Again it would appear that this method is not amenable to strains that cannot be efficiently lysed by conditions close to those originally described by the authors. No evidence was found for the three plasmids reported in *Pseudomonas* AM1 (Warner and Higgins, 1977), or the 8 x $10^7$ molecular weight plasmid reported in *M. organophilum* XX NCIB 11278 (Haber *et al.*, 1983).

The only method by which plasmids in strains P11278 and P9399 could be readily detected was by a maxi-prep (3.2.4) on cells harvested from a litre batch culture. The exactness of this method makes it unsuitable for screening large numbers of unknown strains.

## 8. Discussion and Conclusions

Congruent results from the PAGE-protein analyses, $-\Delta Tm_{(e)}$ values for PPFM hybrids and calibration experiments on the hybridisation system lend support for the validity of the DNA homology data found in this study. It can only be assumed that interference from plasmid DNA was minimal for the strains investigated.

Strains included in some DNA homology groups can be corroborated by experimental observations from previous investigations. Both the PAGE-protein and DNA homology groupings are in agreement with the findings of Urakami and Komagata (1979; 1981) after comparative studies on the fatty acid and coenzyme Q compositions and the electrophoretic mobilities of enzymes from representative PPFM strains. They found that the PPFMs constitute a unique bacterial group. Of special note was their tight clustering of strains (their sub-groups 1, 2 and 3) corresponding to homology groups 1A, 1B and 1C in this study.

Members from DNA homology cluster 1B, exemplified by *P.extorquens* NCIB 9399, have been the organisms of choice for many of the biochemical and genetic studies involving PPFMs. Similar C1 oxidation enzymes, cytochromes and patterns of enzyme induction and suppression have been found in representatives of this group when studied (see Section 1.5). Various workers have also emphasised the very high phenotypic similarity between members of this cluster (Peel and Quayle, 1961; Stocks and

McCleskey, 1964). Some disparaties exist after the discovery of two enzymes for formaldehyde/methanol utilisation in strains NCIB 2879 and NCIB 9399 but only one in strain NCIB 9133, and the presence of three putative plasmids in strain NCIB 9133 and none in strain NCIB 9399. Neither of these offer sufficient reason to cast doubt over the relatedness of this group.

Strains included within DNA homology clusters 1A and 1C are less well studied and little biochemical data is available for comparison.

The type strain of the genus, *Methylobacterium organophilum* XX NCIB 11278 is one of the best studied PPFMs and was found to relate to members from DNA homology groups 1A, 1B and 1C at levels >88% after the phenotypic analysis of Green and Bousfield (1982). Both the DNA homology and PAGE-protein data indicate that strain NCIB 11278 is more loosely associated with these groups. Biochemical studies have indicated that strain NCIB 11278 has a methanol dehydrogenase with properties differing from those found in strains NCIB 9133 and NCIB 9686. Dissimilarities over the number and specificity of C1 assimilation enzymes exist between strains NCIB 11278 and NCIB 9133 but more significant are disparities in the modes of induction and suppression of components of the C1 utilisation pathways (see Sections 1.5, 1.6). Prior to the emendation by Green and Bousfield (1983) the description of the genus *Methylobacterium* (Patt *et al.*, 1976) excluded all non-methane utilising strains but included all facultative methane utilising bacteria irrespective of their overall similarity. With hindsight, such a rigorous classification is clearly inappropriate and

the methane oxidising phenotype has been overweighted for taxonomic purposes. This unstable feature may be plasmid borne, is easily lost and therefore unsuitable as a primary taxonomic criterium. Organisms included in the genus *Methylobacterium* as facultative methanotrophs often bear little other resemblance to the type species and indeed most have since lost this ability or have proved to be mixed cultures of unrelated organisms (Green *et al.*, 1984). The question of whether facultative methanotrophs exist remains to be unambiguously answered.

Excepting cluster 2E, exemplified by strain 617, DNA homology clusters 2-5 constitute a more diverse assemblage of organisms than do members of homology group 1. Some data from previous studies supports this division. Austin and Goodfellow (1979) found significant biochemical and morphological differences when testing their isolate *Pseudomonas mesophilica* NCIB 11561 against strains from DNA homology groups 1B and 1G in this study. Similar deviations were reported by Nishimura *et al.* (1981) comparing their isolate *Pseudomonas radiora* NCIB 10815 with strains from DNA homology group 1B and by Urakami and Komagata (1981) after comparative studies on zymograms.

There is broad general agreement between the strains included in major DNA homology group 1 and groups 2-5 with those in clusters 10 and 11 respectively under the classification of Green and Bousfield (1982). Some internal variation does however occur as to the exact relationships between individual representatives from the DNA homology groups in this study. Several strains associated loosely with DNA homology group 1, namely 317 and 602, formed part of their cluster 11. Such variability

is not untoward as these strains were included at a level proximal to that of the merging of the two major homology groups. No further comment can be made on strains D5, D12 and D21 as these recent isolates have not been included in any previous study.

Despite differences, account must be taken of the overall similarity of DNA homology groups 2-5 with Group 1 as regards their C1 utilisation pathways, DNA base composition, substrate utilisation spectra, phenotypic and other biochemical similarities.

All evidence suggests that the DNA homology data in this study is valid. If so, homology groups can be assigned to taxons of various hierarchial levels. Although each case must be taken on its own merits, if we assume a minimum value of 60% as the species limit (Johnson, 1973), then DNA homology clusters 1A, 1B, 1C, 1E, 1J, 2B and 2D should each be regarded as including closely related strains (genosubspecies) within separate species groups. Cluster 1A includes strains *Pseudomonas* 135 NCIB 10609, NCIB 10610; cluster 1B strains *Pseudomonas* AM1 NCIB 9133, *P. extorquens* NCIB 9399, *P. ruber* NCIB 2879, 539, N-12; cluster 1C strains *Pseudomonas* 1, NCIB 10598, NCIB 10611, M159-1, CS51, *Corynebacterium rubrum* CR; cluster 1E strains NCIB 9141, NCIB 9145; cluster 1J strains 317, 602; cluster 2B strains *P. radiora* NCIB 10815, NCIB 9142, 434; cluster 2E strains 45, 269, 617, 0-31.

Strain 2CR is borderline for inclusion in cluster 2B at a level of 55% homology. The next highest linkages are strains 385B then DNA homolgy group 1E (NCIB 9141, NCIB 9145) at levels of 51 and 47% respectively to

homology group 1C. It is conceivable that these could all be more distantly related strains of the same species. Likewise, the minimum level of relatedness between homology groups 2B, 2C and 2D is 52% indicating that these too could be considered as groups of related strains within a common species. All other strains have sufficiently low homology to warrant exclusion from these groups.

If the PAGE-protein results are combined with the above, the 'species' groups can be extended to include the following strains: cluster 1A - strain NCIB 10604; cluster 1B - strain *Pseudomonas* M27 NCIB 9686; cluster 1C - strain NCIB 10601; cluster 1E - strain D21.

At lower DNA homology values, if we assume a value of 20% (Johnson, 1973) as the absolute limit for inclusion of strains, or species groups, within the same genus then major homology cluster 1 constitutes one genus and major cluster 2 a second. Strains R14, D12 and B46 are excluded from this as single member clusters (genera) but were not tested in all possible pairwise combinations with other PPFMs. Due to the high degree of phenotypic, base composition and cell wall composition similarity shown by all of the PPFMs studied it may be prudent to include all of the strains as members of a single genus. The minimum UPGMA linkage value is 12%, previously considered sufficiently high for inclusion within a common genus (Dopfer *et al.*, 1982). Had further pairwise DNA hybridisations been carried out or the alternative single linkage method been used to analyse the data, these values may have been even greater. The genus most appropriate for the PPFMs would be *Methylobacterium*.

All PPFMs tested are interrelated at DNA homology values significantly greater than the highest values obtained against DNA from an unrelated source (Table 8a).

This study has confirmed an interrelatedness between the PPFMs at a taxonomical level approximating to the genus. But, what of their phylogeny and taxonomic placement relative to other groups of organisms? Unfortunately little can be concluded from the rRNA and rDNA hybridisations attempted in this study but there is no doubt as to the potential of this latter technique.

As long ago as 1970, Quayle suggested that the PPFMs may have evolved from the obligate methane utilisers because of the conservation of C1 utilisation pathways. It was suggested that they had somehow lost the ability to utilise the most reduced of the C1 substrates and retained the potential to utilise the rest. However, studies by Hanson (1980) and this thesis have shown a significant divergence between these two groups, being unrelated below the generic level. However, an in depth study should be undertaken at a more fundamental level to confirm this.

A longer standing association has been that of members of the PPFMs with the genus *Pseudomonas*. Several studies, including the numerical analyses of Green and Bousfield (1982) and Jenkins *et al*. (1984); enzyme and cell wall composition studies of Urakami and Komagata (1979; 1981); base composition and DNA homology values (this study) have demonstrated unequivocally the low relatedness between the PPFMs and

members from the genus *Pseudomonas*. PPFM strains, as well as a spectrum of other equally misnamed bacteria, have been placed in this genus of convenience simply because insufficient data on the organisms prevented them from being assigned their true taxonomic placements. As stated previously, representatives of the PPFMs contain bacteriochlorophyll pigments whose function and ubiquity in the PPFMs has yet to be confirmed. This, and further circumstantial evidence from cell wall analysis led Nishimura *et al*. (1983) to investigate the relatedness between a representative of the PPFMs and the photosynthetic purple non-sulphur bacteria or *Rhodospirillaceae* (Imhoff *et al*., 1984). Some members of the *Rhodospirillaceae* can assimilate methanol or formate via their oxidation product $CO_2$ and the RuBP pathway of $CO_2$ fixation. Representatives can grow anaerobically in the light using the C1 substrate as the reductant and under suitable conditions aerobically in the dark with methanol or formate as the sole carbon and energy source (Truper, 1981). Similarities have been noted in the methanol dehydrogenase from *R. acidophila* when compared with that from the PPFM *Pseudomonas* M27 NCIB 9686 (Sahm *et al*., 1976).

Limited DNA:rRNA hybridisation experiments involving PPFMs have recently been carried out in the laboratory of J. De Ley and coworkers in Belgium. De Ley and his colleagues have been investigating the phylogenetic relationships between the major eubacterial groups for many years, and the relative positions of the *Rhodospirillaceae* have been estimated (Gillis *et al*., 1982; De Ley, personal communication). On the basis of rRNA similarities an rRNA superfamily consisting of the genera *Rhodopseudomonas*, *Rhodomicrobium*, *Rhodospirillum*, *Beijerinckia*,

*Agrobacterium*, *Rhizobium*, *Spirillum*, *Paracoccus*, *Zymomonas*, *Gluconobacter* and *Acetobacter* has been established. This is in agreement with the results obtained after oligonucleotide sequencing studies (Gibson *et al.*, 1979). Because of the putative link of the PPFMs with the purple non-sulphur bacteria, DNA from representative strains *P. radiora* NCIB 10815 and 317 have now been included in their DNA:rRNA studies and indeed have been shown to group within the boundaries of the genus *Rhodopseudomonas* (M. Gillis, personal communication). In particular the PPFMs have been found to be related to *Rhodopseudomonas palustris* and less so to *R. sphaeroides*. However, the *Rhodospirillaceae* form a broad phylogenetic unit so at best the PPFMs can be described as being related in an evolutionary sense to certain members within it. In this study the highest DNA homology values between PPFM DNA and DNA from an unrelated source were indeed found with members of the *Rhodospirillaceae*, but not at significant levels. *R. palustris* was not tested against the PPFMs.

Clearly more detailed analysis is required before firm conclusions can be made as to the true phylogenetic relationship between these two groups of organisms. The *Rhodospirillaceae* are known to be genealogically inter-mixed with other non-phototrophic relatives (Truper, 1981) so a proximal relationship with the PPFMs is feasible. Further supportive evidence such as that from oligonucleotide sequencing of rRNAs from the PPFMs would be desirable as this is in a sense a more absolute measure of phylogenetic distance.

Such evidence may inspire further investigation of the PPFMs for additional vestigal functions from a possible photosynthetic ancestry. No investigations into the possible occurrence of the key enzyme ribulose bisphosphate carboxylase or other components of the Calvin cycle have been reported in these organisms.

In summary, if considered as members of a common genus the pink-pigmented facultative methylotrophs consist of a number of closely related species groups and single membered species. The singularity of some strains must be due at least partly to the randomness of strain selection after the study of Green and Bousfield (1982) and the fact that they were not tested in all possible pairwise combinations with other strains. Evidence suggests that the PPFMs may be most closely linked phylogenetically to members of the purple non-sulphur photosynthetic bcateria.

Allen, R. S. and Hanson, R.S. (1985). Construction of broad host range cosmid cloning vectors. Identification of genes necessary for growth on methanol in *Methylobacterium organophilum* strain XX. J. Bact. 161, 955-962.

Ambler, R. P. and Meyer, T. E. (1979). Cytochrome $C_2$ sequence variation among the recognised species of purple non-sulphur photosynthetic bacteria. Nature 278, 659-660.

Anthony, C. (1975). The microbial metabolism of Cl compounds: The cytochromes of *Pseudomonas* AM1. Biochem. J. 146, 289-298.

Anthony, C. (1981). Electron transport in methylotrophic bacteria. In: Microbial Growth on Cl Compounds, pp. 220-230. H. Dalton (Ed.), Heyden, London.

Anthony, C. (1982). The Biochemistry of Methylotrophs, Academic Press, London.

Anthony, C. and Zatman, L. J. (1964). Microbial oxidation of methanol: Isolation and properties of *Pseudomonas* M27. Biochem. J. 92, 609-614.

Athwal, R. S., Deo, S. S. and Imaeda, T. (1984). DNA relatedness among *Mycobacterium leprae*, *Mycobacterium lepraemurium* and selected bacteria by dot blot and spectrophotometric DNA hybridisation assays. Int. J. Syst. Bact. 34, 371-375.

Attwood, M. M. and Quayle, J. R. (1984). Formaldehyde as a central intermediary metabolite of methylotrophic metabolism. In: Microbial Growth on Cl Compounds, pp. 315-323. R. L. Crawford and R. S. Hanson (Eds.), ASM, Washington.

Austin, B. and Goodfellow, M. (1979). *Pseudomonas mesophilica*, a new species of pink bacteria isolated from leaf surfaces. Int. J. Syst. Bact. 29, 373-378.

Baker, R. F. (1977). Binding of DNA to nitrocellulose under denaturing conditions. Analytical Biochemistry 78, 569-571.

Balch, E., Fox, G. E., Magrum, L. J., Woese, C. R. and Wolfe, R. S. (1979). Methanogens: a re-evaluation of a unique biological group. Microbiol. Revs. 43, 260-296.

Beardmore-Gray, M. and Anthony, C. (1984). Methanol-cytochrome c oxidoreductase activity of methylotrophs. In: Microbial Growth on C1 Compounds. R. L. Crawford and R. S. Hanson (Eds.), ASM, Washington.

Biavati, B., Scardovi, V. and Moore, W. E. C. (1982). Electrophoretic patterns of proteins in the genus *Bifidobacterium* and proposal of four new species. Int. J. Syst. Bact. 32, 358-373.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acid Res. 7, 1513.

Bishop, J. O. (1970). Examination of an equilibrium interpretation of DNA-RNA hybridisation data. Biochem. J. 116, 223-228.

Bishop, J. O., Robertson, F. W., Burns, J. A. and Melli, M. (1969). Methods for analysis of DNA-RNA hybridisation data. Biochem. J. 115, 361-370.

Bousfield, I. J. and Green, P. N. (1985). Reclassification of bacteria of the genus *Protomonas* (Urakami and Komagata, 1984) in the genus *Methylobacterium* (Patt, Cole and Hanson) emend (Green and Bousfield, 1983). Int. J. Syst. Bact. 35, 209.

Bousfield, I. J., Smith, G. L., Dando, T. R. and Hobbs, G. (1983). Numerical analysis of total fatty acid profiles in the identification of *Nocardioform* and some other bacteria. J. Gen. Microbiol. 129, 375-394.

Bradley, S. G. (1980). DNA reassociation and base composition. In: Microbial Classification and Identification, SAB Symposium No. 8, pp. 273-297. M. Goodfellow and R. Board (Eds.), Academic Press, London.

Brenner, D. J. and Cowie, D. B. (1968). Thermal stability of *Escherichia coli* - *Salmonella typhimurium* DNA duplexes. J. Bact. 95. 2258-2262.

Brenner, D. J., Fanning, G. R., Johnson, K. E., Citarella, R. V. and Falkow, S. (1969). Polynucleotide sequence relationships among members of *Enterobacteriaceae*. J. Bact. 98, 637-650.

Burton, K. (1968). Determination of DNA concentration with diphenylamine. Meth Enzymol. 12, 163-166.

Byrom, D. (1981). Taxonomy of the methylotrophs - a reappraisal. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

Cannon, G., Heinhorst, S. and Weissbach, A. (1985). Quantitative molecular hybridisation on nylon membranes. Anal. Biochem. 149. 229-237.

Chan, H. W. and Ross, R. F. (1984). Restriction endonuclease analyses of 2 porcine *Mycoplasma* DNAs: sequence specific methylation in the *Mycoplasma hyopneumoniae* genome. Int. J. Syst. Bact. 34, 16-20.

Chandra, T. S. and Shethna, Y. I. (1975). Isolation and characterisation of some new oxalate decomposing bacteria. Antonie van Leeuwenhoek 41, 101-111.

Christiansen, C., Black, F. T. and Freundt, E. A. (1981). Hybridisation experiments with DNA from *Ureaplasma urealyticum* serovars I to VIII. Int. J. Syst. Bact. 31, 259-262.

Colby, J., Dalton, H. and Whittenbury, R. (1979). Biological and biochemical aspects of microbial growth on C1 compounds. Ann. Rev. Microbiol. 33, 481-517.

Colby, J. and Zatman, L. J. (1973). Trimethylamine metabolism in obligate and facultative methylotrophs. Biochem. J. 132, 101-112.

Corpe, W. A. and Basile, D. V. (1982). Methanol utilising bacteria associated with green plants. Dev. Ind. Microbiol. 23, 483-494.

Crameri, R., Hintermann, G. and Hutter, R. (1983). DNA restriction endonuclease fingerprint characterisation of *Actinomycete* strains. Int. J. Syst. Bact. 33, 652-655.

Crosa, J. H., Brenner, D. J. and Falkow, S. (1973). Use of single strand specific nuclease for analysis of bacterial and plasmid DNA homo- and heteroduplexes. J. Bact. 115, 904-911.

Dalton, H. (1981). Methane monooxygenases from a variety of microbes. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

Davis, R. W., Simon, M. and Davidson, N. (1971). Electron microscope heteroduplex methods for mapping regions of base sequence homology of nucleic acids. Meth. Enzymol. 11, 413-428.

Dekio, S., Yamaski, R., Jidoi, J., Hori, H. and Osawa, S. (1984). Secondary structure and phylogeny of *Staphylococcus* and *Micrococcus* 5S rRNAs. J. Bact. 159, 233-237.

De Ley, J. (1970a). Hybridisation of DNA. In: Methods in Microbiology CHII. J. Norris and D. Ribbons (Eds.), Academic Press, London.

De Ley, J. (1970b). The determination of molecular weight of DNA per bacterial nucleoid. In: Methods in Microbiology 5A, pp.302-309. J. Norris and D. Ribbons (Eds.), Academic Press, London.

De Ley, J. (1970c). Re-examination of the association between melting point, buoyant density and chemical base composition of DNA. J. Bact. 101. 738-754.

De Ley, J., Cattoir, H. and Reynaerts, A. (1970). The quantitative measurement of DNA hybridisation from renaturation rates. Eur. J. Biochem. 12, 133-142.

De Ley, J. and De Smedt, J. (1975). Improvements of the membrane filter method for DNA-rRNA hybridisation. Antonie van Leeuwenhoek 41, 287-307.

De Ley, J. and Tijtgat, R. (1970). Evaluation of membrane filter methods for DNA-DNA hybridisation. Antonie van Leeuwenhoek 36, 461-474.

De Ley, J., Tijtgat, R., De Smedt, J. and Michiels, M. (1973). Thermal stability of DNA:DNA hybrids within the genus *Agrobacterium*. J. Gen. Microbiol. 78, 241-252.

den Dooren de Jong, L. E. (1927). Uber protaminophage Bakterien. Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten and Hygiene 71, 193-232.

Denhardt, D. T. (1966). A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Comm. 23, 641-646.

De Smedt, J., Bauwens, M., Tijtgat, R. and De Ley, J. (1980). Intra- and intergeneric similarities of rRNA cistrons of free-living nitrogen fixing bacteria. Int. J. Syst. Bact. 30. 106-122.

Dickerson, R. E. (1980). Evolution and gene transfer in purple photosynthetic bacteria. Nature 283. 210-212.

Dijkhuizen, L. (1983). Metabolic pathways involved in bacterial growth on C1 compounds. In: Microbiology, 1983. D. Schlessinger (Ed.), ASM, Washington.

Dopfer, H., Stackebrandt, E. and Fielder, F. (1982). Nucleic acid hybridisation studies on *Microbacterium*, *Curtobacterium*, *Agromyces* and related taxa. J. Gen. Microbiol. 128. 1697-1708.

Doudoroff, M. and Palleroni, N. J. (1974). Genus I, *Pseudomonas*. In: Bergeys Manual of Determinative Bacteriology, 8th Edition. R. E. Buchanan and N. E. Gibbons (Eds.), Williams and Wilkins, Baltimore.

Downs, J. and Harrison, D. E. F. (1974). Studies on the production of pink pigment in *Pseudomonas extorquens* NCIB 9399 growing in continuous culture. J. App. Bact. 37, 65-74.

Duine, J. A. and Frank, J. (1981). Methanol dehydrogenase, a quinoprotein. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

Dunstan, P. M., Anthony, C. and Drabble, W. T. (1972). The role of glyoxylate and acetate in the growth of *Pseudomonas* AM1 on ethanol and C1 compounds. Biochem. J. 128. 107-115.

Eady, R. R. and Large, P. J. (1971). Microbial oxidation of amines: spectral and kinetic properties of the primary amine dehydrogenase of *Pseudomonas* AM1. Biochem. J. 123, 757-771.

Eckersley, K. and Dow, C. S. (1980). *Rhodopseudomonas blastica* sp. nov.: a member of the *Rhodospirillaceae*. J. Gen. Microbiol. 119. 465-473.

Eckhardt, T. (1978). A rapid method for the identification of plasmid DNA in bacteria. Plasmid 1, 584-588.

Feltham, R. K. A. and Sneath, P. H. A. (1979). Quantitative comparison of electrophoretic traces of bacterial proteins. Computers and Biomedical Research 12, 247-263.

Fishmann, M. M. and Schiff, H. F. (1968). The use of nitrocellulose filters for DNA samples. Biochim. Biophys. Acta 157. 419-420.

Flavell, R. A., Birfelder, E. J. and Sanders, J. P. M. (1974). DNA-DNA hybridisation on nitrocellulose filters. I. General considerations and non-ideal kinetics. Eur. J. Biochm. 47, 535-543.

Flavell, R. A., Borst, P. and Birfelder, E. J. (1974). DNA-DNA hybridisation on nitrocellulose filters. 2. Concatenation effects. Eur. J. Biochem. 47, 545-548.

Fotos, P., Gerencser, M. A. and Yelton, D. B. (1984). Strain differentiation of *Rothia dentocariosa* and related isolates by SDS-polyacrylamide gel electrophoresis. Int. J. Syst. Bact. 34, 102-106.

Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Balch, W. E., Tanner, R. S., Magrum, L. J., Zablem, L. B., Blakemore, R., Gupta, R., Bowen, L., Lewis, B. J., Stahl, D. A., Luehrsen, K. R., Chen, K. N. and Woese, C. R. (1980). The phylogeny of prokaryotes. Science 209. 457-463.

Fulton, G. L., Nunn, D. N. and Lidstrom, M. E. (1984). Molecular cloning of a malyl coenzyme A lyase gene from *Pseudomonas* strain AM1, a facultative methylotroph. J. Bact. 160. 718-723.

Galchenko, V. F., Andreev, L. V. (1984). Taxonomy of obligate methanotrophs. In: Microbial Growth on C1 Compounds. R. L. Crawford and R. S. Hanson (Eds.), ASM, Washington.

Gibson, J., Stackebrandt, E., Zablen, L. B., Gupta, R. and Woese, C. R. (1979). A phylogenetic analysis of the purple photosynthetic bacteria. Current Microbiology 3, 59-64.

Gillespie, S. and Gillespie, D. (1971). RNA-DNA hybridisation in aqueous solutions and in solutions containing formamide. Biochem. J. 125. 481-487.

Gillespie, D. and Spiegelman, S. (1965). A quantitative assay for DNA-RNA hybrids with DNA immobilised on a membrane. J. Mol. Biol. 12. 829-842.

Gillis, M. and De Ley, J. (1975). Determination of the molecular complexity of double-stranded phage genome DNA from initial renaturation rates. The effect of DNA base composition. J. Mol. Biol. 98, 447-464.

Gillis, M., De Ley, J. and De Cleene, M. (1970). The determination of molecular weight of bacterial genome DNA from renaturation rates. Eur. J. Biochem. 12, 143-153.

Gillis, M., Dejonghe, J., Smet, A., Onghenae, G. and De Ley, J. (1982). Intra- and intergeneric similarities of the rRNA cistrons in the *Rhodospirillaceae*. In: Abstracts of IVth International Symposium on Photosynthetic Prokaryotes, Bambannes, France.

Graf, W. and Bauer, L. (1973). Red bacterial growth in tap water systems. Zentralblatt fur Bakteriologie, Parasitenkunde, Infectionskrankheiten and Hygiene 236. 513-530.

Green, P. N. and Bousfield, I. J. (1981). The taxonomy of the Pink Pigmented Facultative Methylotrophs. In: Microbial Growth on Cl Compounds. H. Dalton (Ed.), Heyden, London.

Green, P. N. and Bousfield, I. J. (1982). A taxonomic study of some gram negative facultatively methylotrophic bacteria. J. Gen. Microbiol. 128, 623-638.

Green, P. N. and Bousfield, I. J. (1983). Emendation of *Methylobacterium*, Patt *et al*., 1976: *Methylobacterium rhodos* (Heumann, 1962): *Methylobacterium radiora* (Ito and Iizuka, 1971): *Methylobacterium mesophilica* (Austin and Goodfellow, 1979). Int. J. Syst. Bact. 33, 875-877.

Green, P. N. and Gibson, D. M. (1984). Carbohydrate metabolism in some methylotrophic bacteria. FEMS Microbiol. Letts. 23, 31-34.

Green, P. N., Hood, D. and Dow, C.S. (1984). Taxonomic status of some methylotrophic bacteria. In: Microbial Growth on C1 Compounds. R. L. Crawford and R. S. Hanson (Eds.), ASM, Washington.

Grimont, P. A. D., Popoff, M. Y., Grimont, F., Coynault, C. and Lemelin, M. (1980). Reproducibility and correlation study of three DNA hybridisation procedures. Curr. Microbiol. 4, 325-330.

Haber, C. L., Allen, L. N., Zhao, S. and Hanson, R. S. (1983). Methylotrophic bacteria: Biochemical diversity and genetics. Science 221. 1147-1153.

Hames, B. D. and Rickwood, D. (1981). Gel electrophoresis of proteins. IRL Press, London.

Hanna, J., Neill, S. D., O'Brien, J. J. and Ellis, W. A. (1983). Comparison of aerotolerant and reference strains of *Campylobacter* species by polyacrylamide gel electrophoresis. Int. J. Syst. Bact. 33, 143-146.

Hanson, R. S. (1980). Ecology and diversity of methylotrophic organisms. Adv. Appl. Microbiol. 26, 3-39.

Harder, W. and Quayle, J. R. (1971a). Aspects of glycine and serine biosynthesis during growth of *Pseudomonas* AM1 on C1 compounds. Biochem. J. 121, 763-769.

Harder, W. and Quayle, J. R. (1971b). The biosynthesis of serine and glycine in *Pseudomonas* AM1 with special reference to growth on carbon sources other than C1 compounds. Biochem. J. 121. 753-762.

Hayward, A. C. (1960). Relationship between *Protaminobacter ruber* and some red pigmented pseudomonads. J. Appl. Bact. 23, 11.

Hersh, L. B. (1973). Malate adenosine triphosphate lyase. J. Biol. Chem. 248, 7295-7303.

Heumann, W. (1962). In: Biologisches Zentralblatt. 81, 341-354.

Higgins, I. J., Best, D. J., Hammond, R. C. and Scott, D. (1981). Methane-oxidising microorganisms. Microbiol. Revs. 45, 556-590.

Hintermann, G., Crameri, R., Keiser, T. and Hutter, R. (1981). Restriction analysis of the *Streptomyces glavescens* genome by agarose gel electrophoresis. Arch. Microbiol. 130. 218-222.

Holloway, B. W. (1981). The application of pseudomonad based genetics to methylotrophs. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

Hookey, J. V., Waitkins, S. A. and Jackman, P. J. H. (1985). Numerical analysis of *Leptospira* DNA restriction endonuclease patterns. FEMS Microbiol. Letts. 29, 185-188.

Hori, H. and Osawa, S. (1979). Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S RNA species. Proc. Nat. Acad. Sci. USA 76, 381-385.

Hotchkin, P. G. and Kaya, H. K. (1984). Electrophoresis of soluble proteins from two species of *Xenorhabdus*, bacteria mutualistically associated with the nematodes *Steinernema* spp and *Heterorhabditis* spp. J. Gen. Microbiol. 130, 2725-2731.

Hou, C. T., Patel, R. N., Laskin, A. I. and Barwabe, N. (1979). Microbial oxidation of gaseous hydrocarbons: Production of methyl ketones from their corresponding secondary alcohols by methane and methanol-grown microbes. Appl. Env. Microbiol. 38, 127-134.

Hoyer, B. H., McCarthy, B. J. and Bolton, E. T. (1964). A molecular approach in the systematics of higher organisms. Science 144, 959-967.

Imhoff, J. F., Truper, H. G. and Pfennig, N. (1984). Rearrangement of the species and genera of the phototrophic purple nonsulphur bacteria. Int. J. Syst. Bact. 34, 340-343.

Ito, H. and Iizuka, H. (1971). Taxonomic studies on a radio-resistant *Pseudomonas*, part XIII. Agricult. Biol. Chem. 35, 1566-1571.

Jackman, P. J. H. (1982). Classification of *Corynebacterium* species from axillary skin by numerical analysis of electrophoretic protein patterns. J. Med. Microbiol. 15, 485-492.

Jackman, P. J. H. (1983). Bacterial Taxonomy based on electrophoretic whole-cell protein patterns. In: The Use of Chemotaxonomic Methods for Bacterial Taxonomy. M. Goodfellow and D. Minnikin (Eds.), Academic Press, London.

Jackman, P. J. H., Feltham, R. K. A. and Sneath, P. H. A. (1982). A program in basic for numerical taxonomy of microorganisms based on electrophoretic protein patterns. Microbios. Letts. 23, 87-98.

Janota-Bassalik, L. and Pedyk, D. (1961). Ability of *Flavobacterium extorquens* to utilise various sources of carbon with particular reference to glucose. Acta Microbiologica Polonica 10, 225-238.

Jarvis, B. D. W., Dick, A. G. and Greenwood, R. M. (1980). DNA homology among strains of *Rhizobium trifolii* and related species. Int. J. Syst. Bact. 30, 42-52.

Jellum, E., Tingelstad, V. and Olsen, I. (1984). Differentiation between *Actinobacillus actinomycetecomitans* and *Haemophilus aphrophilus* by high resolution 2-D protein electrophoresis. Int. J. Syst. Bact. 34, 478-483.

Jenkins, O., Byrom, D. and Jones, D. (1984). Taxonomic studies on some obligate methanol utilising bacteria. In: Microbial Growth on C1 Compounds. R. L. Crawford and R. S. Hanson. (Eds.), ASM, Washington.

Jeyaseelan, K. and Guest, J. R. (1979). Transfer of antibiotic resistance to facultative methylotrophs with plasmid R68.45. FEMS Microbiol. Letts. 6, 87-89.

Johnson, J. L. (1973). Use of nucleic acid homologies in the taxonomy of anaerobic bacteria. Int. J. Syst. Bact. 23, 308-315.

Johnson, J. L. (1981). Genetic Characterisation. In: Manual of Methods for General Bacteriology, Ch. 22. Gerhardt *et al* (Eds.), ASM, Washington.

Johnson, J. L. and Ordal, E. J. (1968). DNA homology in bacterial taxonomy: Effect of incubation temperature on reaction specificity. J. Bact. 95, 893-900.

Johnson, P. A. and Quayle, J. R. (1964). Microbial Growth on C1 Compounds: 6. Oxidation of methanol, formaldehyde and formate by methanol grown *Pseudomonas* AM1. Biochem. J. 93, 281-290.

Kafatos, F. C., Weldon-Jones, C. and Efstratiadis, A. (1979). Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridisation procedure. Nuc. Acids Res. 7, 1541-1522.

Kersters, K. and De Ley, J. (1975). Identification and grouping of bacteria by numerical analysis of their electrophoretic protein patterns. J. Gen. Microbiol. 87, 333-342.

Kersters, K. and De Ley, J. (1980). Classification and identification of bacteria by electrophoresis of their proteins. In: Microbial Classification and Identification, SAB Symposium No. 8, pp.273-297. M. Goodfellow and R. Board (Eds.), Academic Press, London.

Kersters, K., Hinz, K. H., Hertle, A., Segers, P., Lievens, A., Siegmann, O. and De Ley, J. (1984). *Bordetella avium* sp. nov. isolated from the respiratory tracts of turkeys and other birds. Int. J. Syst. Bact. 34, 56-70.

Kingsbury, D. T., Fanning, G. R., Johnson, K. E. and Brenner, D. J. (1969). Thermal stability of interspecies *Neisseria* DNA duplexes. J. Gen. Microbiol. 55, 201-208.

Kirby, K. S., Fox-Carter, E. and Guest, M. (1967). Isolation of DNA and rRNA from bacteria. Biochem. J. 104. 258-262.

Koops, H. and Harms, H. (1985). DNA homologies among 96 strains of ammonia-oxidising bacteria. Arch. Microbiol. 141. 214-218.

Kortstee, G. J. J. (1980). The homoisocitrate-glyoxylate cycle in pink facultative methylotrophs. FEMS Microbiol. Letts. 8, 59-65.

Kortstee, G. J. J. (1981). The second part of the $Icl^-$ serine pathway. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

Kouno, K. and Ozaki, A. (1975). Distribution and identification of methanol-utilising bacteria. In: Microbial Growth on C1 Compounds. G. Tenui *et al.* (Eds.), Society of Fermentation Technology, Japan.

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

Large, P. J. (1981). Microbial growth on methylated amines. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

Large, P. J., Peel, D. and Quayle, J.R. (1961). Synthesis of cell constituents by methanol and formate grown *Pseudomonas* AM1 and methanol grown *Hyphomicrobium vulgare*. Biochem. J. 81, 470-480.

Large, P. J., Peel, D. and Quayle, J. R. (1962). Distribution of radioactivity in metabolism of methanol grown *Pseudomonas* AM1 after incubation with $^{14}C$-methanol and $^{14}C$-bicarbonate. Biochem. J. 82, 483-488.

Large, P. J. and Quayle, J. R. (1963). Enzyme activities in extracts of *Pseudomonas* AM1. Biochem. J. 87, 386-396.

Legault-Demare, J., Desseaux, B., Heyman, T., Seror, S. and Ross, G. P. (1967). Studies on hybrid molecules of nucleic acids. I. DNA-DNA hybrids on nitrocellulose filters. Biochem. Biophys. Res. Comm. 28, 550-557.

Lidstrom, M. E. and Wopat, A. E. (1984). Plasmids in methanotrophic bacteria: Isolation, characterisation and DNA hybridisation analysis. Arch. Microbiol. 140, 27-33.

Lidstrom-O'Connor, M. E., Fulton, G. L. and Wopat, A. E. (1983). *Methylobacterium ethanolicum*: A synotrophic association of two methylotrophic bacteria. J. Gen. Microbiol. 129, 3139-3148.

Lutz, L. H. and Yayanos, A. A. (1985). Spectrofluorometric determination of bacterial DNA base composition. Anal. Biochem. 144, 1-5.

Lynch, M. J., Wopat, A. E. and O'Connor, M. E. (1980). Characterisation of two new facultative methanotrophs. Appl. Environ. Microbiol. 40, 400-407.

Macario, E. C. and Macario, A. J. (1983). Monoclonal antibodies for bacterial identification and taxonomy. ASM News 49, 1.

Manachini, P. L. and Parini, C. (1983). DNA restriction endonuclease patterns, DNA sequence similarity and phenotypical characteristics on some strains of *Lactobacillus helveticus* and *Lactobacillus jugurti*. Antonie van Leeuwenhoek 49, 143-152.

Mandel, M., Schildkraut, C. L. and Marmur, J. (1968). 108. Use of CsCl density gradient analysis for determining the G-C content of DNA. 109. Use of UV absorbance-temperature profile for determining G-C content of DNA. Meth. Enzymol. 12B.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). (Eds.) Molecular Cloning, Cold Spring Harbor Laboratory, USA.

Marmur, J. (1961). A procedure for the isolation of DNA from microorganisms. J. Mol. Biol. 3, 208-218.

Marmur, J. and Doty, P. (1962). Determination of the base composition of DNA from its thermal denaturation temperature. J. Mol. Biol. 5, 109-118.

Marshall, R. B., Winter, P. J. and Yanagawa, R. (1984). Restriction endonuclease DNA analysis of *Leptospira interrogans* serovars icterohaemorrhagiae and hebdomadis. J. Clin. Microbiol. 20, 808-810.

Marshall, T. and Williams, K. (1984). Artifacts associated with 2-mercaptoethanol upon high resolution 2-D electrophoresis. Anal. Biochem. 139, 502-505.

McConaughy, B. L., Laird, C. D. and McCarthy, B. J. (1969). Nucleic acid reassociation in formamide. Biochemistry 8, 3289-3295.

McNerney, T. and O'Connor, M. L. (1980). Regulation of C1 enzymes associated with C1 metabolism in 3 facultative methylotrophs. App. Environ. Microbiol. 40, 370-375.

Merabtine, H., Scott, C. D. B., Harrison, D. E. F. and Oliver, S. G. (1982). The role of pigmentation in the growth and survival of the methanol utilising bacterium, *Pseudomonas extorquens*. J. Chem. Tech. Biotech. 32, 1059-1064.

Mocca, L. and Frasch, C. E. (1982). SDS-polyacrylamide gel typing system for characterisation of *Neisseria meningitidis* isolates. J. Clin. Microbiol. 16, 240-244.

Moore, W. E. C., Hash, D. E., Holdeman, L. V. and Cato, E. P. (1980). Polyacrylamide slab electrophoresis of soluble proteins for studies of bacterial floras. Appl. Environ. Microbiol. 39, 900-907.

Mordarski, M. and Szyba, K. (1976). DNA reassociation in the classification of the 'rhodochrous' complex and allied taxa. J. Gen. Microbiol. 94, 235-245.

Nichols, J. M., Foulds, I. J., Crouch, D. H. and Carr, N. G. (1982). The diversity of cyanobacterial genomes with respect to rRNA cistrons. J. Gen. Microbiol. 128, 2739-2746.

Nishimura, Y., Shimadzu, M. and Iizuka, H. (1981). Bacteriochlorophyll formations in radiation resistant *Pseudomonas radiora*. J. Gen. Appl. Microbiol. 27, 427-430.

Nishimura, Y., Yamanaka, S. and Iizuka, H. (1983). Ubiquinone, fatty acid, polar lipid and DNA base composition in *Pseudomonas radiora*. J. Gen. Appl. Microbiol. 29, 421-427.

Ochs, D. (1983). Protein contaminants of SDS-polyacrylamide gels. Anal. Biochem. 135. 470-474.

O'Connor, M. L. (1981a). Regulation and genetics in facultative methylotrophic bacteria. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

O'Connor, M. L. (1981b). Extension of the model concerning linkage of genes coding for C1 related functions in *Methylobacterium organophilum*. Appl. Environ. Microbiol. 41, 437-441.

O'Connor, M. L. and Hanson, R. S. (1975). Serine transhydroxymethylase isoenzymes from a facultative methylotroph. J. Bact. 124. 985-996.

O'Connor, M. L. and Hanson, R. S. (1977). Enzyme regulation in *Methylobacterium organophilum*. J. Gen. Microbiol. 101. 327-332.

O'Connor, M. L. and Hanson, R. S. (1978). Linkage relationships between mutants of *Methylobacterium organophilum* impaired in their ability to grow on one carbon compounds. J. Gen. Microbiol. 104. 105-111.

O'Connor, M. L., Wopat, A. and Hanson, R. S. (1977). Genetic transformation in *Methylobacterium organophilum*. J. Gen. Microbiol. 98, 265-272.

O'Keeffe, D.T. and Anthony, C. (1980). The two cytochromes C in the facultative methylotroph *Pseudomonas* AM1. Biochem. J. 192. 411-419.

Owen, R. J. and Jackman, P. J. H. (1982). The similarities between *Pseudomonas paucimobilis* and allied bacteria derived from analysis of DNAs and electrophoretic protein patterns. J. Gen. Microbiol. 128. 2945-2954.

Owen, R. J. and Leaper, S. (1981). Base composition, size and nucleotide sequence similarities of genome DNAs from species of the genus *Campylobacter*. FEMS Microbiol. Letts. 12. 395-400.

Patel, R. N. (1984). Methane monooxygenase from *Methylobacterium* sp. strain CRL 26. In: Microbial Growth on C1 Compounds. R. L. Crawford and R. S. Hanson (Eds.), ASM, Washington.

Patel, R. N., Ching, T. H. and Felix, A. (1978). Isolation of methane utilising bacteria and characterisation of a facultative methane utilising isolate. J. Bact. 136. 352-358.

Patt, T. E., Cole, G. C., Bland, J. and Hanson, R. S. (1974). Isolation and characterisation of bacteria that grow on methane and organic compounds as sole sources of carbon and energy. J. Bact. 120. 955-964.

Patt, T. E., Cole, G. C. and Hanson, R. S. (1976). *Methylobacterium*, a new genus of facultatively methylotrophic bacteria. Int. J. Syst. Bact. 26, 226-229.

Patt, T. E. and Hanson, R.S. (1978). Intracytoplasmic membrane, phospholipid and sterol content of *Methylobacterium organophilum* cells grown under different conditions. J. Bact. 134. 636-644.

Peel, D. and Quayle, J. R. (1961). Microbial growth on C1 compounds. I. Isolation and Characterisation of *Pseudomonas* AM1. Biochem. J. 81, 465-469.

Pfennig, N. (1977). Phototrophic green and purple bacteria: A comparative systematic survey. Ann. Rev. Microbiol. 31, 275-290.

Phillips, A. P. (1969). A study of factors influencing the retention of DNA by membrane filters. Biochim. Biophys. Acta 195. 186-196.

Quayle, J. R. (1972). The metabolism of one carbon compounds in microorganisms. Adv. Microbial Physiol. 7, 119-203.

Quayle, J. R. and Ferenci, T. (1978). Evolutionary aspects of autotrophy. Microbiol. Revs. 42, 251-275.

Razin, S., Harasawa, R. and Barile, M. F. (1983). Cleavage patterns of the *Mycoplasma* chromosome, obtained by using restriction endonucleases. Int. J. Syst. Bact. 33, 201-206.

Ridell, M. and Williams, S. T. (1983). Serotaxonomical analysis of some *Streptomyces* and related organisms. J. Gen. Microbiol. 129. 2857-2861.

Rigby, P. W. J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). Labelling DNA to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113. 237-251.

Rock, J. S., Goldberg, I., Ben-Bassat, A. and Mateles, R. I. (1976). Isolation and characterisation of two methanol utilising bacteria. Ag. Biol. Chem. 40, 2129-2135.

Sahm, H., Cox, R. B. and Quayle, J. R. (1976). Metabolism of methanol by *Rhodopseudomonas acidophila*. J. Gen. Microbiol. 94, 313-322.

Salem, A. R., Hacking, A. J. and Quayle, J. R. (1973). Cleavage of malyl-CoA into acetyl-CoA and glyoxylate by *Pseudomonas* AM1 and other C1 unit utilising bacteria. Biochem. J. 136. 89-96.

Salem, A. R., Hacking, A. J. and Quayle, J. R. (1974). Lack of malyl-CoA lyase in a mutant of *Pseudomonas* AM1. J. Gen. Microbiol. 81, 525-527.

Sato, K. (1978). Bacteriochlorophyll formation by facultative methylotrophs, *Protaminobacter ruber* and *Pseudomonas* AM1. FEBS Letts. 85, 207-210.

Sato, K. and Shimizu, S. (1979). The conditions for bacteriochlorophyll formation and the ultrastructure of a methanol-utilising bacterium *Protaminobacter ruber*, classified as non-photosynthetic bacteria. Ag. Biol. Chem. 43, 1669-1675.

Schildkraut, C. L., Marmur, J. and Doty, P. (1961). The formation of hybrid DNA molecules and their use in studies of DNA homologies. J. Mol. Biol. 3, 595-617.

Schleifer, K. H., Ludwig, W., Kraus, J. and Festle, H. (1985). Cloned rRNA genes from *Pseudomonas aeruginosa* as probes for conserved DNA sequences. Int. J. Syst. Bact. 35, 231-236.

Schleifer, K. H. and Stackebrandt, E. (1983). Molecular systematics of prokaryotes. Ann. Rev. Microbiol. 37. 143-187.

Seldin, L. and Dubnau, D. (1985). DNA homology among *Bacillus polymyxa*, *B. macerans*, *B. azotofixans*, and other nitrogen fixing *Bacillus* strains. Int. J. Syst. Bact. 35, 151-154.

Shiba, T. and Simidu, U. (1982). *Erythrobacter longus* gen. nov. sp. nov., an aerobic bacterium which contains bacteriochlorophyll a. Int. J. Syst. Bact. 32, 211-217.

Shimizu, S., Sato, K., Hiraoka, M., Yamashita, F. and Kobayashi, T. (1982). Carotenoids formation by a facultative methylotroph, *Protaminobacter ruber*. J. Ferm. Tech. 60, 163-166.

Simon, R. (1984). High frequency mobilisation of gram negative bacterial replicons by the *in vitro* constructed Tn5-Mob transposon. Mol. Gen. Genet. 196. 413-420.

Sneath, P. H. A. (1972). Computer taxonomy. In: Methods in Microbiology, pp.29-98. J. Norris and D. Ribbons (Eds.), Academic Press, London.

Sneath, P. H. A. (1978). Classification of microorganisms. Essays in Microbiology 9, 9/1-9/31.

Sneath, P. H. A. and Johnson, R. (1972). The influence of numerical taxonomic similarities of errors in microbiological tests. J. Gen. Microbiol. 72, 377-392.

Sneath, P. H. A. and Sokal, R. R. (1973). Numerical Taxonomy. Freeman, San Francisco.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments seperated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Stackebrandt, E. and Kandler, O. (1979). Taxonomy of the genus *Cellulomonas*, based on phenotypic characters and DNA-DNA homology, and proposal of seven neotype strains. Int. J. Syst. Bact. 29, 273-282.

Stackebrandt, E. and Woese, C. R. (1981). The evolution of prokaryotes. In: Molecular and Cellular Aspects of Microbial Evolution. M. Carlisle, J. Collins and B. Moseley (Eds.), Cambridge Press, Cambridge.

Stackebrandt, E. and Woese, C. R. (1984). The phylogeny of prokaryotes. Microbiol. Sci. 1, 117-122.

Stackebrandt, E., Wunner-Fussl, B., Fowler, V. J. and Schleifer, K. (1981). DNA homologies and rRNA similarities among spore forming members of the order *Actinomycetales*. Int. J. Syst. Bact. 31, 420-431.

Stanley, S. and Dalton, H. (1982). Role of ribulose-1,5-bisphosphate carboxylase/oxidase in *Methylococcus capsulatus* (Bath). J. Gen. Microbiol. 128, 2927-2935.

Stocks, P. K. and McCleskey, C. S. (1964). Identity of the pink-pigmented methanol oxidising bacteria as *Vibrio extorquens*. J. Bact. 88, 1065- 1070.

Studier, F. W. (1973). Analysis of bacteriophage T7 early RNAs and proteins in slab gels. J. Mol. Biol. 79, 237-248.

Swenson, C. E., Van Hamont, J. and Dunbar, B. S. (1983). Specific differences among strains of *Ureaplasma urealyticum* as determined by 2-D gel electrophoresis and a sensitive silver stain. Int. J. Syst. Bact. 33, 417-421.

Swings, J., Kersters, K. and De Ley, J. (1976). Numerical analysis of electrophoretic protein patterns of *Zymomonas* strains. J. Gen. Microbiol. 93, 266-271.

Takamiya, K. and Okamura, K. (1984). Photochemical activities and photosynthetic ATP formation in membrane preparations from a facultative methylotroph *Protaminobacter ruber* strain NR-1. Arch. Microbiol. 140, 21-26.

Tamaoka, J. and Komagata, K. (1984). Determination of DNA base composition by reversed-phase high performance liquid chromatography. FEMS Microbiol. Letts. 25, 125-128.

Tatra, P. K. and Goodwin, P. M. (1983). R plasmid mediated chromosome mobilisation in the facultative methylotroph *Pseudomonas* AM1. J. Gen. Microbiol. 129, 2629-2632.

Tatra, P. K. and Goodwin, P. M. (1984). R factor mediated chromosome mobilisation in the facultative methylotroph *Pseudomonas* sp. strain AM1. In: Microbial Growth on C1 Compounds. R. L. Crawford and R. S. Hanson (Eds.), ASM, Washington.

Tobari, J. (1984). Blue copper proteins in electron transport in methylotrophic bacteria. In: Microbial Growth on C1 Compounds. R. L. Crawford and R. S. Hanson (Eds.), ASM, Washington.

Torres, A. R., Li, M., Ward, D. C. and Edberg, S. C. (1984). Differentiation of *Neisseria gonorrhoeae* from other *Neisseria* species by use of the restriction endonuclease Hae III. J. Clin. Microbiol. 20, 687-690.

Truper, H. (1981). Versatility of carbon metabolism in the phototrophic bacteria. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

Urakami, T. and Komagata, K. (1979). Cellular fatty acid and coenzyme Q system in gram negative methanol utilising bacteria. J. Gen. Appl. Microbiol. 25., 343-360.

Urakami, T. and Komagata, K. (1981). Electrophoretic comparison of enzymes in the gram negative methanol utilising bacteria. J. Gen. Appl. Microbiol. 27, 381-403.

Urakami, T. and Komagata, K. (1984). *Protomonas*, a new genus of facultatively methylotrophic bacteria. Int. J. Syst. Bact. 34, 188-201.

Van Landschoot, A., De Vos, P. and De Ley, J. (1984). An alternative parameter for $Tm_{(e)}$ in DNA:rRNA hybridisation studies. Syst. Appl. Microbiol. 5, 143-156.

Warner, P. J. and Higgins, I. J. (1977). Examination of obligate and facultative methylotrophs for plasmid DNA. FEMS Microbiol. Letts. 1, 339-342.

Warner, P. J., Higgins. I. J. and Drozd, J. W. (1980). Conjugative transfer of antibiotic resistance to methylotrophic bacteria. FEMS Microbiol. Letts. 7, 181-185.

Warnaar, S.O. and Cohen, J. A. (1966). A quantitative assay for DNA-DNA hybrids using membrane filters. Biochem. Biophys. Res. Comm. 24, 554-558.

Watabe, J., Benno, Y. and Mitsuoka, T. (1983). Taxonomic studies of *Bacteroides oralis* and related organisms and proposal of *Bacteroides veroralis* sp. nov. Int. J. Syst. Bact. 33, 57-64.

Weckesser, J., Mayer, H., Metz, E. and Biebl, H. (1983). Lipopolysaccharide of *Rhodocyclus purpureus*: Taxonomic implication. Int. J. Syst. Bact. 33, 53-56.

Whitta, S., Sinclair, M. I. and Holloway, B. W. (1985). Transposon mutagenesis in *Methylobacterium* AM1 (*Pseudomonas* AM1). J. Gen. Microbiol. 131, 1547-1549.

Whittenbury, R. and Dalton, H. (1981). The methylotrophic bacteria. In: The Prokaryotes. Vol. I. M. P. Starr *et al*. (Eds.), Springer Verlag, Berlin.

Whittenbury, R. and Dow, C. S. (1977). Morphogenesis and differentiation in *Rhodomicrobium vannielii* and other budding and prosthecate bacteria. Bact. Revs. 41, 754-808.

Whittenbury, R., Philips, K. C. and Wilkinson, J. F. (1970). Enrichment, isolation and some properties of methane utilising bacteria. J. Gen. Microbiol. 61, 205-218.

Woese, C. R., Maniloff, J. and Zablen, L. B. (1980). Phylogenetic analysis of the *Mycoplasmas*. Proc. Nat. Acad. Sci. USA 77, 494-498.

Wolf, H. J. and Hanson, R. S. (1978). Alcohol dehydrogenase from *Methylobacterium organophilum*. Appl. Environ. Microbiol. 36, 105-114.

Wolf, H. J. and Hanson, R. S. (1979). Isolation and characterisation of methane utilising yeasts. J. Gen. Microbiol. 114. 187-194.

Wray, W., Boulikas, T., Wray, V. P. and Hancock, R. (1981). Silver staining of proteins in polyacrylamide gels. Anal. Biochem. 118. 197-203.

Zatman, L. J. (1981). A search for patterns in methylotrophic pathways. In: Microbial Growth on C1 Compounds. H. Dalton (Ed.), Heyden, London.

COMPUTER

PROGRAMS

(a)

**JOYCE L**

```
L.
 10 REM DATA CAPTURE FORM JOYCE LOERL
 20 REM FAGE SHOULD BE SET AT &2300
 25 REM JOYCE-L 2/9/85
 26 REM  Mel Whiteside
 30 REMONERROR GOTO720
 40 DIM F%(999):DRIVE%=1:*DRIVE0
 50 a%=&1A02:b%=&1A03
 60REPEAT:MODE7
 70FRINTTAB(5,22)CHR$129:"CURRENT DRIVE = ";CHR$131;DRIVE%
 80FRINTTAB(5,9)"1. ACCUMULATE"
100FRINTTAB(5,11)"2. CHANGE DRIVE"
110ON INSTR("12",GET$) GOTO120,180 ELSE110
120 INFUTTAB(5,9)"LENGTH OF SCAN IN SECONDS ";T%
130DELAY%=INT(T%/10+.5)
140CLS:FRINTTAB(9,9)"FRESS ANY KEY TO START"
150A=GET
160MODE0:FROCREAD:MODE7
170FROCSAVE:UNTIL FALSE
180IF DRIVE%=1 DRIVE%=2 ELSE DRIVE%=1
190ON DRIVE% GOTO200,220
200*DRIVE0
210UNTIL FALSE
220*DRIVE2
230UNTIL FALSE
270DEF FROCREAD
280MOVE1280,0
290REFEAT UNTIL ADVAL(0)=1
300I%=0
310 TIME=0
320REFEAT
330N%=ADVAL(1)/16:FROC_FOKE
340 DRAW1280-I%,N%
350K%=K%+1
360I%=I%+1
370REFEAT UNTIL TIME>=DELAY%
380TIME=0
390UNTIL I%=999
400ENDFROC
410DEF FROCSAVE
420REFEAT:CLOSE£0
430CLS:INFUTTAB(9,9)"FILENAME ";FILENAME$
440X=OFENIN(FILENAME$)
450UNTIL X=0
460CLOSE£X
470X=OFENOUT(FILENAME$)
480 OSCLI"*SA."+FILENAME$+" 1A00 +800"
490CLOSE£X
500ENDFROC
510DEF FROCRECAL
520REFEAT
530CLS:INFUTTAB(9,9)"FILENAME ";FILENAME$
540X=OFENIN(FILENAME$)
550UNTIL X<>0
560INFUT£X,A$
570FRINTTAB(2,5)A$
580FRINTTAB(9,15)"FRESS ANY KEY TO CONTINUE"
590A=GET
600FOR I%=0 TO 999
610F%(I%)=0
620INFUT£X,F%(I%)
630NEXT
640ENDFROC
650DEF FROCFLOT
660MOVE1280,0
670FOR I%=0 TO 499
680X%=1280-2*I%
690 Y%=(F%(I%*2)+F%((I%*2)+1))/8
700DRAWX%,Y%
710NEXT:ENDFROC
720CLOSE£0:FRINTERR,ERL:REFORT
725 END
730 FROC_FOKE
740 ?a%=N% MOD 256
750 ?b%=N% DIV 256
760 ENDFROC
```

(b)

**COMP 4**

```
10 REM COMPARE [illegible]
20REM COMPARES AND PRINTS OUT TRACES ON PLOTTER WITH COMPARISON TRACE - BACKGR
OUND OPTION - 50% BACKGROUND SUBTRACTION
30DIMO%(20),FLN$(3),M 999,D 999,HI%(2),TW%(2),NO%(3),DD%(2),[illegible]
40MODE7:CLOSE#0:NT=FALSE
50PROCINITIATE
60*FX5,0
70PROCINPUT:PROCPRINTINIT
80MODE4:STM%=900:V%=[illegible]70=7
90REM
100REM SET SEARCH PARAMETERS
110MARGIN=[illegible]
120 FOR J%=0 TO 1
130[illegible]ZA=HI%(J%)MOD256:[illegible]=HI%(J[illegible])
140[illegible]ZC=TN%(J%)
150MODE4:PROCPRINTINIT
160PROCGENTRACE
170PROCTRANS:PROCSCREENTRACE(500,M)
180PROCFINDBAND:PROCLOCATE
190PROCCOMPAND:IF NT GOTO220
200CLS:PROCPRINTOUT
210N%=0:PROCPRINTTRACE(0):PROCPRINT(0)
220NT=TRUE:NEXT J%:PROCNOP:PROCCOMPARE
230J%=1
240V%=1:PROCPLOT
250PROCINFO:END
260DEF PROCAXIS
270MOVE 2[illegible],0
280FOR I%=280 TO 1280 STEP 50
290DRAW I%,0:DRAWI%,24:MOVEI%,0
300NEXT:ENDPROC
310DEF PROCBACKGROUND
320LOCAL J%:CLS:PRINTTAB(5,9)"BACKGROUND SUBTRACTION"
330FOR I%=0 TO 19
340ST%=0:[illegible]=I%*50
350FOR J%=0 TO 49
360ST%=ST%+FNP(Q%+J%)
370NEXT J%:O%(I%)=ST%/50
380NEXTI%
390ENDPROC
400DEF PROCTS
410 TP%=FNP(0)
420P%=0
430FOR I%=1 TO 250
440IF FNP(I%) TP% P%=I%:TP%=FNP(I%)
450NEXT:ENDPROC
460DEF PROCDRAWBG
470REM DRAW BACKGROUND
480MOVE280,0:DRAW1280,0
490FOR I%=0 TO 19
500DRAW 280+I%*50,O%(I%)/20
510NEXTI%:ENDPROC
520DEF PROCLOSS2
530LOCAL I%
540FOR I%=0 TO 19
550REM CALCULATE SLOPE
560SLOPE=(O%(I%+1)-O%(I%))/50
570M%=I%*50
580REM SUBTRACT BACKGROUND
590FOR J%=0 TO 49:N% J%+M%
600P%=FNP(N%)-(SLOPE*J%+O%(I%))[illegible]
605PROCPOKE(N%,P%)
610NEXT J%,I%:PROCDRAWBG
620ENDPROC
630DEF PROCMAX
640NMAX%=0
650FOR I%=0 TO 19
660NMAX%=NMAX%+O%(I%)
670NEXT:NMAX%=NMAX%/20:ENDPROC
680DEF PROCTRANS
690FOR I%=0 TO 999
700IF ?(&1907+I%*2)AND?[illegible] M(I%) [illegible]
710?(&1907+I%*2)=?(&1907+I%*2)[illegible]
720NEXT:ENDPROC
730DEF PROCINITIATE
740FOR I%=&1900 TO &[illegible]D4 STEP 4
750!I%=0:NEXT
760FOR I%=0 TO 999 STEP 4
770M!I%=0:IF NOT NT O!I%=0
780NEXT:ENDPROC
790DEF PROCINPUT
800
810[illegible]%=0
820FOR I%=0 TO 1:REPEAT:CLS
830INPUTTAB(7,9)"SPECIFY DRIVE [illegible]":DR%(I%)
840PROCDRIVE(I%)
850CLS:INPUTTAB([illegible],9)"FILENAME ":FILENAME$
860[illegible]OPENIN(FILENAME$)
870IF X 0 PRINTTAB(5,1[illegible])"FILE NOT [illegible]":A=GET
880UNTIL X 0:CLS
890PRINTTAB([illegible])"BACKGROUND SUBTRACTION"
900PRINTTAB(7,11)"[illegible]"
910[illegible]GET[illegible]
920IF[illegible]%(I%)[illegible]GOTO 8[illegible]
930[illegible]%(I%)=[illegible]
940CLOSE#X:FLN$(I%)=FILENAME$
950[illegible](I%)
960NEXT:ENDPROC
```

```
930@%=&20,0
940VDU2,1,18,1,13,1,65,1,13,1,18,1,13
950PROCdraw("M",50,-150)
960VDU1,73,1,13 :REM SET ORIGIN
970VDU3:ENDPROC
980DEF PROCPRINTTRACE(U)
990X%=V%*100+50:VDU2,1,18,1,13
1000J%=E1%-1
1010VDU1,81,1,49,1,13:REM WRITE T-R
1020VDU1,67,1,ASC(STR$(J%)),1,13:REM COLOUR CHANGE
1030PROCdraw("M",X%,100):REM ABOVE TRACE
1040PROClabel(FLN$(J%))
1050PROCdraw("M",X%,0):REM TOP OF TRACE
1060ENDPROC
1070DEFPROCPRINT(U)
1080FOR I%=0 TO 999:Y%=-I%
1090IF ?(U+I%)=&F PROCBAND
1100NEXTI%
1110VDU1,72,1,13:REM HOME
1120VDU3:ENDPROC
1130DEFPROCdraw(D$,X%,Y%)
1140LOCAL I%
1150D$=D$+STR$(X%)+","+STR$(Y%)
1160FORI%=1TOLEN(D$)
1170VDU1,ASC(MID$(D$,I%,1))
1180NEXT:VDU1,13
1190ENDPROC
1200DEFPROClabel(F$)
1210LOCAL I%:L$="P"+D$
1220FORI%=1TOLEN(L$)
1230VDU1,ASC(MID$(L$,I%,1))
1240NEXT:VDU1,13
1250ENDPROC
1260DEF PROCLOCATE
1270SP%=890:REM END OF SEARCH
1280MHT%=1000:REM MIN HEIGHT
1290PP%=SP%:PH%=MHT%
1300FOR I%=999 TO SP% STEP -1
1310IFFNP(I%)>PH% PP%=I%:PH%=FNP(I%)
1320NEXT:B2%=PP%
1330PROCLABEL(PP%):ENDPROC
1340DEF PROCBAND
1350 IF U=M H%=I%:GOTO 1410
1360FLAG=FALSE
1370REPEAT:H%=H%+1
1380IF ?(M+H%)=&F FLAG=TRUE
1390UNTIL FLAG
1400H%=H%+1
1410SIZE%=(FNP(H%)/400)*8+.5
1420PROCdraw("M",X%-SIZE%/2,Y%)
1430PROCdraw("D",X%+SIZE%,Y%)
1440ENDPROC
1450DEF PROCGENTRACE
1460PROCINITIATE
1470PROCDELETE(J%)
1480OSCLI "*LOAD "+FLN$(J%)
1485IF BS%(J%) PROCBACKGROUND:PROCSET...:A=GET
1490NOF%(J%)=USR(&2200)AND &FFFF
1500PRINTTAB(9,9)"NO. OF PEAKS ";NOF%(J%)
1510PROCAXIS
1520ENDPROC
1530DEF PROCBANDS
1540SIZE%=(FNP(I%)/400)*8+.5
1550X%=1278-I%
1560Y%=V%-SIZE%/2
1570MOVE X%,Y%
1580DRAW X%,Y%+SIZE%
1590ENDPROC
1600DEF PROCSCREENTRACE(V%,U)
1610FOR I%=0 TO 999
1620IF ?(U+I%)=&F PROCBANDS
1630NEXT:ENDPROC
1640DEF PROCLABEL(MF%)
1650X%=1278-MF%
1660MOVE X%,1000
1670DRAWX%,940
1680DRAWX%+20,1000
1690DRAWX%-20,1000
1700DRAWX%,940:ENDPROC
1710DEF PROCFIRSTBAND
1720PP%=20:PH%=1000
1730FOR I%=20 TO 60
1740IF FNP(I%)>PH% PP%=I%:PH%=FNP(I%)
1750NEXT:B1%=PP%
1760PROCLABEL(B1%):ENDPROC
1770DEF PROCCOMBAND
1780FF%=PP%-B1%:F=SIM%/FF%
1790FOR I%=B1% TO 999
1800IF ?(M+I%)=&F PROCCOMPCALC ELSE 1870
1810IF I%<999 AND NOT ?(U+NF%)=?(U+NF%+1)... :GOTO1870
1820IF I%>999 ?(U+NF%)=&F
1830NEXT:ENDPROC
1840DEF FNP(F%)
1850=!(&1900+F%*2) AND &FFFF
1860DEF PROCPRINTOUT
1870@%=&20,0
1880VDU2,1,18,1,13,1,65,1,13,1,1 ,1,1
1890PROCdraw("M",50,-100)
1900VDU1,73,1,13
```

```
1910PROCdraw("Y1,",40,10)
1920VDU1,72,1,13
1930PROCdraw("X0,",-50,20)
1940FOR I%=0 TO 999
1950X%=FNP(I%)/10
1960Y%=-999+I%
1970PROCdraw("D",X%,Y%)
1980IFI%=B1%ORI%=B2% PROCdraw("D",X%+50,Y%):PROClabel("X")
1990NEXTI%
2000PROCdraw("M",0,-999)
2010VDU1,18,1,13,1,1,65,1,17,1,17,1,13
2020PRINT"FILENAME              ";FUNF(Z%)
2030PRINT"NO. OF PEAKS          ";NOP%
2040PRINT"HEIGHT THRESHOLD  ";HT%
2050PRINT"WIDTH THRESHOLD   ";TW%
2060VDU3:ENDPROC
2070DEF PROCCOMPCALC
2080NP%=INT(F*(I%-B1%)+.5)
2090IF NP%>1000 I%=999
2100ENDPROC
2110DEF PROCCOMPARE
2120FOR I%=0 TO 999
2130IF?(D+I%)AND&F PROCSEARCH
2140NEXT:ENDPROC
2150DEF PROCSEARCH
2160START%=I%-MARGIN:NR%=0
2170IF I%<MARGIN START%=0
2180FOR K%=START% TO I%+MARGIN
2190IF?(D+K%)AND&80 PROCNEAR
2200NEXTK%
2210IF NR%=0 ENDPROC
2220IF NOT?(D+NR%)AND&40 ?(D+NR%)=?(D+NR%)OR &40
2230ENDPROC
2240DEF PROCPCT
2250PROCPRINTTRACE(D)
2260FOR I%=0 TO 999:Y%=-I%
2270IF ?(D+I%)AND&80 PROCBAND
2280NEXTI%
2290VDU1,72,1,13
2300PROCMARKER
2310ENDPROC
2320DEF PROCMARKER
2330 X%=400:N%=0
2340PROCdraw("M",X%,0)
2350FORI%=0 TO 999
2360IF?(D+I%)AND&40 PROCMARK
2370NEXTI%
2380VDU3:ENDPROC
2390DEF PROCMARK
2400N%=N%+1
2410Y%=-I%
2420PROCdraw("M",X%,Y%)
2430PROCdraw("D",X%+40,Y%)
2440ENDPROC
2450DEF PROCINFO
2460VDU2,1,17,1,17
2470PRINT:PRINT:PRINT
2480PRINT"NUMBER OF SIMILAR BANDS ";N%
2490PRINT
2500 PRINT"NO OF PEAKS IN TRACE 1 ";NOP%(0)
2510PRINT"NO OF PEAKS IN TRACE 2 ";NOP%(1)
2520PRINT
2530PRINT;SPC(20);"TRACE 1";SPC(4);"TRACE 2"
2540PRINT"WIDTH THRESHOLD";SPC(5);TW%(0);SPC(9);TW%(1)
2550 PRINT"HEIGHT THRESHOLD";SPC(4);HT%(0);SPC(8);HT%(1)
2555 PRINT"BACKGROUND SUB. ";SPC(4);BS%(0);SPC(9);BS%(1)
2560PRINT"MARGIN = ";MARGIN
2570VDU3:ENDPROC
2580DEFPROCREA(N%)
2590CLS:REPEAT:PRINTTAB(28,9)"
2600 PRINTTAB(5,9)CHR$130;"PEAK HEIGHT THRESHOLD ";:INPUT;HT%(N%)
2610UNTIL HT%(N%)>0 AND HT%(N%)<4096
2620CLS:REPEAT:PRINTTAB(27,9)"
2630 PRINTTAB(5,9)CHR$130;"PEAK WIDTH THRESHOLD ";:INPUT;TW%(N%)
2640UNTIL TW%(N%)>0 AND TW%(N%)<1000
2650ENDPROC
2660DEF PROCNEAR
2670IFABS(I%-K%)<ABS(I%-NR%) NR%=K%
2680ENDPROC
2690DEF PROCNOP
2700 NOP%(0)=0:NOP%(1)=0
2710FOR I%=0 TO 999
2720IF ?(D+I%) AND &F NOP%(0)=NOP%(0)+1
2730IF ?(D+I%) AND &80 NOP%(1)=NOP%(1)+1
2740NEXT:ENDPROC
2750DEF PROCDRIVE(Z%)
2760ON DD%(Z%)+1 GOTO2770,2790,2810,2830
2770*DR.0
2780ENDPROC
2790*DR.1
2800ENDPROC
2810*DR.2
2820ENDPROC
2830*DR.3
2840ENDPROC
3000DEF PROCPOKE(C%,A%)
3010?(&1500+2*C%)=A% MOD 256
3020?(&1900+2*C%)=A% DIV 256
```

(c)

**AUTO CA**

```
>L.
   10 REM AUTOCA 3/5/84
   15 REM PAGE=&2400:*LO. MC
   20REM COMPANDS AND PRINTS OUT TRACKS ON PLOTTER WITH COMPARISON TRACK- BACKGR
OUND OPTION - 50% BACKGROUND SUBTRACTION  STORES COEFF FILE FROM CF1 DATA
   30DIMQ%(20),F$(1),M 999,O 999,NOF%(2),BS%(2),CAT$(32),BUFFER%255,PARAM%20,D%(
1),C$(64),D 64,a% 2025
   40MODE7:CLOSE#0
   60*FX5,1
   70PROCINPUT:PROCINFO1
   80MODE4:STM%=900:V%=0:?&7C=3
   90?&7A=HT%MOD256:?&7B=HT%DIV256:?&7C=TW%
   96 FOR I%=0 TO 2025:I%!a%=0:NEXT
  100PRINT"WHATS ON":PROCCAT
  110FOR i%=1TOL%-1:J%=0
  120F$(0)=C$(i%):D%(0)=?(D+i%)
  125PROCLOOP
  126FOR j%=1+i%TOL%:F$(1)=C$(j%):D%(1)=?(D+j%):J%=1
  127 IF j%>L% GOTO130
  128PROCLOOP
  129PROCCOMPARE:PROCMARKER:PROCINFO
  130NEXTj%,i%:VDU2:PROCPRINT:VDU3:END
  140DEFPROCLOOP
  145N%=0:PROCINITIATE
  150 IF J%=1 FOR M%=0 TO 999:O?M%=O?M% AND &7F:NEXT
  160PROCGENTRACK
  170PROCTRANS
  180PROCFIRSTBAND:PROCLOCATE
  190PROCCOMPAND
  200CLS
  220 PROCNOP
  250ENDPROC
  310DEF PROCBACKGROUND
  320LOCAL J%:CLS:PRINTTAB(5,9)"BACKGROUND SUBTRACTION"
  330FOR I%=0 TO 19
  340ST%=0:O%=I%*50
  350FOR J%=0 TO 49
  360ST%=ST%+FNP(O%+J%)
  370NEXTJ%:Q%(I%)=ST%/50
  380NEXTI%
  390ENDPROC
  400DEF PROCTS
  410 TR%=FNP(0)
  420PP%=0
  430FORI%=1 TO 998
  440IF FNP(I%)>TR% PP%=I%:TR%=FNP(I%)
  450NEXT:ENDPROC
  460DEF PROCDRAWBG
  470REM DRAW BACKGROUND
  480MOVE280,0:DRAW1280,0
  490FOR I%=0 TO 19
  500DRAW 1280-I%*50,Q%(I%)/8
  510NEXTI%:ENDPROC
  520DEF PROCBGS2
  530LOCAL J%
  540FOR I%=0 TO 19
  550REM CALCULATE SLOPE
  560SLOPE=(Q%(I%+1)-Q%(I%))/50
  570M%=I%*50
  580REM SUBTRACT BACKGROUND
  590FOR J%=0 TO 49:N%=J%+M%
  600P%=FNP(N%)-(SLOPE*J%+Q%(I%))/2
  605PROCPOKE(N%,P%)
  610NEXT J%,I%:PROCDRAWBG
  620ENDPROC
  630DEFPROCNMAX
  640NMAX%=0
  650FORI%=0 TO 19
  660NMAX%=NMAX%+Q%(I%)
  670NEXT:NMAX%=NMAX%/20:ENDPROC
  680DEF PROCTRANS
  690FOR I%=0 TO 999
  700IF ?(&1903+I%*2)AND&80 ?(M+I%)=&F
  710?(&1903+I%*2)=?(&1903+I%*2)AND&7F
  720NEXT:ENDPROC
  730DEF PROCINITIATE
  740FORI%=&1900 TO &20D4 STEP 4
  750!I%=0:NEXT
  760FOR I%=0 TO 999 STEP 4
  770M!I%=0:IF J%=0 O!I%=0
  780NEXT:ENDPROC
  790DEF PROCINPUT
  810BS%=0:CLS
  830INPUTTAB(3,9)"SPECIFY DRIVE(0-3) ";DD%
  840PROCDRIVE(DD%)
  850CLS:INPUTTAB(3,9)"MARGIN ";MARGIN
  882PRINTTAB(3,9)"BACKGROUND SUBTRACTION"
  884PRINTTAB(8,11)"1.ON - 2.OFF"
  885ON INSTR("12",GET$) GOTO887,889 ELSE885
  887BS%=1:GOTO 890
  889BS%=0
  900PROCPARA
  910ENDPROC
 1260DEF PROCLOCATE
 1270SP%=890:REM END OF SEARCH
 1280MHT%=1000:REM MIN HEIGHT
 1290FP%=SP%:FH%=MHT%
 1300FOR I%=999 TO SP% STEP -1
```

```
1310IFFNP(I%) FH% PF%=I%:FH%=FNP(I%)
1320NEXT:B2%=PF%
1330PROCLABEL(PF%):ENDPROC
1450DEF PROCGENTRACK
1470PROCDRIVE(D%(J%))
1480OSCLI "*LOAD "+F$(J%)
1485IF BS% PROCBACKGROUND:PROCBGS2:A=GET
1490NOP%(J%)=USR(&2200)AND &FFFF
1520ENDPROC
1530DEF PROCBANDS
1540SIZE%=(FNP(I%)/400)*8+.5
1550X%=1279-I%
1560Y%=V%-SIZE%/2
1570MOVEX%,Y%
1590DRAW X%,Y%+SIZE%
1590ENDPROC
1600DEF PROCSCREENTRACK(V%,U)
1610FOR I%=0 TO 999
1620IF ?(U+I%)=&F PROCBANDS
1630NEXT:ENDPROC
1640DEF PROCLABEL(MR%)
1650X%=1278-MR%
1660MOVEX%,1020
1670DRAWX%,960
1680DRAWX%+20,1000
1690DRAWX%-20,1000
1700DRAWX%,960:ENDPROC
1710DEF PROCFIRSTBAND
1720PF%=20:FH%=1000
1730FOR I%=20 TO 60
1740IF FNP(I%) FH% PF%=I%:FH%=FNP(I%)
1750NEXT:B1%=PF%
1760PROCLABEL(B1%):ENDPROC
1770DEF PROCCOMPAND
1780PF%=PF%-B1%:F=STM%/PF%
1790FOR I%=B1% TO 999
1800IF ?(M+I%)=&F PROCCOMPCALC ELSE 1830
1810IFI% 999 AND J%=1 ?(O+NP%)=?(O+NP%)OR&80:GOTO1830
1820IF I% 999 ?(O+NP%)=&F
1830NEXT:ENDPROC
1840DEF FNP(P%)
1850=!(&1902+P%*2) AND &FFF
2070DEF PROCCOMPCALC
2080NP%=INT(F*(I%-B1%)+.5)
2090IF NP%>=1000 I%=999
2100ENDPROC
2110DEF PROCCOMPARE
2120FOR I%=0 TO 999
2130IF?(O+I%)AND&F PROCSEARCH
2140NEXT:ENDPROC
2150DEF PROCSEARCH
2160START%=I%-MARGIN:NR%=0
2170IF I% MARGIN START%=0
2180FOR K%=START% TO I%+MARGIN
2190IF?(O+K%)AND&80 PROCNEAR
2200NEXTK%
2210IF NR%=0 ENDPROC
2220IF NOT?(O+NR%)AND&40 ?(O+NP%)=?(O+NP%)OR &40
2230ENDPROC
2240DEF PROCPCT
2250PROCPRINTTRACK(O)
2260FOR I%=0 TO 999:Y%=-I%
2270IF ?(O+I%)AND&80 PROCBAND
2280NEXTI%
2290VDU1,72,1,13
2300PROCMARKER
2310ENDPROC
2320DEF PROCMARKER
2330 N%=0
2350FORI%=0 TO 999
2360IF?(O+I%)AND&40 N%=N%+1
2370NEXTI%
2380ENDPROC
2450DEFPROCINFO
2460VDU2
2470CF1=((2*(N%-1))/(NOP%(0)+NOP%(1)-4))*100
2473CF1=INT(CF1*10+.5)/10
2480IF NOP%(0) NOP%(1) I%=NOP%(0) ELSE I%=NOP%(1)
2490CF2=((2*(N%-1))/(I%*2-4))*100
2500CF2=INT(CF2*10+.5)/10
2510PRINT'"TRACK 1,2 ";F$(0);"  ";F$(1);"  NOP 1,2  ";NOP%(0);"  ";NOP%(1);"  NO
S ";N%;"  CF1 ";CF1;"  CF2 ";CF2;
2515?(L%*i%+j%+a%)=CF1
2520VDU3:ENDPROC
2525DEFPROCINFO1:VDU2:PRINT
2540PRINT"WIDTH THRESHOLD";SPC(5):TW%;SPC(9);TW%
2550 PRINT"HEIGHT THRESHOLD";SPC(1);HT%
2555 PRINT"BACKGROUND SUB. ";SPC(4);BS%
2560PRINT"MARGIN = ";MARGIN
2570VDU3:ENDPROC
2580DEFPROCPARA
2590CLS:REPEAT:PRINTTAB(28,9)"
2600 PRINTTAB(5,9)CHR$134;"PEAK HEIGHT THRESHOLD ";:INPUT;HT%
2610UNTIL HT% 0 AND HT% 4096
2620CLS:REPEAT:PRINTTAB(27,9)"
2630 PRINTTAB(5,9)CHR$134;"PEAK WIDTH THRESHOLD ";:INPUT;TW%
2640UNTIL TW% 0 AND TW% 1000
2650ENDPROC
```

```
2660DEF PROCNEAR
2670IFABS(I%-F%)<ABS(I%-NF%) NF%=I%
2680ENDPROC
2690DEF PROCNOF
2700 NOF%(0)=0:NOF%(1)=0
2710FOR I%=0 TO SIM%
2720IF ?(D+I%) AND &F NOF%(0)=NOF%(0)+1
2730IF ?(D+I%) AND &80 NOF%(1)=NOF%(1)+1
2740NEXT:ENDPROC
2750DEF PROCDRIVE(Z%)
2760OSCLI"*DR."+STR$(Z%)
2770ENDPROC
3000DEF PROCPOKE(C%,A%)
3010?(&1902+2*C%)= A% MOD 256
3020?(&1903+2*C%)= A% DIV 256
3030ENDPROC
5000DEF PROCPRINT
5010VDU1,27,1,15
5012PRINT"    ";
5015FORI%=1 TO L%:A$=STR$(I%):PROCPAD:NEXT:PRINT
5020FOR J%=1 TO L%
5025A$=STR$(J%):PROCPAD
5030FOR I%=1 TO J%
5040 A$=STR$(?(a%+J%+(I%*L%))):PROCPAD
5055 NEXT:PRINT:NEXT
5056FORI%=0 TO L%:A$=STR$(I%):PROCPAD:NEXT:PRINT
5060 B%=&1902:FORI%=0 TO2000 STEP 4
5065 I%!B%=I%!a%:NEXT
5080 *DR.2
5085 OSCLI"*SA. COEFF 1902 +7D0"
5090 ENDPROC
5100DEF PROCPAD
5120IF LEN(A$)=0 A$="    ":GOTO5160
5130IF LEN(A$)=1 A$=A$+"   ":GOTO5160
5140IF LEN(A$)=2 A$=A$+" ":GOTO5160
5160PRINTA$;:ENDPROC
10110REM
10120REM
10130NEXTI%
10140DEFFNcat(DRIVE%,BUFFER%,X%)
10150LOCALY%,L%,FILE%,T%,A$,DIF%
10160Y%=X%DIV256
10170A%=&7F
10180?X%=DRIVE%
10190X%!1=BUFFER%
10200X%?5=3
10210X%?6=&53
10220X%?7=0
10230X%?8=1
10240X%?9=&21
10250X%?10=0
10260CALL&FFF1
10270IFX%?10L%=-1:=L%
10280L%=(BUFFER%?5)DIV8
10290IFL%=0THEN=0
10300X%?8=0
10310X%?10=0
10320CALL&FFF1
10330IFX%?10L%=-1:=L%
10340FORFILE%=1TOL%
10350DIR%=?(BUFFER%+FILE%*8+7)
10360A$=""
10370FORT%=0TO6
10380A$=A$+CHR$(?(BUFFER%+FILE%*8+T%))
10390NEXTT%
10400CAT$(FILE%)=A$
10410NEXTFILE%
10420=L%
10430DEFPROCCAT
10440L%=FNcat(DD%,BUFFER%,PARAM%)
10450IFL%=-1PRINTTAB(2,9)"PLEASE CHECK DISC/DRIVE"'"    THEN HIT A KEY":A=GET:GOT
O10440
10455IFDD%=2 GOTO10490
10460IFL%=0PRINT"DISC EMPTY"
10470FORF%=1TOL%:C$(F%)=CAT$(F%):?(D+F%)=DIF%:NEXT
10480LI%=L%+1:DD%=2:GOTO10440
10490LI%=LI%:FORF%=LI%TOLI%+L%-1
10500C$(F%)=CAT$(F%-LI%+1):?(D+F%)=DD%:NEXT
10510L%=L%+LI%-1
10520FORI%=1 TO L%
```

(d)

AUTO EF

```
   20 REM AUTOEF 29/1/84
   25REM MUST *LO."TRNSOBJ"
   30 REM PAGE SET TO &2000
   40REM COMPANDS AND COMPARES EVERY TRACE ON A DISC - GIVES COEFS OF SIMILARITY
BASED ON ALL POINTS + 1% SHUFFLE
   50D1%=100
   60DIMO%(20),F$(1),NOF%(2),BS%(2),CAT$(72),BUFFER%255,PARAM%20,D%(1),C$(64),D
64,A%(D1%),B%(D1%),a% 2025
   70MODE7:CLOSE#0
   80*FX5,1
   90PROCINPUT
  100STM%=900:V%=0:?%20=7
  110 D2%=STM%/D1%:D3%=D2%-1
  120 FOR I%=0 TO 2020:I%!a%=0:NEXT
  130PRINT"WHATS ON":PROCCAT:CLS
  140TIME=0
  150FOR i%=1TOL%-1:J%=0
  160F$(0)=C$(i%):D%(0)=?(D+i%)
  170PROCLOOP:Xm%=FNMEAN+0.5
  180PROCXG
  190FOR j%=1+i%TOL%:F$(1)=C$(j%):J%=J%+1
  200IF j%=L% GOTO270
  210D%(1)=?(D+j%)
  220PROCLOOP:Xm%=FNMEAN+0.5
  230PROCXT
  240 PROCCOEFCAL(100,0,0):P1%=R%
  250 PROCCOEFCAL(97,1,0):P2%=R%
  260 PROCCOEFCAL(99,0,1):P3%=R%
  270NEXTj%,i%:VDU2:PROCPRINT:VDU3
  280PRINT"RUN TIME ";TIME/100,"SECS":END
  290DEFPROCLOOP
  300N%=0
  310PROCGENTRACE
  320PROCFIRSTBAND:PROCLOCATE
  330 PROCCOMPAND:PRINTF
  340 ENDPROC
  400DEF PROCBACKGROUND
  410LOCAL J%:CLS:PRINTTAB(5,9)"BACKGROUND SUBTRACTION"
  420FOR I%=0 TO 19
  430ST%=0:O%=I%*50
  440FOR J%=0 TO 49
  450ST%=ST%+FNP(O%+J%)
  460NEXTJ%:B%(I%)=ST%/50
  470NEXTI%
  480ENDPROC
  490DEF PROCTS
  500 TR%=FNP(0)
  510PP%=0
  520FORI%=1 TO 998
  530IF FNP(I%)>TR% PP%=I%:TR%=FNP(I%)
  540NEXT:ENDPROC
  550DEF PROCDRAWBG
  560REM DRAW BACKGROUND
  570MOVE280,0:DRAW1280,0
  580FOR I%=0 TO 19
  590DRAW 1280-I%*50,B%(I%)/8
  600NEXTI%:ENDPROC
  610DEF PROCPB2
  620VDU2:PRINT,I%,"  ";F$(I%);"  ";?(D+I%),:NEXT:VDU3:ENDPROC
  630FOR I%=0 TO 19
  640REM CALCULATE SLOPE
  650SLOPE=(B%(I%+1)-B%(I%))/50
  660M%=I%*50
  670REM SUBTRACT BACKGROUND
  680FOR J%=0 TO 49:N%=J%+M%
  690P%=FNP(N%)-(SLOPE*J%+B%(I%))/2
  700PROCPOKE(N%,P%)
  710NEXT J%,I%:PROCDRAWBG
  720ENDPROC
  730DEFFNNMAX
  740NMAX%=0
  750FORI%=0 TO 19
  760NMAX%=NMAX%+B%(I%)
  770NEXT:NMAX%=NMAX%/20:ENDPROC
  780DEF PROCINITIATE
  790FOR I%=&1900 TO &7D04 STEP 4
  800!I%=0:NEXT
  810 FOR I%=&2400 TO &2EF8 STEP 4
  820 !I%=0:NEXT
  830 ENDPROC
  840DEF PROCINPUT
  850CLS:DR%=0
  860PROCDRIVE(DR%)
  870REMPRINTTAB(3,7)"BACKGROUND SUBTRACTION"
  880ENDPROC
  890DEF PROCLOCATE
  900SF%=850:REM END OF SEARCH
  910MHT%=1000:REM MIN HEIGHT
  920PP%=SF%:PH%=MHT%
  930FOR I%=999 TO SF% STEP -1
  940IFFNP(I%)>PH% PP%=I%:PH%=FNP(I%)
  950NEXT:D2%=PP%
```

```
 960ENDPROC
 970DEF PROCGENTRACK
 980PROCINITIATE
 990PROCDRIVE(D%(J%))
1000OSCLI "*LOAD "+F$(J%)
1010PRINT'F$(J%)
1020ENDPROC
1100DEF PROCFIRSTBAND
1110FF%=20:FH%=1000
1120FOR I%=20 TO 60
1130IF FNF(I%)<FH% FF%=I%:FH%=FNF(I%)
1140NEXT:B1%=FF%
1150ENDPROC
1160DEF PROCCOMPAND
1170FF%=FF%-B1%:F=STM%/FF%
1180FOR I%=B1% TO 999
1190 NF%=F*(I%-B1%)+.5:IF NF%>1000 I%=999
1201U%=&2402+NF%*2:W%=&1902+I%*2
1202?U%=?W%:U%?1=W%?1
1204
1210NEXT:PROCTRANSFER:ENDPROC
1220DEF FNF(F%)
1230=!(&1902+F%*2) AND &FFF
1240DEF PROCCOMPCALC
1250NF%=INT(F*(I%-B1%)+.5)
1260IF NF%>1000 I%=999
1270ENDPROC
1280DEFPROCINFO
1290VDU2
1300IFB1%<B2% N%=B1% ELSE N%=B2%
1310IFB% N% N%=B%
1320PRINT'F$(0);"  ";F$(1);"  CF1 ";B1%;"  ";B2%;"  ";R%;" = ";N%;
1330 ?(L%*1%+J%+a%)=N%
1340VDU3:ENDPROC
1350DEF PROCDRIVE(Z%)
1360OSCLI"*DR."+STR$(Z%)
1370ENDPROC
1380DEF PROCPOLE(C%,A%)
1390?(&1902+2*C%)= A% MOD 256
1400?(&1903+2*C%)= A% DIV 256
1410ENDPROC
1420DEF FNMEAN
1430 S%=0:FORI%=0 TO STM%
1440S%=S%+FNF(I%)/4:NEXT
1450=S%/STM%
1460DEF PROCXA
1465 LOCAL X%
1470FORI%=0TOD1%-1
1480S%=0:X%=I%*D2%
1490FORF%=0TOD3%
1500S%=S%+FNF(X%+F%)/4
1510 NEXT:A%(I%)=(S%/D2%)-Xm%
1520NEXT:ENDPROC
1530DEF PROCXB
1535 LOCAL X%
1540FORI%=0TOD1%-1
1550S%=0:X%=I%*D2%:FORF%=0TOD3%
1560S%=S%+FNF(X%+F%)/4
1570NEXT:B%(I%)=(S%/D2%)-Xm%
1580NEXT:ENDPROC
1590DEF PROCCORRCAL(V1%,V2%,V3%)
1600REM CALC SUM OF X4-X4m*X5-X5m
1610 E1%=0:E2%=0:E3%=0
1620 FORI%=0TOV1%-1
1630 E2%=E2%+A%(I%+V2%)^2
1640 E3%=E3%+B%(I%+V3%)^2
1650E1%=E1%+A%(I%+V2%)*B%(I%+V3%)
1660 NEXT
1670R%=(E1%/SQR(E2%*E3%)*100)+.5
1680ENDPROC
1690DEF PROCPRINT
1700VDU1,27,1,15
1710PRINT"   ";
1720FOR I%=1 TO L%:A$=STR$(I%):PROCPAD:NEXT:PRINT
1730FOR J%=1 TO L%
1740A$=STR$(J%):PROCPAD
1750FOR I%=1 TO J%
1760 A$=STR$(?(a%+J%+(I%*L%))):PROCPAD
1770 NEXT:PRINT:NEXT
1780FOR I%=0 TO L%:A$=STR$(I%):PROCPAD:NEXT:PRINT
1790 P%=&1902:FOR I%=0 TO2000 STEP 4
1800 I%!P%=I%!a%:NEXT
1810 *DR.2
1820 OSCLI"*SA. CORRS 1902 +2000"
1830 ENDPROC
1840DEF PROCPAD
1850IF LEN(A$)=0 A$="   ":GOTO1880
1860IF LEN(A$)=1 A$=A$+"  ":GOTO1880
1870IF LEN(A$)=2 A$=A$+" ":GOTO1880
1880PRINTA$;:ENDPROC
1890DEF FNF(Z%)
```

```
1900LOCAL I%
1910S%=0
1920FORI%=Z% TO 1 STEP -1
1930S%=S%+I%:NEXT
1940=S%
1950 IF ERL=1480 GOTO1970
1960 CLS:END
1970*.
1980 GOTO1000
1990DEF PROCPLOT
2000MOVE0,0
2010FORI%=0 TO 999 STEP 4
2020 A%=0
2030 FOR J%=0 TO 3:A%=A%+FNP(I%+J%):NEXT:ENDPROC
2040DRAW I%,A%DIV4:NEXT
2050 VDU2:CALL &A00:VDU3
2060 ENDPROC
2070DEF PROCPOKE2(C%,A%)
2080?(&2402+2*C%)= A% MOD 256
2090?(&2403+2*C%)= A% DIV 256
2100ENDPROC
2110DEF PROCTRANSFER
2150CALL &1402
2160IF F 1 PROCREMOVEZEROS
2170ENDPROC
2180DEF PROCREMOVEZEROS
2190FOR I%=0 TO SIM%
2200IF FNP(I%) 0 GOTO2250
2210 B%=FNP(I%-1)
2220 C%=FNP(I%+1)
2230 A%=(B%+C%)/2
2240 PROCPOKE(I%,A%)
2250NEXT:ENDPROC
2260REM
2270REM
2280NEXTI%
2290DEFFNcat(DRIVE%,BUFFER%,X%)
2300LOCALY%,L%,FILE%,I%,A$,DIR%
2310Y%=X%DIV256
2320A%=&7F
2330?X%=DRIVE%
2340X%!1=BUFFER%
2350X%?5=3
2360X%?6=&53
2370X%?7=0
2380X%?8=1
2390X%?9=&21
2400X%?10=0
2410CALL&FFF1
2420IFX%?10L%=-1:=L%
2430L%=(BUFFER%?5)DIV8
2440IFL%=0THEN=0
2450X%?8=0
2460X%?10=0
2470CALL&FFF1
2480IFX%?10L%=-1:=L%
2490FORFILE%=1TOL%
2500DIR%=?(BUFFER%+FILE%*8+7)
2510A$=""
2520FORI%=0TO6
2530A$=A$+CHR$(?(BUFFER%+FILE%*8+I%))
2540NEXTI%
2550CAT$(FILE%)=A$
2560NEXTFILE%
2570=L%
2580DEFPROCCAT
2590L%=FNcat(DD%,BUFFER%,PARAM%)
2600IFL%=-1PRINTTAB(2,9)"PLEASE CHECK DISC/DRIVE"'"   THEN HIT A KEY":A=GET:GOT
O2590
2610IFDD%=2 GOTO2650
2620IFL%=0PRINT"DISC EMPTY"
2630FORF% 1TOL%:C$(F%)=CAT$(F%):?(D+F%)=DD%:NEXT
2640LI%=L%+1:DD%=2:GOTO2590
2650LI%=LI%:FORF%=LI%TOLI%+L%-1
2660C$(F%)=CAT$(F%-LI%+1):?(D+F%)=DD%:NEXT
2670L%=L%+LI%-1
2680FORI%=1 TO L%
2690VDU2:PRINT,I%,"   ";C$(I%);"  ";?(D+I%),:NEXT:VDU3:ENDPROC
```

(e)

# DENDO

```
 5REM ********     DENDO    ********
 6REM ********************************
 7REM M.C.WHITESIDE & D.WOOD
 8REM           9/4/84
 9*FX16,0:A.=&210 [illegible]
 10PROCINIT
 20CLS:PROCEM(1):PRINTN2%
 30PROCTR
 40CLS:PROCDM(1)
 50PRINT' :PROCDM(0)
 60G=GET
 70REPEAT:CLS
 80PROCEM(0)
 90PROCTR
 100PRINT'':PRINTN2%
 110UNTIL N2% 2
 120END
 140FOR I%=1TON%:L%=LEN(X2$(I%))
 142FOR J%=1TOL%:A$=STR$(ASC(MIDS(X2$(I%),J%,1)))
 144IF A$="128" A$=" "
 145IF A$="129" A$="*"
 146PRINTA$;
 147NEXT:PRINT" ";I%:NEXT
 150[illegible]:END
 160DEF PROCEM(I%)
 170[illegible]=0:XI%=1:YJ%=1
 180FOR I%=1 TO N%
 190FOR J%=I%TO N%
 200A%=ASC(I$(1%,I%,J%))
 210 IFH%<A% XI%=I%:YJ%=J%:H%=A%
 220NEXTJ%
 230NEXTI%:VDU2:PRINTTAB(5,5) [illegible]:VDU3
 240N2%=N2%-1:[illegible]
 250[illegible]
 260[illegible]
 270[illegible]
 280ENDPROC
 290DEF PROCTR
 300FOR I%=1 TO N%-1
 310L%=LEN(X2$(I%))
 320IFL%=0 PROCCL:GOTO 440
 330REM IFL%=1 PROCDIRECT:GOTO440
 340FORJ%=I%+1 TO N%
 350REM  REMOVES NULL ROW
 360IFX2$(J%)="" I$(0,I%,J%)="":GOTO430
 370[illegible]=0:D%=0
 380L J%=LEN(X2$(J%))
 390FORT%=1 TO LJ%
 400XZ%=ASC(MIDS(X2$(J%),T%,1))
 410FORK%= 1 TO L%
 420[illegible]=ASC(MIDS(X2$(I%),K%,1))
 430[illegible]
 440[illegible]
 450[illegible]
 460I$(0,I%,J%)=CHR$[illegible]
 470NEXTJ%
 480NEXTI%:ENDPROC
 490DEF PROCDIRECT
 500FORJ%=1TO N%
 510IFX2$(J%)="" I$(0,I%,J%)="":GOTO490
 520I$(0,I%,J%)=I$(1,I%,J%)
 530NEXT:ENDPROC
 540DEF PROCINIT
 550N%=42:N2%=N%
 560DIM I$(1,N%,N%+1),X1$(N%)
 570DIMX2$(N%),Y2$(N%),[illegible]N%
 580[illegible]=STRING$(N%/2," ")
 590FOR I%=1 TO N%
 600X2$(I%)=A$:Y2$(I%)=CHR$(I%)
 610X1$(I%)=A$:X2$(I%)=CHR$(I%):NEXT
 620PROCINPUT:ENDPROC
 630DEF PROCDM(I%)
 640VDU2:VDU1,27,1,15,1,12
 650FORJ%=1 TO N%
 660A$=STR$(J%):PROCPAD
 670FORI%=1TO J%
 680PRINT:ASC(I$(1%,I%,J%))[illegible]
 690NEXTI%:PRINT:NEXTJ%
 700FORI%=0TON%:A$=STR$(I%):PROCPAD:NEXT
 710[illegible]:ENDPROC
 800DEF PROCCL
 900FORJ%=1TO N%
 910I$(0,I%,J%)=""
 920NEXT:ENDPROC
 1000DEF PROCINPUT
 1010*DR.2
 1020LOAD"COEFS" 1900
 1030FOR I%=1TO N%-1
 1040FOR J%=I%+1 TO N%
 1050I$(1,I%,J%)=CHR$(?(I%*N%+J%+[illegible]
 1060NEXT:NEXT
 1070ENDPROC
 1100DEF PROCPAD
 1110IF LEN(A$)=0 A$=" ":GOTO1140
 1120IF LEN(A$)=1 A$=A$+" ":GOTO1140
 1130IF LEN(A$)=2 A$=A$+" ":GOTO1140
 1140PRINT;A$;:ENDPROC
```