Cloning, Sequencing, and Expression of the puc genes of two strains of *Rubrivivax gelatinosus*

by

Adrian Edward Simmons

A thesis submitted for the degree of doctor of philosophy.

UNIVERSITY
of
GLASGOW

Division of Biochemistry & Molecular Biology

Institute of Biomedical & Life Sciences

Declaration

The composition of this thesis and the work described within it was carried out entirely by myself unless otherwise cited or acknowledged. Its contents have not previously been submitted for any other degree. The research for this thesis was carried out between October 1991 and October 1994.

Signed: .....

A. E. Simmons

Tue, Oct 3, 1995

Abbreviations

Bchl	Bacteriochlorophyll
Bphe	Bacteriophaeophytin
Car	Carotenoid
CM	Cytoplasmic membrane
HTH	Helix-turn-helix
ICM	Intracytoplasmic membrane
LHC	Light-harvesting Complex
LH1	Light-harvesting complex one
LH2	Light-harvesting complex two
Mb, kb, bp	Mega base pairs, Kilo base pairs and base pairs
MCS	Multiple Cloning Site
NIR	Near Infrared
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PGC	Photosynthesis Gene Cluster
PU	Photosynthetic Unit
RNA	Ribose Nucleic Acid
RC	Reaction Centre
SDS–PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
WT	Wild Type

Genus abbreviations for Bacterial species

Er.	*Erythrobacter*
Rb.	*Rhodobacter*
Ro.	*Roseobacter*
Rps.	*Rhodopseudomonas*
Rs.	*Rhodospirillum*
Rv.	*Rubrivivax*

Abstract

An initial observation that the two strains 149 and 151 of *Rv. gelatinosus* synthesise differing amounts of LH2 (B800-850) when cultured under the same conditions was confirmed. Strain 151 produces approximately one third more LH2 than strain 149 when cultured at any given light intensity. It was also found that both strains vary the level of LH2 to a similar extent in response to changing irradiance levels, both synthesising twice the amount of LH2 under low irradiance than was synthesised at high irradiance.

Southern analysis indicated the presence of a single copy of the pucBA genes in each strain, but suggested there may be a difference between them with regard to the pucC gene. Genomic libraries were constructed from the two strains of *Rv. gelatinosus* using the λ replacement vector λGEM-11. The pucBA genes coding for the α and β-polypeptides of the LH2 complex were cloned and sequenced from both strains, as was the pucC gene from strain 151 and a substantial region of sequence probably involved in transcriptional control. The predicted amino acid sequence of the strain 151 and 149 LH2 α and β-polypeptides matched that achieved by protein sequencing (Brunisholz *et al.*, 1994), whilst the predicted sequence of the strain 151 PucC protein was shown to have high sequence and structural homology to other PucC proteins.

The arrangement of the *Rv. gelatinosus* puc genes was found to be somewhat different from that found in other bacteria from which they have been cloned. With the pucC gene being present downstream of the pucBA genes, but in the opposite orientation. Analysis of the sequence upstream of the pucBA and pucC genes has identified possible *E. coli* σ^{70} like promotor elements, two upstream of pucBA and one upstream of pucC. Near palindromes similar to the *Rhodobacter* Pps oxygen sensitive repressor binding sites were also found, two in the promotor more proximal to pucB, and a single one straddling the start of the pucC gene.

Northern analysis was carried out on cultures of the two strains grown under a variety of conditions, these indicate that the *Rv. gelatinosus* pucBA genes are repressed in the presence of oxygen, and are expressed more under low irradiance conditions than at high irradiance. Thus *Rv. gelatinosus* appears to respond to light and oxygen in a similar manner to other photosynthetic bacteria. The northern analysis also indicates a transcript size of around 600 bp for the puc genes of *Rv gelatinosus*.

Acknowledgements

A big acknowledgement must go to Dr. Christopher MacKenzie for help and advice in the construction of the strain 151 genomic library, and for instilling me with some molecular biological common sense. Thanks also go to Dr Stuart Barrett and Dr Alastair Gardiner help and advice in the lab. I would also like to thank my supervisor Prof. Richard Cogdell for his support. Thanks go to Norman Tait for developing my photographs, and to Dr Peter Dominy and Dr Joel Milner for occasional suggestions and advice. An acknowledgement is also due to Prof. John Allen for starting it all off by stimulating my interest in photosynthesis.

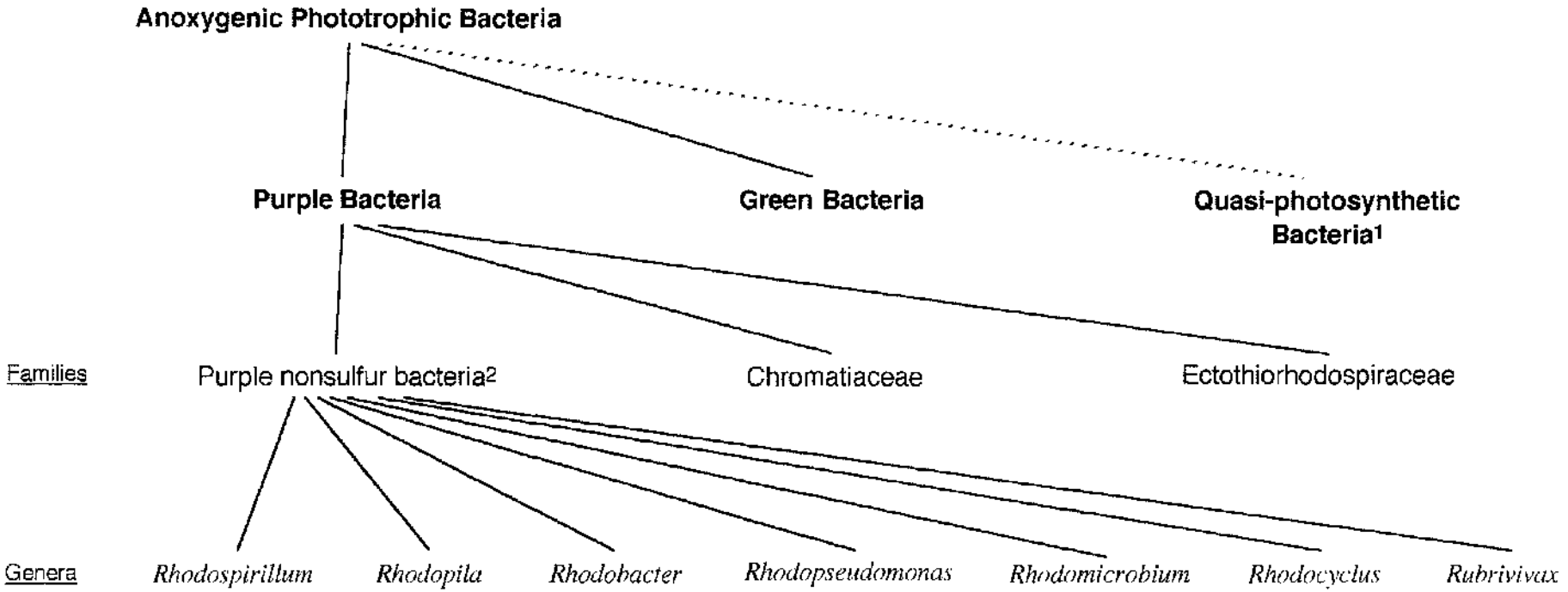
Many thanks to my mother for *really* starting it all off, and for all those years of emotional and financial support. Thanks to Helena for putting up with the stress, the moods, and just for being there. I would also like to thank Mrs Kyra Schumann for her interest and financial support, and Morgan Shaw for his computer related financial assistance. Thanks also to Fran, Declan, Steve and Sean for dragging me away from it all every once in a while. Thanks to Gaye and Andy for the hill walking (amongst other things). And lastly thanks to the past and present members of the Hooker, Arnott, Wolfson and B4a laboratories for their help, advice and friendship (you know who are).

Index

Table of Contents

Figure 1

Classification of the Purple nonsulfur bacteria. Adapted from Pfennig and Trüper, 1989.



Note : Further genera have been suggested for the purple nonsulfur bacteria more recently, usually to encompass new wild-type isolates, these include: *Rhodoferax* (Hochkoeppler *et al.*, 1995), *Rhodobium* (Hiraishi *et al.*, 1995), *Rhodovulum* (Hiraishi & Ueda, 1994).

[1] There are a number of taxa grouped with anoxygenic phototrophic bacteria under the title Quasi-photosynthetic Bacteria. These species are clearly related to purple photosynthetic bacteria but appear to use photosynthesis only as a secondary energy source. Genera currently grouped under this heading include *Roseobacter*, *Roseococcus*, *Erythrobacter*, *Erythromicrobium* and *Rhizobium* (Gest, 1993).

[2] The Purple nonsulfur bacteria were previously known as the Rhodospirillaceae, this family name has now been discarded, please see the main text for details.

Introduction

Photosynthesis, in its widest sense, can be defined as a series of processes in which light energy is converted to chemical energy used for biosynthesis of organic cell materials; a 'photosynthetic organism' is thus defined as one which uses light to provide a major fraction of the energy required for cellular syntheses (Gest, 1993).

The simplicity of this description belies the structural and physiological complexity found in photosynthetic organisms, which are now considered to include higher plants, cyanobacteria, green and purple photosynthetic bacteria, and 'quasi-photosynthetic bacteria' (Gest, 1993). The photochemical systems of this immense range of organisms could not be described in depth without writing several books. However photosynthesis can be conceptually and physically divided into the light reactions, involving the photochemical conversion of light energy to chemical energy, and the dark reactions, which utilise this chemical energy to "fix" carbon dioxide (or carbon in the form of organic compounds) into carbohydrate. My work has involved the molecular biology of the light reactions in photosynthetic purple bacteria, thus I will first discuss the taxonomy of the photosynthetic purple bacteria and related species, before moving on to describe the structure and functioning of cellular components involved in the light reactions, and the molecular biology underlying photosynthesis in these organisms.

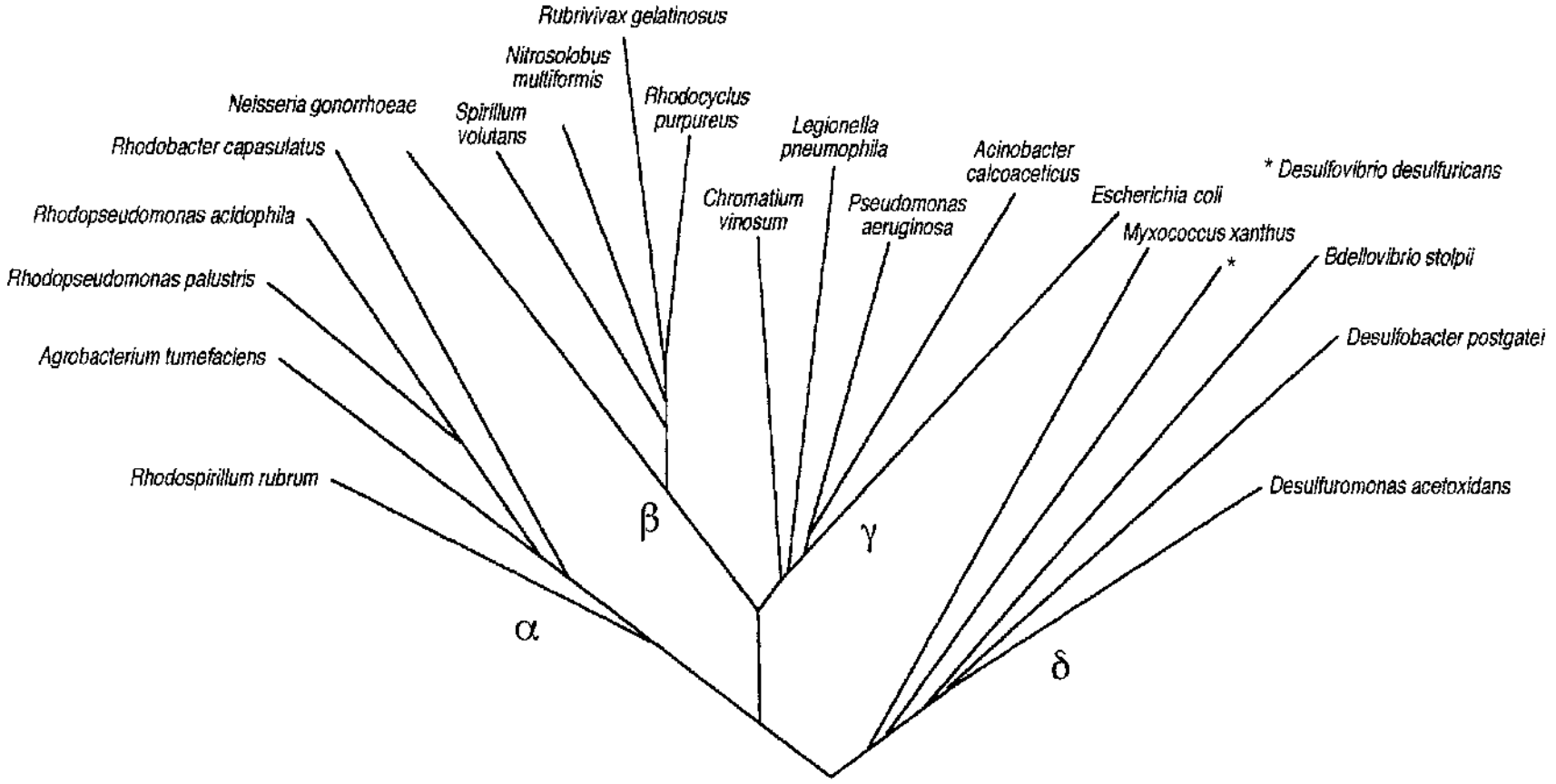
Taxonomy of the Purple Photosynthetic Bacteria

Currently two systems of bacterial classification are in use, in the older determinative system bacteria are grouped together into assemblies of genera which have in common cytological, physiological and morphological characteristics (Staley *et al.*, 1989). These assemblies are known by vernacular names and provide a practical basis by which the bacteria can be identified, organised and named. The grouping of taxa in this way does not however imply any phylogenetic (evolutionary) relationship between genera (Staley and Krieg, 1989). More recently molecular biology has encouraged the construction of a systematic system of classification which attempts to define an evolutionary framework for bacteria and prokaryotes (see Woese, 1987).

The organisation of anoxygenic phototrophic bacteria provides an example of the state of bacterial taxonomy and is summarised (with a bias towards the purple nonsulfur bacteria) in Figure 1 (opposite). The anoxygenic phototrophic bacteria consist of a group of bacteria which are usually aquatic, able to grow by photosynthesis under anaerobic conditions and contain Bchl, carotenoids, cytochromes, quinones and non-heme iron proteins (Pfennig & Trüper, 1989). Most species are also capable of growth under aerobic conditions using respiration, or in the dark via fermentative metabolic pathways. This group is divided into two major subsets, the Purple and Green Bacteria, on the basis of LHC structure. In Purple bacteria the photosynthetic apparatus is located in the ICM (intracytoplasmic membrane) and CM (cytoplasmic membrane) (Imhoff & Trüper, 1989) whilst in Green bacteria the photosynthetic apparatus is located in the CM and in Chlorosomes (Pfennig, 1989; Cohen-Bazire *et al.*, 1964).

Figure 2

Phylogenetic tree for the Proteobacteria (purple bacteria) constructed using 16s rRNA sequences, indicating the four rRNA based subdivisions. Adapted from Woese, 1987.



The Purple Bacteria were previously arranged into three families, the Rhodospirillaceae, Chromatiaceae and Ectothiorhodospiraceae (Imhoff *et al.*, 1984). However it has been found that most purple photosynthetic bacteria are more closely related to various non-photosynthetic bacterial species than to each other (Woese *et al.*, 1984a, 1984b; Stackebrandt *et al.*, 1988). Thus the designation Rhodospirillaceae is not taxonomically correct, however for the time being these bacteria are still grouped together under the name purple nonsulfur bacteria (Imhoff & Trüper, 1989).

Constituting a further group which is clearly related to the anoxygenic phototrophic bacteria but shows obvious differences are the Quasi-photosynthetic bacteria (Gest, 1993), these were previously defined as 'Aerobic photosynthetic bacteria' (Shiba & Harashima, 1986). Members of this group are characterised by the production of Bchl and photosynthetic apparatus when growing under aerobic conditions (in contrast to the anoxygenic photosynthetic bacteria), the designation quasi-photosynthetic refers to the fact that they appear to use photosynthesis as a secondary source of energy for growth under aerobic conditons, and appear unable to photosynthesise under anerobic conditions (Gest, 1993).

Whilst there is no generally accepted scheme for the classification of the prokaryotes at higher taxonomic levels, several suggestions have been made (Murray, 1989). One of the more promising schemes and one increasingly adopted as a basis for bacterial classification, is that based on the analysis of ribosomal RNA sequences (Woese & Fox, 1977; Woese, 1980; Woese, 1987).On the basis of an analysis of 16s rRNA sequences from a range of bacteria, four rRNA based subdivisions have been defined within the 'purple bacteria'(Proteobacteria), the α, β, γ, and δ subdivisions. A phylogenetic tree created using the 16s rRNA sequences of many photosynthetic and non-photosynthetic purple bacterial species is shown in Figure 2 (Opposite). It is interesting to note that almost all the presently known purple nonsulfur bacterial species can be found within the α subdivision, only the species *Rc. purpureus*, *Rc. tenue* and *Rv. gelatinosus* are located in the β subdivision (Woese *et al.*, 1984a &1984b). These three species differ from those in the α subclass not only in their 16s rRNA sequence but in the type of cytochrome *c* which they synthesise (Dickerson, 1980a &1980b), and possibly in the structure of their photosynthetic reaction centres (Clayton & Clayton, 1978; Fukushima *et al.*, 1988).

The use of 16s rRNA sequences alone to determine the phylogenetic history of bacteria would clearly be undesirable, and unlikely to show a completely true history of evolution. However an analysis of cytochrome *c* sequences and structure (Dickerson, 1980b), is in good agreement with that of 16s rRNA sequences. Thus as the range of bacterial species from which these two genes have been sequenced increases, and as more genes are found to be relatively stable in structure throughout evolution, a wholly molecular system of bacterial classification is arising.

The Light Reactions of Photosynthesis

In essence the light reactions of photosynthesis have two phases, firstly the capture of light energy and secondly its conversion to chemical energy. In those photosynthetic organisms studied thus far both these steps have been found to occur in assemblies of pigment–protein complexes (Kaplan & Arntzen, 1982; Drews, 1985). These complexes have been studied in most detail in photosynthetic bacteria where structural simplicity (in comparison to higher plants), ease of isolation and the ability to apply molecular biological techniques, have facilitated a more rapid advance in knowledge. However the similarity in structure between the photosynthetic systems of higher plants, algae, cyanobacteria, and those of photosynthetic bacteria allows much of the knowledge gained in the study of photosynthetic bacteria to be applied to all photosynthetic organisms (Zuber & Brunisholz, 1991; Margulies, 1991; Ford, 1992; Otsuka, 1992).

Structure and function of the bacterial PSU

The purple bacterial photosynthetic apparatus contains pigment molecules bound to integral membrane proteins and arranged in a highly ordered fashion, which appears to be essential for efficient energy capture and conversion (see following sections) (Zuber & Brunisholz, 1991). Those purple bacteria studied thus far have a photosynthetic unit which consists of a Reaction Centre complex (RC) which carries out a charge separation (Feher *et al.*, 1975; Norris *et al.*, 1975), surrounded by one or more Light-harvesting complexes (LHC), which as their name suggests are involved in the absorption of light energy followed by its transfer to the RC (Knox, 1977; Pearlstein, 1982; Sauer, 1986). Following the transfer of an exciton to the RC and charge separation, the electron produced is used to reduce a quinone molecule (Debus *et al.*, 1986). This process is repeated and a second electron passed to the quinone (Dracheva *et al.*, 1988), the quinone then takes up two protons from the cytoplasm and diffuses out of the RC complex (Crofts & Wraight, 1983). The membrane bound cytochrome bc_1 complex transfers the electrons and protons from the quinone to the periplasm (Trumpower, 1990), the electrons are then returned to the RC by cytochrome proteins (Rich, 1984; Deisenhofer, 1985), and the protons used by the ATP synthase complex to synthesise ATP. Thus light driven cyclic electron transport builds a proton gradient across the membrane, which can then be used for the synthesis of ATP (Mitchell, 1979).

Figure 3

Figures 3a, b and c show structural views of the RC from the carotenoidless mutant strain R-26 of *Rb. sphaeroides*, and were adapted from Rees, *et al.*, 1989.

Figure 3a

Stereo veiw of the RC. This shows the cofactors and the carbon backbone of the protein subunits. The twofold axis of symmetry is in the plane of the paper, with the cytoplasmic side at the bottom of the figure.

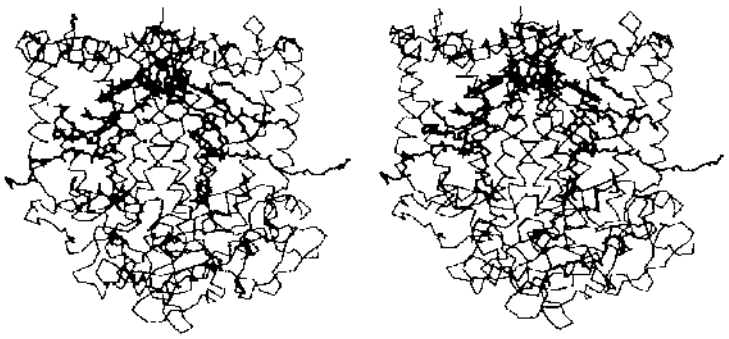


Figure 3b

A second view of the RC, showing cofactors and protein subunits. The α-helices are represented as cylinders. Helices of the L subunit are labelled in plain type, those of the M subunit in italic and H subunit in bold. The phytyl and isoprenoid tails of the cofactors have been truncated to clarify the view. The RC is in the same orientation as figure 3a.

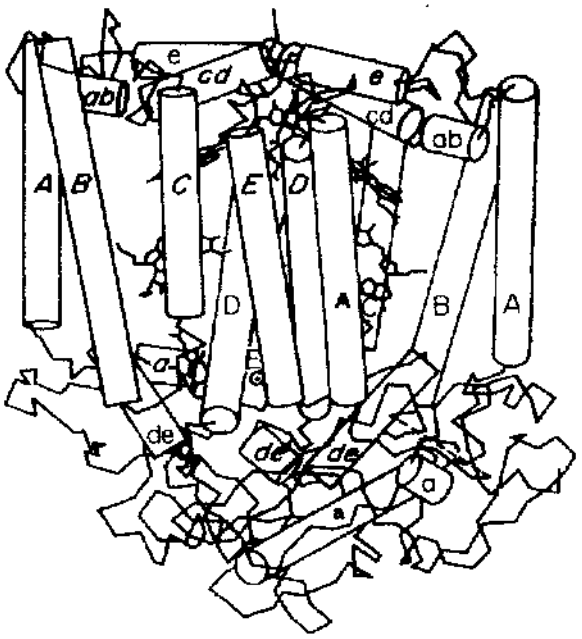
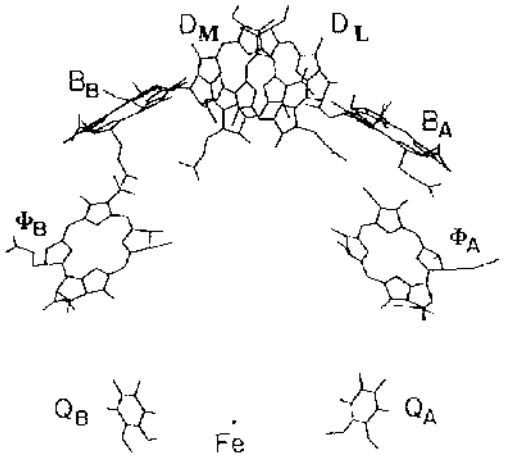


Figure 3c

Cofactor structure of the RC shown in the same orientation as figure 3a. The phytyl and isoprenoid tails of the cofactors have been truncated. D, Bchl dimer; B, accessory Bchl; Φ, Bphe; Q, Quinone; Fe, non-heme iron. For further details of nomenclature please see main text.



The reaction centre and energy conversion

The determination to atomic resolution of the RC structures from *Rhodopseudomonas viridis* (Deisenhofer *et al.*, 1985; Deisenhofer & Michel, 1989) and those from *Rhodobacter sphaeroides* strain R26 (Allen *et al.*, 1987a & b; Yeates *et al.*, 1987; Chang *et al.*, 1991), wild type strain 2.4.1 (Allen *et al.*, 1988a; Yeates *et al.*, 1988) and wild type strain Y (Arnoux *et al.*, 1990; Reiss-Husson *et al.*, 1990), has provided detailed information about the way in which photosynthetic organisms carry out the primary events of photosynthesis. The structure and function of reaction centres has been reviewed extensively (Feher *et al.*, 1989; Deisenhofer & Michel, 1991a; Deisenhofer & Michel, 1991b; Ermler *et al.*, 1994; Deisenhofer *et al.*, 1995), for a more detailed description than that given below the reader should consult these reviews.

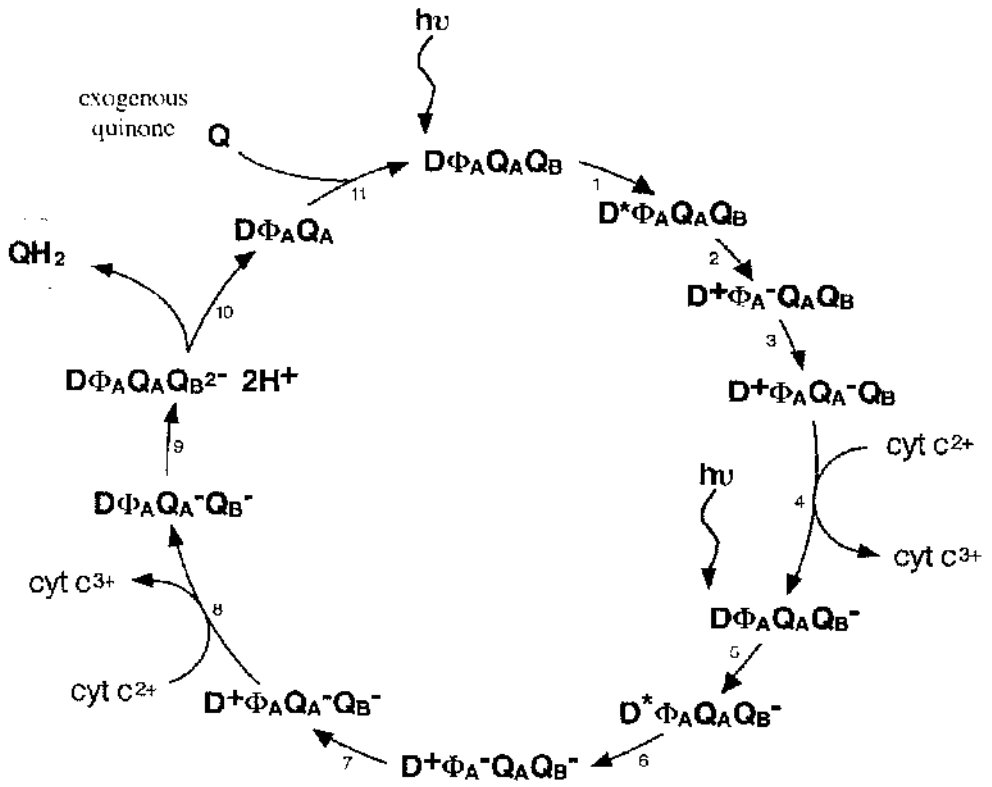
The structure of the RC has proved to be highly conserved not only between the different strains of *Rb. sphaeroides*, but also between *Rb. sphaeroides* and *Rps. viridis*. Indeed so homologous are the structures that the coordinates of the *Rps. viridis* RC were used to solve the phase problem of the initial *Rb. sphaeroides* crystallographic structure (Allen *et al.*, 1986; Chang *et al.*, 1986). The main difference between the two species concerns the number of protein subunits present within the crystallised complex. Both species have three polypeptides termed L (light), M (medium) and H (heavy), according to their apparent molecular weights as determined by SDS–PAGE (Feher & Okamura, 1978). But *Rps. viridis* has an additional fourth subunit, a c-type cytochrome which is bound to the RC on the periplasmic side of the membrane, now termed as subunit C (Deisenhofer *et al.*, 1995).

Bound to the L and M polypeptides are ten cofactors that take part in the charge separation that is the function of a reaction centre. These vary slightly between the two species but include 4 Bchl molecules (Bchl *a* in *Rb. sphaeroides*, Bchl *b* in *Rps. viridis*), 2 Bphe (Bphe *a* in *Rb. sphaeroides*, Bphe *b* in *Rps. viridis*), 2 quinone molecules (both ubiquinone-10 in *Rb. sphaeroides*, one menaquinone-9 and one ubiquinone-9 in *Rps. viridis*), a ferrous iron ion, and a single carotenoid molecule (*Rb. sphaeroides* uses sphaeroidene whilst *Rps. viridis* has dihydroneurosporene).In addition to this *Rps. viridis* has four heme units covalently bound to the C subunit (cytochrome c), in *Rb. sphaeroides* a water soluble cytochrome c_2 interacts with the reaction centre in place of a bound cytochrome (Deisenhofer *et al.*, 1995; Ermler *et al.*, 1994).

Figure 3 (Opposite) illustrates the structure of the RC from *Rb. sphaeroides* strain R26, showing the three polypeptides L, M and H, together with their associated cofactors (excepting the carotenoid). Figure 3a provides a view of the overall structure of the RC, Figure 3b also gives an overall view but highlights the protein subunits. In total the 3 protein subunits have 11 transmembrane α-helices, the L and M subunits contribute 5 helices each (labelled A, B, C, D and E) and the H subunit 1 (Deisenhofer *et al.*, 1995; Ermler *et al.*, 1994). The remainder of the H subunit consists of a globular domain which lies within cytoplasm, contains 2 antiparallel β sheets, and appears to stabilise the complex whilst also creating part of the environment necessary for efficient electron transfer between the two quinone molecules (Debus *et al.*, 1985).

Figure 4

Schematic diagram of energy transfer and cyclic electron transport through the reaction centre. Adapted from Feher *et al.*, 1989.



Note: Following the initial excitation of D (step 1), the primary electron transfer steps between D^* and Φ_A (step 2) take place in ~3-4 ps (Flemming *et al.*, 1988). Although not illustrated here B_A does play a role, electron transfer from D to B_A occurs in ~3.5 ps, followed by a 0.9 ps transfer between B_A and Φ_A (Creighton *et al.*, 1988; Dressler *et al.*, 1991; Hamm *et al.*, 1995). Electron transfer from Φ_A to Q_A (step 3) is somewhat slower at ~200 ps (Kirmaier & Holten, 1987), and that between Q_A and Q_B slower still at around 100 ms (Debus *et al.*, 1986). Before a second electron transfer to Q_B can take place the dimer (D) must be re-reduced by cytochrome[1] (step 4). A second excitation can then occur (step 5) and drive a second electron transfer to Q_B (steps 6, 7 & 8), thus producing a double reduced quinone Q_B^{2-} (Dracheva *et al.*, 1988). This reduced quinone picks up a complement of two protons (step 9), diffuses out of the reaction centre (QH_2, circled) (step 10), and is replaced by an exogenous quinone from the membrane pool (step 11) (Crofts & Wraight, 1983). A typical total cycle time would be ~10^{-3} s (Feher *et al.*, 1989). For details of abbreviations and nomenclature please see main text.

[1] In *Rb. sphaeroides* the reduction of the special pair is carried out by a water soluble cytochrome c_2 which interacts with the RC (Rich, 1984), in *Rps. viridis* the cytochrome bound to the RC carries out the reduction (Deisenhofer, 1985).

One of the most striking features of the RC is the two-fold symmetry of the L and M subunits, this symmetry is also visible in the arrangement of the cofactors bound to them. The L and M subunits hold the cofactors in two almost identical symmetrical 'arms', as illustrated in Figure 3c, despite the similarity between the two arms electron transfer appears to occur preferentially along the chain of cofactors bound to the L subunit. A close examination of the two subunits shows that although their overall similarity is high, minor differences in the structure of the proteins, orientation of cofactors and in the protein environment surrounding the cofactors, result in the L chain being the favoured route for electron transfer (Deisenhofer *et al.*, 1995; Ermler *et al.*, 1994). This asymmetry has given gives rise to a cofactor nomenclature that distinguishes between the two chains, with those cofactors bound to the L subunit being termed as the 'A' chain, and those bound to the M subunit as the 'B' chain. The cofactors themselves have recognised abbreviations (D, Bchl dimer; B, accessory Bchl; Φ, Bphe; Q, Quinone; Fe, non-heme iron), each cofactor is referred to as, for example, Φ_A or Φ_B, thus indicating the cofactor chain to which it belongs. The exception to this rule relates to the two Bchl molecules which form the dimer or "special pair". Since both these molecules are involved in the charge separation reaction and thus can be considered as part of the A chain, they are labelled D_L and D_M in accordance with the protein subunit to which they are bound (Deisenhofer *et al.*, 1995).

The three dimensional arrangement of the cofactors is illustrated in Figure 3c. At the periplasmic side of the complex forming and lying on the central axis of the 2 fold symmetry, are two Bchl molecules (D_L & D_M) which form a dimer. Next along the chains are two monomeric Bchl molecules (B_A & B_B), below these are Bphe molecules (Φ_A & Φ_B) and following those the two quinone molecules (Q_A & Q_B). Between Q_A & Q_B the non heme iron is bound, in *Rps. viridis* the iron is centrally placed between the two quinone molecules, but in *Rb. sphaeroides* it is asymmetrically placed and lies slightly closer to Q_B than to Q_A (Deisenhofer *et al.*, 1995; Ermler *et al.*, 1994). The carotenoid molecule present in the wild type RC is not shown in Figure 3c, since the Figure shows the cofactors from strain R26, a carotenoidless mutant. However the long chain carotenoid molecule has now been identified on the M subunit side of the RC, lying in the plane of the membrane close to, and in van de Waals contact with, B_B in both *Rb. sphaeroides* and *Rps. viridis* (Arnoux *et al.*, 1990; Deisenhofer *et al.*, 1995).

Energy transfer through the reaction centre and beyond

Cyclic electron transfer through the RC is summarised in Figure 4 (opposite). Initially energy is passed from the light-harvesting complexes to the Bchl dimer or "special pair", where a charge separation occurs (Feher *et al.*, 1975; Norris *et al.*, 1975). The energy, now in the form of an electron migrates rapidly along the A branch via B_A (Creighton *et al.*, 1988; Zinth *et al.*, 1985; Dressler *et al.*, 1991; Hamm *et al.*, 1995) and Φ_A to the primary quinone Q_A, where it is used to reduce the secondary quinone, Q_B (Debus *et al.*, 1986). A second charge separation event followed by the transfer of a second electron down the A chain to Q_B (Dracheva *et al.*, 1988), is necessary before Q_B is able to take up two protons from the cytoplasm and diffuse out of the the RC into the lipid membrane (Crofts & Wraight, 1983). Once in the membrane Q_B passes to the

membrane spanning cytochrome bc_1 complex, this oxidises the quinol and simultaneously transfers the electrons plus three to four protons to the periplasm (Trumpower, 1990), thus generating the proton gradient necessary for ATP synthesis (Mitchell, 1979). The electrons are shuttled back to the RC via a water soluble cytochrome c_2 (Rich, 1984), and the bound cytochrome in *Rps. viridis* (Deisenhofer, 1985), where they are used to reduce the Bchl special pair prior to a new excitation event.

In comparison to the purple nonsulfur bacteria the light driven cyclic electron transfer of the quasi-photosynthetic bacteria seems to be impaired. The RC of the quasi-photosynthetic bacterium *Ro. denitrificans*[1] differs from other purple bacterial RCs in the structure of its cytochrome subunit, which is associated with the RC as in *Rps. viridis* (Shimada *et al.*, 1985; Takamiya *et al.*, 1987). It appears that the RC cytochrome is unable reduce the special pair under anaerobic conditions (Garcia *et al.*, 1994), thus impeding cyclic electron transfer through the RC, and preventing this organism carrying out photosynthesis under anaerobic conditions (Garcia *et al.*, 1994; Okamura *et al.*, 1985). Thus an impaired photosynthetic electron transfer pathway appears to lead to a requirement for aerobic conditions in order that cell growth and photosynthesis (at low effeciency) can occur (Okamura *et al.*, 1986), and probably represents the core difference between quasi-photosynthetic and purple photosynthetic bacteria (Okamura *et al.*, 1985; Gest, 1993).

The role of cofactor environment in electron transfer

Electron transfer has been found to occur down only one of the two apparently symmetrical arms of cofactors found in the RC (Michel-Beyerle *et al.*, 1988). From analysis of the 3 dimensional structures and amino acid sequences of the proteins, it seems likely that differences in the amino acid sequences of the L and M subunit proteins result in deviations from the axis of symmetry (Bélanger *et al.*, 1988; Ermler *et al.*, 1994; Deisenhofer *et al.*, 1995). The resulting differences in relative cofactor orientation, and interactions of amino acid side chains with the cofactors (Deisenhofer & Michel, 1989), appear to provide an explanation for the observed route of electron transfer down the A branch. Systems suitable for the site directed mutagenesis of RC proteins in several species were constructed (Youvan *et al.*, 1985; Farchaus & Oesterhelt, 1989; Nagarajan *et al.*, 1990; Paddock *et al.*, 1989; Takahashi *et al.*, 1990; Laussermair & Oesterhelt, 1992; Jones *et al.*, 1992). These made possible studies involving site directed mutagenesis of RC proteins in combination with spectroscopic analysis of mutation effects, which have confirmed that the route and speed of energy transfer through the reaction centre is greatly influenced by the arrangement and environment of the cofactors . A review of mutagenesis studies can be found in (Ermler *et al.*, 1994). A summary of the main cofactor-protein interactions is given below.

The Bchls forming the special pair are covalently bound to Histidine residues in the L and M proteins via the magnesium atoms at the centre of their porphyrin rings (Michel *et al.*, 1986), hydrogen bonding also occurs between the special pair and several amino acid residues (Ermler *et al.*, 1994). In the region of the special pair in

[1] This species was previously known as *Erythrobacter* sp. OCh 114 (Shiba, 1991).

Rps. viridis are Tyrosine (M subunit) and Histidine (L subunit) residues, which form hydrogen bonds to the D_M and D_L respectively. In *Rb. sphaeroides* the residue comparable to the Histidine is a Phenylalanine residue that is unable to form hydrogen bonds, thus asymmetry is introduced into the binding of the cofactors (Ermler *et al.*, 1994).

Two residues, a Tyrosine residue on the A side of the complex and a symmetry related Phenylalanine on the B side, lie near the centre of the complex and are each in van der Waals contact with 4 pigment molecules (Tiede *et al.*, 1988). That these residues are conserved in the RC's of the species *Rps. viridis*, *Rb. sphaeroides*, *Rb. capsulatus*, and *Rs. rubrum*, suggests that they have an important function. Mutagenesis studies in *Rb. sphaeroides* suggest that these residues are involved in electron transfer from the special pair to Φ_A (Nagarajan *et al.*, 1990; Gray *et al.*, 1990).

The two Bchl monomers (B_A and B_B) are bound to the proteins in a similar manner to the special pair, forming covalent bonds between their Mg^{2+} atoms and Histidine residues on the L and M subunits (Ermler *et al.*, 1994; Deisenhofer *et al.*, 1995). As well as interacting with the proteins the monomers appear to interact with the special pair and Bphe molecules (Creighton *et al.*, 1988). The Bphe molecule Φ_A forms a hydrogen bond with a conserved Glutamine residue on the L subunit, no similar residue is found near Φ_B (Ermler *et al.*, 1994).

Between Φ_A and Q_A in both *Rb. sphaeroides* and *Rps. viridis* is a conserved Tryptophan residue that forms a hydrogen bond with a Threonine (both residues are on the M subunit), these residues appear to play a role in binding of Q_A and possibly in electron transfer from Φ_A to Q_A (reviewed by Ermler *et al.*, 1994). The binding site of Q_A in *Rps. viridis* contains two residues, an Alanine and a Histidine, which hydrogen bond to the menaquinone (Q_A; Michel *et al.*, 1986). *Rb. sphaeroides* uses a ubiquinone, but this is found in a similar position (Allen, *et al.*, 1988b).

The structure of the Q_B binding site and position of the Fe^{2+} atom vary between *Rb. sphaeroides* and *Rps. viridis*. In both species the Fe^{2+} is covalently bound to 4 histidine residues (2 on M and 2 on L), and a glutamic acid residue (M subunit), it has been suggested that these interactions with Fe^{2+} may play a role in stabilising the RC structure (Michel *et al.*, 1986; Feher *et al.*, 1989; Ermler *et al.*, 1994). In *Rps. viridis* the Fe^{2+} atom lies almost equidistant from both Q_A and Q_B on the cofactor axis of symmetry which runs perpendicular to the membrane and centres on the special pair (Michel *et al.*, 1986). In *Rb. sphaeroides* the Fe^{2+} atom is asymmetrically positioned and lies closer to Q_B than to Q_A, both of which are ubiquinone molecules (Allen, *et al.*, 1988b). This asymmetry is thought to provide an energetically favourable environment for electron transfer from Q_A to Q_B (Feher *et al.*, 1989). Such asymmetry is not required in the *Rps. viridis* RC because of the inherent difference in redox potential between Q_A (menaquinone) and Q_B (ubiquinone; Feher *et al.*, 1989). Several studies have focused on the path of proton transfer to Q_B, two residues in the Q_B binding pocket are of particular importance, a Serine residue (L subunit) is thought to donate the first proton, and a Glutamine (L subunit) the second (Okamura and Feher, 1992; Leibl *et al.*, 1993).

Refinement of the *Rps. viridis* reaction center crystallographic structure has

revealed the presence of 201 water molecules within the RC (Deisenhofer *et al.*, 1995). More than half of these are bound at the interfaces between protein subunits, and thus probably contribute to subunit interactions, water molecules were also found to form part of the environment around some of the cofactors. The water molecules surrounding the cofactors interact with amino acid side chains and cofactors to form networks of hydrogen bonds. Such networks involving water molecules are found between the special pair and first heme group of the cytochrome c subunit, also in the region of B_B and B_A (where hydrogen bonds arise not only to B_B and B_A but to residues which interact with the special pair), and surrounding the Q_B binding site (Deisenhofer *et al.*, 1995). These bound water molecules clearly have roles to play in the structure and functioning of the RC.

Light Harvesting Complexes

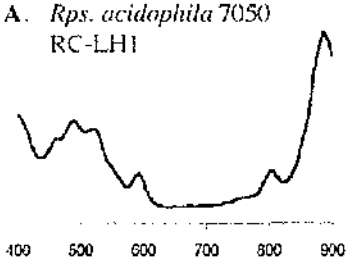
Surrounding the RC are one or more light-harvesting complexes which capture light energy and transfer it to the RC by exciton transfer (Sauer, 1986). All species of photosynthetic purple bacteria studied thus far appear to synthesise a 'core' LHC which is associated with the RC, and which is present in fixed amounts relative to the RC, some species also synthesise 'peripheral' complexes which are present in variable amounts depending on the prevailing environmental conditions (Aagard & Sistrom, 1972; Takemoto & Huang Kao, 1977; Thornber *et al.*, 1978; Cogdell & Thornber, 1979; Drews & Oelze, 1981; Ohad & Drews, 1982; Chory & Kaplan,1983; Hawthornwaite & Cogdell, 1991). The LHCs are identified, named and classified according to their NIR absorption spectra (Cogdell & Thornber, 1980), thus a B890 core complex has a single absorbance peak at 890 nm, whilst the B800-850 and B800-820 peripheral complexes have two peaks in the NIR. It should be noted that certain species, notably *Rps. palustris*, are able to synthesise two spectrally different types of B800-850, these are termed type I and type II B800-850 (Hayashi, *et al.*, 1982; Thornber *et al.*, 1983; Evans *et al.*, 1990). *Rps. acidophila* was also thought to produce a similar type II LH2, but the identification of this complex may have resulted from incomplete separation of the type I LH2 and LH3 during LHC isolation (Cogdell *et al.*, 1983).

There is an alternative nomenclature by which the complexes are also known, where the B890 type complex is called the Light-harvesting one (or LH1) complex, the B800-850 complex the Light-harvesting two (or LH2) complex and the B800-820 complex the Light-harvesting three (or LH3) complex. For the purposes of this discussion the latter nomenclature will be used to describe the various LHCs.

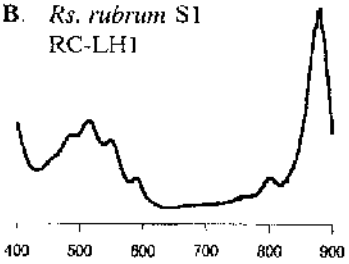
Figure 5

Absorption spectra of solubilised chromatophore membranes and isolated light-harvesting complexes from a range of purple photosynthetic bacteria. Data suplied by Dr. Alastair Gardiner.

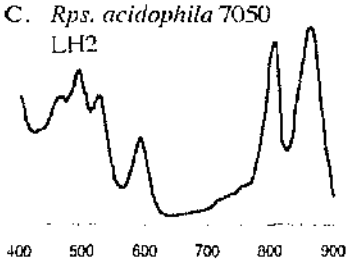
A. *Rps. acidophila* 7050
RC-LH1



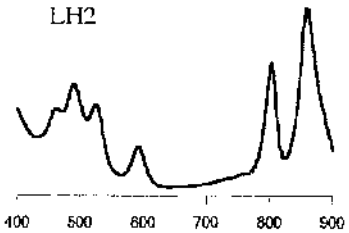
B. *Rs. rubrum* S1
RC-LH1



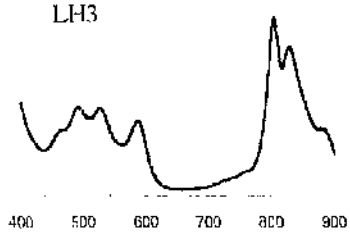
C. *Rps. acidophila* 7050
LH2



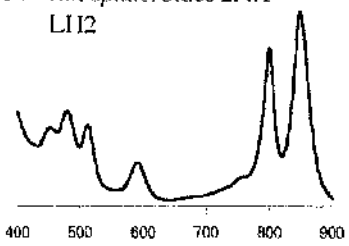
D. *Rps. acidophila* 10050
LH2



E. *Rps. acidophila* 7050
LH3

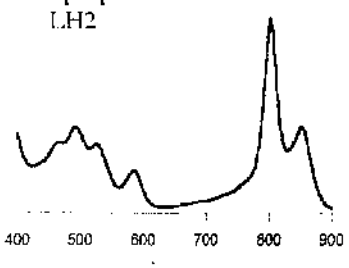


F. *Rb. sphaeroides* 2.4.1
LH2



Note: Each of the spectra (A to G) is labelled with the species and strain from which it was taken, plus the type of LHC represented. The y axes (absorbance units) have been omitted for clarity. Spectra were recorded on a Shimadzu UV-2101-PC spectrophotometer, from solubilised membranes in the case of spectra B, C, D & E (prepared according to the method described by Gardiner *et al.*, 1993), or from isolated complexes, spectra A, F & G, prepared using sucrose density gradient centrifugation (based on the method of Firsow & Drews, 1977).

G. *Rps. palustris* French
LH2



The characteristic absorption spectra for LH1, LH2 and LH3 from a range of species are illustrated in Figure 5 (opposite)[2]. *Rps. acidophila* strain 7050 produces all three LHC types, as shown in spectra A, C and E. The actual absorbance peaks may vary from the 'name' values in any given complex or species, for example, the LH1 (B890) from *Rps. acidophila* strain 7050 (spectra A, Figure 5) has its absorbance peak at 880 nm, the LH2 (B800-850; spectra C) has peaks at 800 nm and 860 nm, and the LH3 (B800-820) peaks at 800 nm and 830 nm. For comparison Figure 5 shows several other spectra, *Rs. rubrum* strain S1 RC-LH1 (B), is clearly very similar to that found in *Rps. acidophila* strain 7050 (spectra A). Spectra D, F and G show LH2 from *Rps. acidophila* strain 10050, *Rb. sphaeroides* strain 2.4.1 and *Rps. palustris* strain French respectively. The latter spectra is visibly different, the 800 nm peak dominating that at 850 nm in contrast to the other species. This complex is in fact a secondary type of LH2 complex (defined as a type II LH2) synthesised by *Rps. palustris* strain French at low irradiance levels, at higher light levels a complex with an absorption spectrum more closely resembling that of *Rb. sphaeroides* strain 2.4.1 (F) is produced (a type I LH2) (Hayashi *et al.*, 1982; Mantele *et al.*, 1988; Evans *et al.*, 1990).

The structural basis for these observed differences in absorption spectra, the environmental factors influencing their occurrence, and their importance in energy transfer are discussed in the following sections.

Structure and Function of Purple Bacterial Light-harvesting Complexes

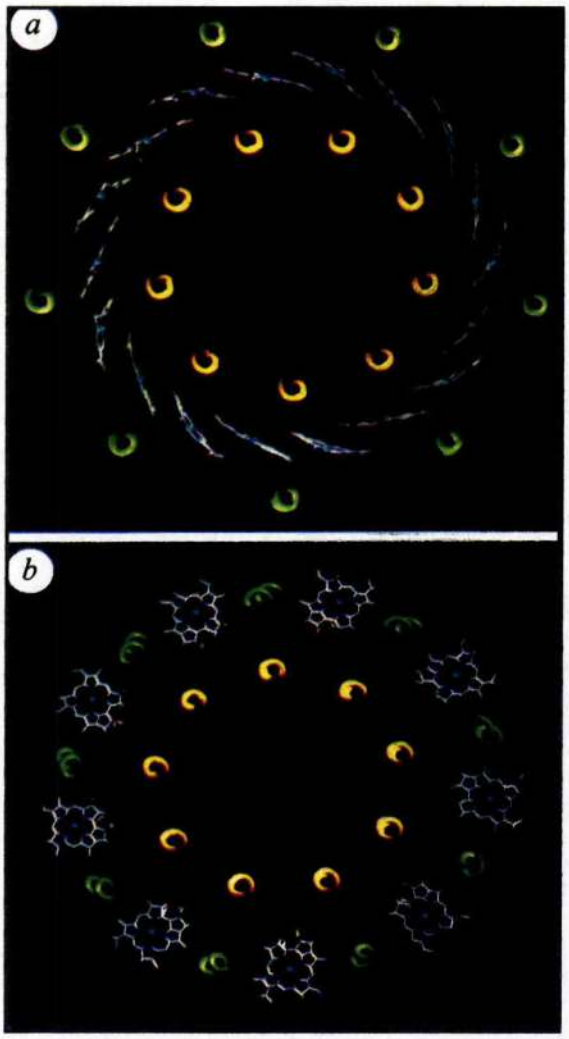
The structure of LH1, LH2 and LH3 proteins appears to be very similar, they all appear to consist of membrane spanning oligomeric assemblies of two polypeptides, α and β, which bind bacteriochlorophyll and carotenoid molecules (Aagard & Sistrom, 1972; Thornber *et al.*, 1978; Cogdell & Thornber, 1979; Zuber and Brunisholz, 1991). The α- and β-polypeptides are present in a 1:1 ratio in each complex (Cogdell *et al.*, 1980). In contrast the number of Bchl and carotenoid molecules bound appears to be complex and species specific. Although the number of Bchl molecules bound within a complex appears fixed, LH1 binding 2 Bchl, LH2 and LH3 binding 3 Bchl per α/β-polypeptide pair (Cogdell & Crofts, 1978; Cogdell & Thornber, 1979; Cogdell & Thornber, 1980), the number of Car molecules bound seems to be more variable (Cogdell, 1986; Thornber, 1986; Zuber, 1986). Thus per α/β-polypeptide pair LH1 from *Rs. rubrum* (Picorel, 1983; Evans *et al.*, 1988), *Chr. vinosum* (Cogdell & Thornber, 1979) and *Rps. acidophila* (Cogdell *et al.*, 1983) binds 2 Bchl and 1 Car, whereas LH1 from *Rb. sphaeroides* (Cogdell & Crofts, 1978; Cogdell & Thornber, 1980), *Rb. capsulatus* (Peters & Drews, 1983) and *Rv. gelatinosus* (Hawthornthwaite & Cogdell, 1993; Jirsakova & Reiss-Husson, 1993) bind 2 Bchl and 2 Car.

[2] Please note that in those spectra recorded from solubilised membranes (see Figure 5 legend), more than one LHC may be present. This is particularly noticable in spectra E where a shoulder can be seen at 890 nm representing the LH1 complex, however the predominant peaks are still representative of the named LHC.

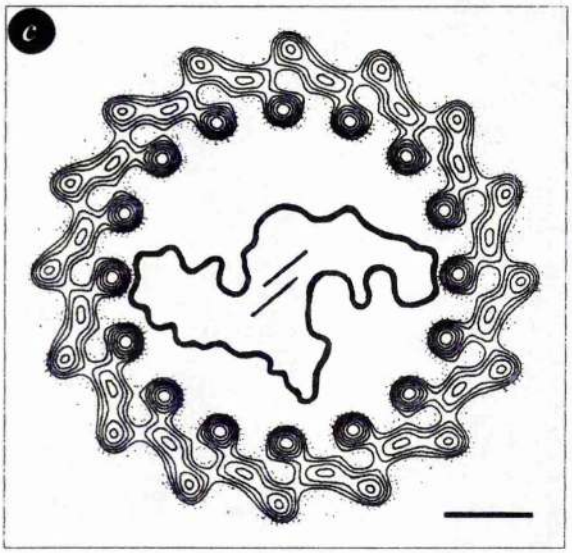
Figure 6

Images from crystallographic structures of purple bacterial light-harvesting complexes.

a **&** ***b*** : Two views of the 3D crystallographic model of LH2 from *Rps. acidophila* strain 10050. Adapted from McDermott *et al.*, 1995. The α–apoproteins are coloured yellow and the β-apoproteins green. Looking down on the complex from above ***a*** shows the ring of B850 Bchl *a* molecules, with their ring structures aligned perpendicular to the membrane surface. In ***b*** the B800 Bchl *a* molecules can be seen, located between the β-apoproteins, with their ring structures parallel to the plane of the membrane. The inner cylinder has a diameter of approximately 36 Å, whilst the outer is 68 Å in diameter.



c : Image processed projection map of LH1 from *Rs. rubrum* obtained from 2D crystals. Adapted from Karrasch *et al.*, 1995. The outer surface has a 116 Å diameter, the inner 68 Å. An approximate van der Waals projection of the LM subunits and single transmembrane α-helix of the H subunit from *Rps. viridis* (Deisenhofer *et al.*, 1985) has been placed arbitrarily within the LHC ring. The structure is a clearly similar to that of the the LH2 complex (*a* above), with a presumed inner ring of 16 α–apoproteins, outer ring of 16 β–apoproteins and the pigment molecules in between. Bar represents 20 Å.



Similarly LH2 and LH3 bind Car in a complex and species specific manner, LH2 from *Rb. capsulatus* (Shiozawa *et al.*, 1982) and *Chr. vinosum* (Cogdell & Thornber, 1979) appear to contain 3 Bchl and 1 Car per α/β- pair, as do the LH3 of *Chr. vinosum* (Thornber, 1970) and *Rps. acidophila* (Hawthornthwaite & Cogdell, 1991) and the type II LH2 of *Rps. palustris* (Evans *et al.*, 1990). In contrast the type I LH2 from *Rps. palustris* (Evans *et al.*, 1990) and LH2 from *Rb. sphaeroides* (Evans *et al.*, 1988) and *Rps. acidophila* (Hawthornthwaite & Cogdell, 1993) appear to contain 3 Bchl and 2 Car per α/β- pair.

The amino acid sequence of the apoproteins is highly conserved between species as illustrated by alignments of the α- and β-polypeptides from a range of photosynthetic purple bacterial species in Tables 1 and 2 (Following pages). The alignments highlight several conserved regions, suggesting that those amino acids have important roles to play in the functioning of LHCs. A comprehensive review of conserved regions plus biochemical and spectroscopic data from a range of LHCs was carried out by Zuber & Brunisholz (1991). These analyses provided a good picture of the structure and function of bacterial LHCs.However our understanding of LHC structure has improved considerably following the recent determination of the 3D crystallographic structure of LH2 from *Rps. acidophila* strain 10050 (McDermott *et al.*, 1995), and a low resolution 2D crystal structure of LH1 from *Rs. rubrum* S1 (Karrasch *et al.*, 1995). Some elements of these two stuctures are illustrated in Figure 6 (opposite).

It is clear that LH1 and LH2 have a very similar structure, as was suggested by their highly conserved polypeptide sequences and pigment composition (Zuber & Brunisholz, 1991). The overall structure of the LHCs can be described as follows. The α- and β-polypeptides form two concentric cylinders, the α-polypeptides form the inner cylinder and the β-polypeptides the outer cylinder, the pigment molecules are located between the two cylinders (McDermott *et al.*, 1995; Karrasch *et al.*, 1995). The β-polypeptides are tilted at an angle relative to the plane of the membrane, whilst the α-polypeptides are more perpendicular (McDermott *et al.*, 1995; Karrasch *et al.*, 1995).

The main difference between LH1 and LH2 concerns the number of α- and β-polypeptides which constitute the complex, the *Rps. acidophila* strain 10050 LH2 structure has nine α/β-pairs (McDermott *et al.*, 1995), whilst the *Rs. rubrum* LH1 complex appears to have 16 pairs (Karrasch *et al.*, 1995). The larger ring structure of the LH1 complex is a necessary requirement for enclosure of the RC complex which is known to be located within the LH1 ring (Sauer & Austin, 1978; Miller, 1982). A secondary difference, mentioned above, concerns the number of bound Bchl molecules in each complex, LH1 binding 2 Bchl and LH2 binding 3 Bchl per α/β-polypeptide pair (Cogdell & Thornber, 1979; Drews, 1985). In the *Rps. acidophila* strain 10050 LH2 structure the Bchl molecules form 2 rings, the ring situated towards the periplasmic side of the complex consists of 18 overlapping Bchls arranged with their porphyrin rings perpendicular to the plane of the membrane, in the ring located nearer the cytoplasmic side of the complex the 9 Bchls lie with their porphyrin rings in the plane of the membrane (McDermott *et al.*, 1995).

Table 1 : Alpha Polypeptides

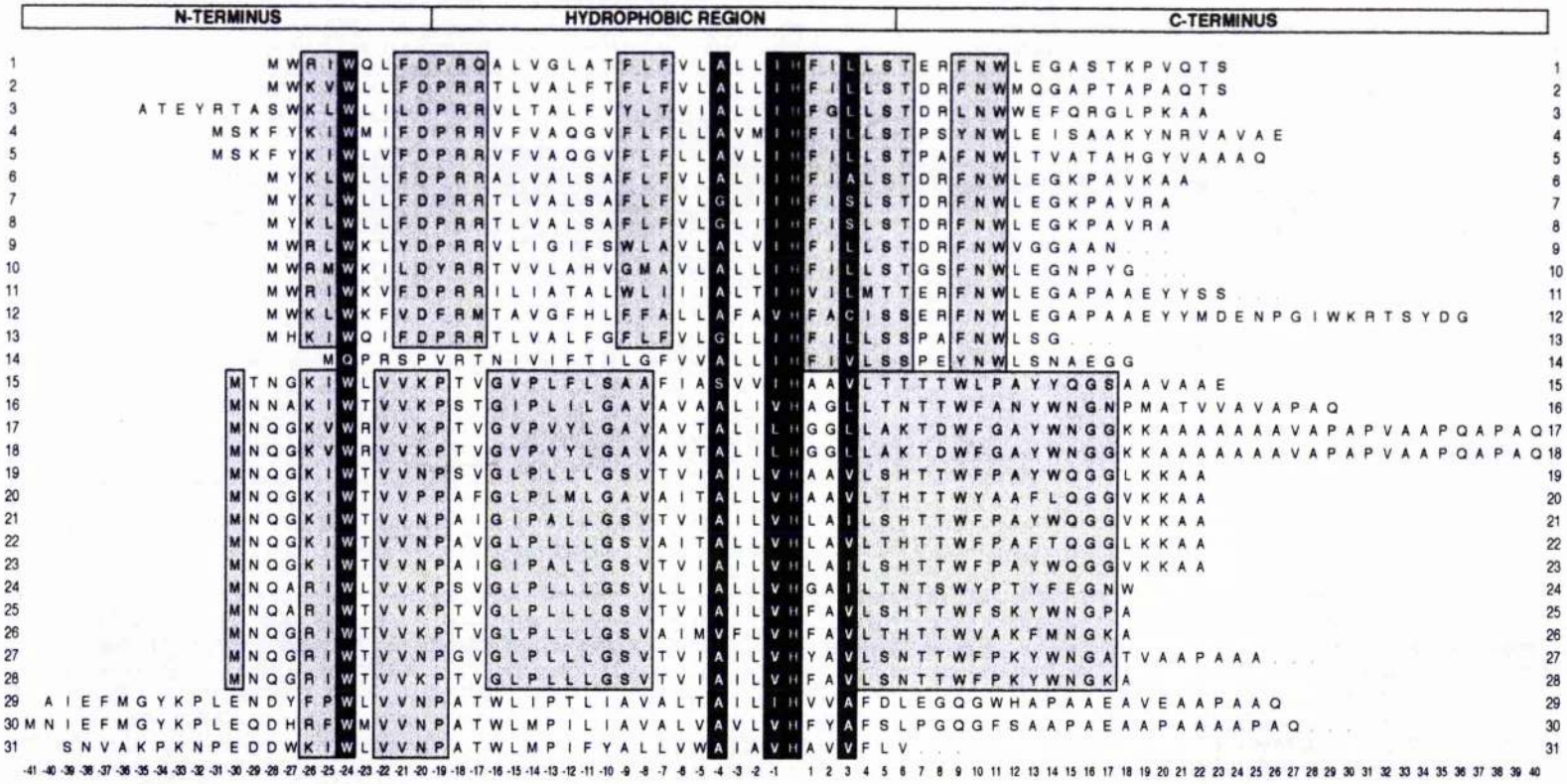


Table 1 Similarities in the primary amino acid structures of Purple Bacterial α-polypeptides. Taken from (Gall, 1994), data were collated from Zuber and Brunisholz, 1986; Zuber, 1990; Wagner-Huber et al., 1992. The boxes indicate consensus sequences. The histidine (H) residues are regarded as the binding site for bacteriochloropyll. The core polypeptides (1-13) and peripheral (14-31) are aligned using the histidine residues. Dotted sequences indicate incomplete primary sequence information. *1. Rs. rubrum B890-α 2. Rp. marina B880-α 3. Rp. viridis 1015-α 4. Rb. sphaeroides B870-α 5. Rb. capsulatus B870-α 6. Rp. acidophila Ac7050 B880-α 7. Rp. acidophila Ac7750 B880-α 8. Rp. acidophila Ac10050 B880-α 9. E. halophila B890α1 10. E. halophila B890-α2 11. E. halochloris MWR 12. E. halochloris MKW 13. Chr. vinosum B890- α 14. Cf. aurantiacus J-10-fl B806-866-α 15. Rb. sphaeroides B800-850-α 16. Rb. capsulatus B800-850-α 17. Rv. gelatinosus DSM149 B800-850-α 18. Rv. gelatinosus DSM151 B800-850-α 19. Rp. acidophila Ac7050 B800-850-α 20. Rp. acidophila Ac7050 B800-820-α 21. Rp. acidophila Ac7750 B800-850-α 22. Rp. acidophila Ac7750 B800-820-α 23. Rp. acidophila Ac10050 B800-850-α 24. E. halophila B800-850-α 25. Rp. palustris 2.6.1 B800-850-α1 26. Rp. palustris 2.6.1 B800-850-α2 27. Rp. palustris 2.6.1 B-800-850-α3 28. Rp. palustris 2.6.1 B800-850-α4 29. Chr. vinosum B800-850-α1 30. Chr. vinosum B800-820-α 31. Chr. vinosum B800-850-α2*

Table 2 : Beta Polypeptides

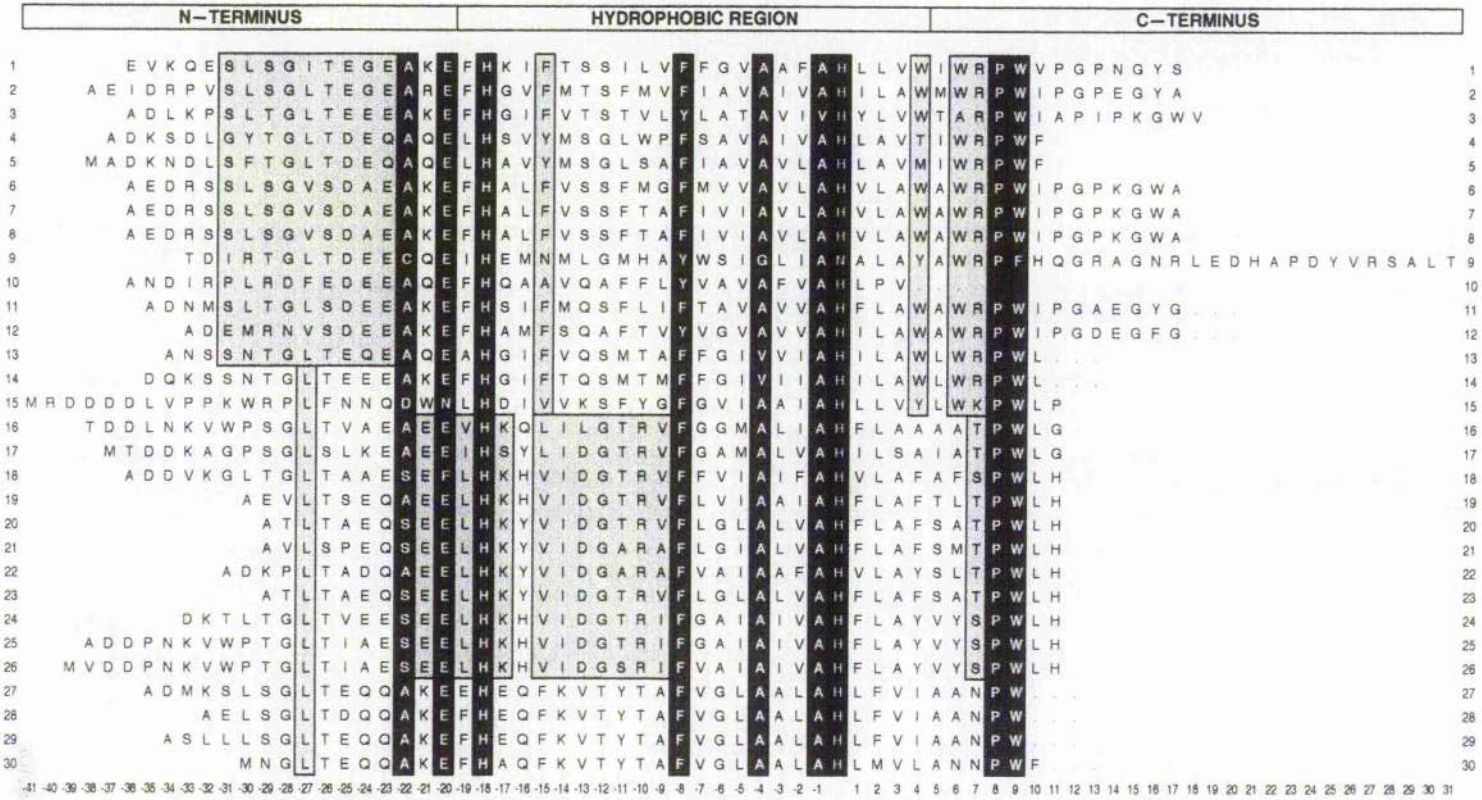


Table 2 Similarities in the primary amino acid structures of Purple Bacterial β-polypeptides. Taken from (Gall, 1994), data were collated from Zuber and Brunisholz, 1986; Zuber, 1990; Wagner-Huber et al., 1992. The boxes indicate consensus sequences. The histidine (H) residues are regarded as the binding site for bacteriochloropyll. The core polypeptides (1-13) and peripheral (14-30) are aligned using the histidine residues. Dotted sequences indicate incomplete primary sequence information. *1. Rs. rubrum B890–β 2. Rp. marina B880-β 3. Rp. viridis 1015-β 4. Rb. sphaeroides B870-β 5. Rb. capsulatus B870-β 6. Rp. acidophila Ac7050 B880-β 7. Rp. acidophila Ac7750 B880-β 8. Rp. acidophila Ac10050 B880-β 9. E. halochloris TDI 10. E. halochloris AND 11. E. halophila B890-β1 12. E. halophila B890-β2 13. Chr. vinosum B890-β1 14. Chr. vinosum B890-β2 15 Cf. aurantiacus J-10-fl B806-866-β 16 Rb. sphaeroides B800-850-β 17. Rb. capsulatus B800-850-β 18. Rp. acidophila Ac7050 B800-850-β 19 Rp. acidophila Ac7050 B800-820-β 20 Rp. acidophila Ac7550 B800-850-β 21. Rp. acidophila Ac7550 B800-820-β2 22. Rp. acidophila Ac7750 B800-820-β1 23 Rp acidophila Ac10050 B800-850-β 24. Rp. palustris 2.6.1 B800-850-β1 25. Rp. palustris 2.6.1 B800-850-β2 26. Rp. palustris 2.6.1 B800-850-β3 27. Chr. vinosum B800-850-β1 28. Chr. vinosum B800-820-β1 29. Chr. vinosum B800-850-β2 29. Chr. vinosum B800-820-β2*

Since the LH1 complex only binds 18 Bchl molecules, and since these have been shown to interact in a manner similar to the group of 18 in the LH2 complex (Cogdell & Scheer, 1985) it is seems likely that they also form an overlapping ring.

The *Rps. acidophila* LH2 appears to bind 2 Car molecules per α/β-pair (Hawthornthwaite & Cogdell, 1993), however only one Car molecule was visible in the initial model of the crystallographic structure (McDermott *et al.*, 1995), although 2 appear to be present in a refined model (Prof. R.J. Cogdell, personal communication). The single carotenoid visible in the published *Rps. acidophila* LH2 structure spans the depth of the complex in an 's' shape (all-trans configuration), interacting with a 'set' of 3 Bchl molecules (bound to one α/β-pair), and with adjacent α- and β-polypeptides (McDermott *et al.*, 1995). Thus carotenoid is not only involved in light harvesting and energy transfer, but potentially in stabilising the structure of the complex (Jirsakova & Reiss-Husson, 1994; McDermott *et al.*, 1995). Given the important role that the carotenoid(s) may play in LHC stability it seems plausible that the Car will be found in a similar position in LH1 when a high resolution structure becomes available. The structure of LH1 can not be described in any greater depth at present because of the low resolution so far obtained in structural studies, however much of the structural detail described below for the LH2 complex is likely to be relevant to the LH1 complex also.

Polypeptide topography and interactions in LH2

The main features of the *Rps. acidophila* strain 10050 LH2 are described below with reference to the alignments of LHC apoproteins in Tables 1 & 2, the *Rps. acidophila* LH2 α- and β-polypeptides are positioned at number 23 in both tables.

As was originally suggested (Brunisholz *et al.*, 1984), the α- and β-polypeptides have a tri-partite domain structure with a central hydrophobic membrane spanning α–helix, the three domains are clearly marked on table 1 and 2. The N and C terminii turn to lie flat against the membrane and contain hydrophobic residues which anchor the terminii to the surface, these anchors appear to hold the helices at the correct angle within the membrane (McDermott *et al.*, 1995). In the N-terminus of the β-polypeptide (Table 2) three such hyrophobic residues occur, an Alanine (-29), a Leucine (-27, highly conserved amongst LH2 and LH3 complexes) and a second Alanine (-25). The C-terminus is shorter, but also contains three hydrophobic residues - Proline, Tryptophan and Leucine (8, 9, 10) which are highly conserved amongst both LH2, LH3 and LH1 β-polypeptides. The C-terminus of the α-polypeptide (Table 1) forms an amphipathic helix, but again hydrophobic residues anchor the helix to the membrane surface, in this case Tryptophan, Phenylalanine, Tyrosine and a second Tryptophan (9, 10, 13, 14). These four residues are well conserved amongst LH2 α-polypeptides. The N-terminus of the α-polypeptide is similar to the C-terminus, forming an amphipathic helix with Isolcucine (-25) Tryptophan (-24) and Valine (-22) anchoring it, but differs in that the actual N-terminus (formylmethionine) is buried within the membrane where it coordinates with the central magnesium atom of the 800 nm absorbing Bchl (McDermott *et al.*, 1995). The anchoring residues of the N-terminus are highly conserved amongst LH2 and LH3 α-polypeptides, and to a lesser extent in the LH1 N-terminus (Table 1). The N-terminal Methione is also highly conserved, though this may

to some extent result from the preference for Methionine as the start codon of many genes in bacteria.

There are no direct helix-helix interactions between the α- and β-apoproteins within the membrane spanning region, however indirect interactions occur via the pigment molecules and water (McDermott *et al.*, 1995). In the N-terminii interactions occur between pairs of α- and β-polypeptides, but not between them. In contrast at the C-terminii interactions occur both between α/β-pairs and between adjacent pairs, in particular the four aromatic α-polypeptide residues Tryptophan (9), Phenylalanine (10), Tyrosine(13) and Tryptophan (14) bind the apoproteins together via hydrogen bonds and hydrophobic interactions (McDermott *et al.*, 1995). These are the same residues that anchor the α-polypeptide C-terminus, this further explains why those residues are so highly conserved in LH2 α-polypeptides from several species (Table 1).

Pigment-protein interactions in LH2

The most striking feature in tables 1 and 2, in many core and peripheral LHCs, is the presence of conserved histidine residues towards the C-terminal (periplasmic) side of the hydrophobic membrane spanning region (point 0 in Table 1 and 2). Resonance Raman studies suggested that these histidine residues bind the 850 nm absorbing Bchl molecules (Robert & Lutz, 1985), and circular dichroism studies suggested that Bchl molecules formed a dimer (Sauer & Austin, 1978). In the *Rps. acidophila* LH2 structure this has been shown to be essentially correct, the 2 Bchl molecules bound to an α/β-pair do indeed dimerise, but the dimers lie so close to each other that they can essentially be considered as a ring of Bchl molecules (McDermott *et al.*, 1995) (this is discussed in greater detail below). A number of other residues also interact with the 850 nm absorbing Bchls, forming a hydrophobic pocket around them, on the α-polypeptide the residues Alanine (-4), Isoleucine (3) and Tryptophan (14) and on the β-polypeptide residues Phenylalanine (-8), Alanine (-4) and Tryptophan (9) all make contacts (McDermott *et al.*, 1995). The latter three residues are all highly conserved amongst LHCs, as is the α-Alanine (-4; Table 1 and 2). Two residues form hydrogen bonds to the 850 nm absorbing Bchls, α-Tryptophan (14) forms a bond with the α-Histidine coordinated Bchl, and α-Tyrosine (13) of an *adjacent* α-polypeptide forms a hydrogen bond to the β-Histidine coordinated Bchl (McDermott *et al.*, 1995).

A second conserved histidine located towards the N-terminal or cytoplasmic side of the membrane spanning region of the β-polypeptide is visible in table 2 (-18, shown in bold type). This was postulated as the ligand for the 800 nm absorbing Bchl (Zuber, 1985; Zuber *et al.*, 1987), but Resonance Raman studies suggested that other residues may be involved (Robert & Lutz, 1985). The latter assertion has proved correct, as previously mentioned the N-terminus of the α–apoprotein coordinates with the Magnesium at the centre of the 800 nm Bchl, but the conserved histidine also interacts with the Bchl, via a water molecule located between the Histidine residue and Bchl molecule (McDermott *et al.*, 1995). A hydrogen bond is formed between the 800 nm Bchl and and the β– residue Arginine (-10), this residue is part of a group of highly conserved LH2 residues.

The observed Car spans the depth of the membrane, one end is associated with

α–polypeptide residues Lysine (-26) and Threonine (-23), the other end is associated with the 850 nm Bchl binding Histidine (0) of an adjacent α–polypeptide (McDermott *et al.*, 1995). On its path through the complex the Car makes several van der Waals contacts with adjacent α- and β-polypeptides, and with both the 800 nm and 850 nm absorbing Bchl molecules (McDermott *et al.*, 1995).

One of the more interesting questions regarding LHC structure concerns the manner in which small changes in polypeptide sequences give rise to profound differences in the NIR absorption peak of the 850 nm absorbing Bchls. On the basis of sequence homologies, and deviations therefrom, it was suggested that changes from the two highly conserved α-polypeptide residues Tyrosine (13) and Tryptophan (14) were responsible for the blue shift in Bchl absorbance found in LH3 (820 nm) (Brunisholz *et al.*, 1987; Bissig *et al.*, 1988; Brunisholz & Zuber, 1988). More specifically, these residues were shown to be different in the LH3 α-polypeptides of *Rps. acidophila* strain 7050 (Table 1 row 20) and strain 7750 (Table 1 row 22), where they were found to have changed to Phenylalanine plus Leucine (7050) and Phenylalanine plus Threonine (7750). Furthermore site directed mutagenesis of these residues in *Rb. sphaeroides* LH2 causes a blue shift similar to that observed in Rps. acidophila LH3 (Fowler *et al.*, 1992; Fowler *et al.*, 1994; Hess *et al.*, 1994). Given the roles that these residues play in the LH2 structure of *Rps. acidophila* strain 10050 (they form H-bonds to the 850 nm Bchls and to other apoproteins, holding the complex together, and anchor the C-terminus to the membrane surface - as described above), it is perhaps not surprising that they have a large effect on the functional properties of LH2 and LH3.

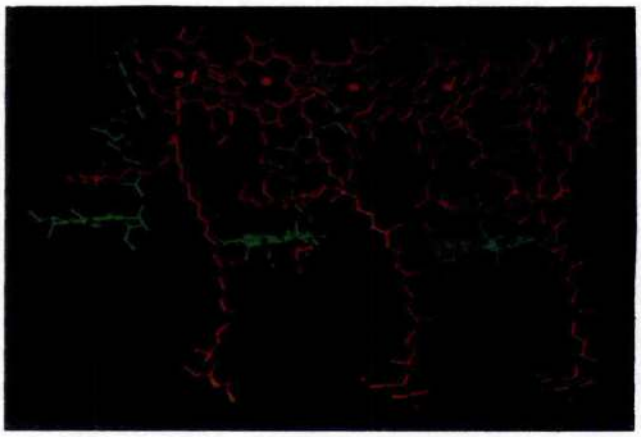
A similar set of highly conserved residues is found in the c-terminus of the LH1 α-polypeptides (table 1), these are Phenylalanine (9), Asparagine (10) and Tryptophan (11) (Zuber & Brunisholz, 1991). Since those LH1 α-polypeptides which deviate from this belong to complexes which have distinctive absorption maxima (e.g. *Rps. viridis* LH1 α-polypeptide, Table 1 Row 3, absorption maxima at 1015 nm), it seems likely that these residues play a similar role in LH1.

Pigment interactions in LH2

It was predicted on the basis of circular dichroism studies that the 850 nm absorbing Bchl molecules would form dimers *in vivo* (Sauer & Austin, 1978), this is indeed the case in the *Rps. acidophila* strain 10050 LH2 structure, but the dimers also lie close enough to each other to form van der Waals contacts, and should be considered as a 'ring' of Bchl molecules (McDermott *et al.*, 1995) (see Figure 6). As previously stated the 18 850 nm absorbing Bchls are arranged with their ring structures perpendicular to the membrane, whilst the 9 800 nm absorbing Bchl molecules lie in the plane of the membrane. There is an interaction between the 800 nm and β-coordinated 850 nm absorbing Bchl molecules in that their phytyl chains twist together and pass across the face of the opposing porphyrin ring, close enough to form van der Waals contacts (McDermott *et al.*, 1995).

Figure 7

Arrangement of pigments in the *Rps. acidophila* stain 10050 LH2 structure model. The pigment molecules bound to 3 α/β-polypeptide pairs are illustrated below.
Adapted from McDermott *et al.*, 1995.



Note: The carotenoid is coloured yellow; B800 Bchl green; B850 bound to the β-apoproteins red and those bound to the α-apoproteins orange. The carotenoid can be seen to extend through the depth of the complex, passing close to both a B800 Bchl and B850 Bchl. Phytyl chains of the B800 Bchl molecules (green) twist up to coordinate with the α-liganded B850 molecules (red), whilst those of the β-liganded B850 molecules (orange) pass down and across the ring structures of the B800 molecules. Only one carotenoid is pictured, but there are two present in a refined model of the complex (see main text).

Each Car molecule also forms van der Waals contacts with both an 800 nm and 850 nm Bchl bound to each polypeptide pair, such that all the pigment molecules within LH2 can be considered to be interconnected (McDermott *et al.*, 1995). Figure 7 (opposite) illustrates the arrangement of pigments in the LH2 structure.

Energy transfer within and through LHCs to the RC

The absorption peaks of the various LHCs hint that transfer of captured light energy is an energetically "downhill" process, an exciton being passed to progressively longer wavelength absorbing pigment molecules until it reaches the RC (Monger & Parson, 1977; Pearlstein, 1982; van Grondelle, 1985). Light harvesting is carried out in the NIR (Q_y), UV (~380 nm, soret band) and visible (~590 nm, Q_x) regions by Bchl, and in the visible region of the spectrum (around 450-550 nm) by Car, which acts as an accessory light-harvesting pigment as well as playing a photoprotective role (Cogdell, 1985). Pigment molecules located in any of the LHC complexes present may harvest light and thus an exciton may begin its journey to the RC at a number of points along the transfer pathway, for the purpose of describing the pathway I will assume that excitons follow the generalised path Car > B800 > B850 > B890 > RC (where B800 indicates the 800 nm absorbing Bchl), it is however, important to bear in mind that Car molecules may also transfer energy to B850 and B890.

Exciton transfer from Car to Bchl occurs with variable efficiency (Goedheer, 1959; Nishimura & Takamiya, 1965), depending on the apoprotein structure and type of Car bound to the complex (Boucher *et al.*, 1977; Cogdell *et al.*, 1981; Hayashi *et al.*, 1989; Noguchi *et al.*, 1990). Certain species e.g. *Rb. sphaeroides* and *Rv. gelatinosus*, appear to have much higher transfer efficiencies (70-100%), than species such as *Rs. rubrum*, *Chr. vinosum* and *Rps. palustris* (30-40%) (Kramer *et al.*, 1984; Evans *et al.*, 1990; Noguchi *et al.*, 1990). The intimate contacts between Car and Bchl in the LH2 complex (McDermott *et al.*, 1995) allow transfer to the B800 to occur very quickly (Gillbro *et al.*, 1988; Gillbro & Cogdell, 1989). Exciton transfer between B800 and B850 absorbing molecules occurs in around 0.7 ps (van Grondelle *et al.*, 1994). In LH3 and type II LH2 complexes the transfer from B800 to B820 or B850 appears to be faster than in type I LH2 complexes (Bergström *et al.*, 1988). Upon transfer of an exciton to a B850 the energy is rapidly delocalised over the complete ring of 18 B850 molecules, from any point on the ring transfer can occur to the B890 ring within LH1. The transfer between complexes occurs on a longer time scale of 5-20 ps (van Grondelle *et al.*, 1994), the exact transfer time being dependent on the topological arrangement of LHCs within the membrane - this varies with environmental conditions (discussed in the next section).

The ring-like structure of LHCs, and the positioning of Bchl molecules to allow delocalisation of exciton energy, provides an elegant method of arranging LHCs within the ICM, such that transfer can occur between any adjacent LH2 and LH1 complex regardless of their orientation (Pearlstein, 1985; McDermott *et al.*, 1995).

Plasticity of the Photosynthetic Unit.

Many purple bacteria synthesise extensive amounts of lipid bilayer by invagination of the cytoplasmic membrane, this is known as the intracytoplasmic membrane, into which the photosynthetic apparatus is inserted. These invaginations may take the form of vesicles (e.g. *Rb. capsulatus*, *Rb. sphaeroides*, *Rs. rubrum*, *Chr. vinosum*), lamellae (e.g. *Rps. acidophila*, *Rps. palustris*, *Rps. viridis*), stacks (e.g. *Rs. photometricum*, *Rs. molischianum*) or tubules (e.g. *Thiocapsia pfennigii*) (Pfennig & Trüper, 1973). Some species however appear to have relatively weak development of the ICM (e.g. *Rv. gelatinosus*, *Rc. purpureus*, *Rc. tenue*) suggesting that much of their photosynthetic apparatus may be located in the cytoplasmic membrane (Pfennig & Trüper, 1973; Oelze & Drews, 1972; Zuber & Brunisholz, 1991).

Both the amount of ICM and components of the photosynthetic apparatus vary in response to environmental factors, the two most important factors being oxygen tension and light intensity (Ohad and Drews, 1982). The exact response to these factors is species specific, *Rb. sphaeroides* synthesises no ICM or photosynthetic apparatus in the presence of oxygen, whilst *Rb. capsulatus* appears to produce small quantities even under aerobic conditions (Klug, 1993), however only under anoxygenic conditions are they produced in large quantities (Drews, 1991). It should be noted however that small amounts of photosynthetic apparatus are synthesised in the dark as long as the oxygen tension is low and the bacteria are able to grow chemotrophically (Cohen-Bazire & Kunisawa, 1960).

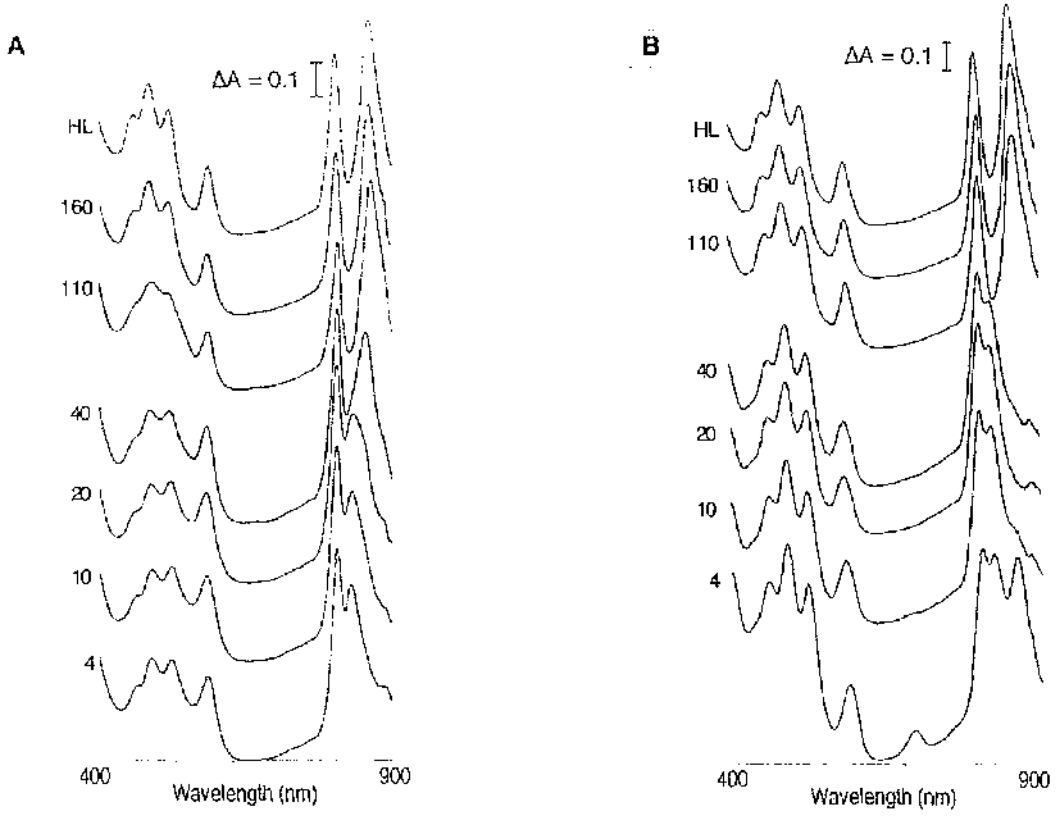
One exception to this is the species *Rb. sulfidophilus*[3] which synthesises large amounts of ICM and LHCs under both anaerobic and aerobic dark conditions (Hagemann *et al.*, 1995). Further exceptions can be found in the quasi-photosynthetic bacteria, species such as *Ro. denitrificans* and *Er. longus* appear to synthesise small amounts of photosynthetic apparatus under fully aerobic conditions (Shiba & Simidu, 1982; Harishima & Nakada, 1983; Shiba, 1987). However quasi-photosynthetic bacteria are unable to utilise photosynthesis as a major source of energy and thus represent a special case (Gest, 1993), it is clear however that their photosynthetic apparatus is highly homologous to that of purple photosynthetic bacteria (Shimada *et al.*, 1985; Liebetanz *et al.*, 1991) but regulated in a different manner, thus they represent an interesting comparison to purple photosynthetic bacteria.

Once anaerobic conditions have been established the amount of ICM, number of RC-LH1 aggregates, and also the presence of, and relative amounts (to RC-LH1) of the peripheral LHCs (LH2 and LH3) vary according to the incident light intensity (Aagard & Sistrom, 1972; Takemoto & Huang Kao, 1977; Drews & Oelze, 1981; Chory & Kaplan,1983). The synthesis of these photosynthetic components being inversely proportional to the light intensity, i.e. the relative size of a PU and the number of PUs increase as the light intensity is decreased (Takemoto & Huang Kao, 1977).

[3] A proposal has recently been made to rename this species *Rhodovulum sulfidophilus* (Hiraishi & Ueda, 1994).

Figure 8

Absorbance spectra of solubilised chromatophore membranes from *Rps. acidophila* strains 7050 and 7750 cultured under varying conditions. Adapted from Gardiner *et al.*, 1993.



A Strain 7050 cultured under a range of light-intensities.

B Strain 7750 cultured under the same range of light-intensities.

C Strain 7750 cultured under a range of temperatures whilst the light intensity is held at ~50 μmol s m^{-2}

Note: Absorbance spectra recorded from solubilised membranes prepared as described in Gardiner *et al.*, 1993. The topmost spectra in each frame represents a 'benchmark' culture grown under high light (HL, ~160 μmol s m^{-2}) at 30°C. In **A** and **B** the numbers to the left of the spectra represent the light intensity under which the bacteria were cultured (μmol s m^{-2}), in **C** these represent the Temperature in °C.

The exact response to changing irradiance levels is species specific, a species such as *Rs. rubrum* which synthesises only a B875 core complex and hence has a PU of a fixed size, will synthesise more ICM and more PUs to compensate for a reduction in irradiance (Oelze & Drews, 1972; Aagard & Sistrom, 1972). In contrast species like *Rb. sphaeroides* and *Rb. capsulatus* which produce a peripheral LH2 are able not only to alter their ICM structure and number of PUs, but also to vary the size of the PU by synthesising variable proportions of LH2 (Takemoto & Huang Kao, 1977). A further type of response is found in species such as *Rps. acidophila* and *Rps. palustris* where under conditions of very low irradiance an LH3 complex is produced (Hawthornthwaite & Cogdell, 1991; Tadros, 1990).

Figure 8 (opposite) illustrates the response of *Rps. acidophila* strains 7050 and 7750 when cultured under a selection of growth conditions. In frame A absorbance spectra of strain 7050 cultured under progressively lower light intensities are displayed, it can be seen that there is a gradual shift from production of LH2 (B800-850) to production LH3 (B800-820) as the light intensity is lowered. This strain also alters the type of Car which it incorporates into its LHCs, as can be seen if one examines the triplet carotenoid peaks of the spectra in frame A. This change in Car is visible by eye, when cultured under high light intensity the cells are a red-brown colour (Rhodopin and Rhodopin-glucoside predominate), whilst under low light they are a deep purple colour (Rhodopinal and Rhodopinal-glucoside are incorporated) (Heinmeyer & Schmidt, 1983; Cogdell *et al.*, 1983). Strain 7750 shows a similar responses to changes in light-intensity (frame B, Figure 8), synthesising LH3 (B800-820) at low light intensities. At the very lowest light intensity strain 7750 appeared to synthesise LH2 as well as LH3, the reasons for this remain obscure (Gardiner *et al.*, 1993). A second difference between the two strains concerns the change in Car, strain 7750 also synthesises a different Car at low light intensity, but in this case there is no colour change (Rhodopin-glucoside is replaced by Rhodopin at low light) (Cogdell *et al.*, 1983). Strain 7750 also shows a response to temperature, this can be visualised in frame C of Figure 8. At moderate light intensities (~50 μmol s m^{-2}) when the culture temperature is reduced, LH3 replaces LH2 in a similar manner to that seen with decreasing light intensity (Gardiner et al., 1993).

These examples illustrate some of the ways in which purple bacteria modify the size and character of their PSU in response to environmental conditions, in the following sections I will describe what is known about the molecular biology underpinning these changes.

Molecular Biology of Purple Photosynthetic Bacteria

Purple photosynthetic bacteria form part of a group of bacteria which includes many non-photosynthetic species, including well studied species such as *Agrobacterium*, *Rhizobium*, and the ubiquitous *E. coli* ('Proteobacteria', see the section "Taxonomy of the purple photosynthetic bacteria", page 2). Like these other species their nucleic acids are relatively easy to isolate and manipulate, allowing identification of photosynthetic unit components in photosynthetic bacteria and their manipulation for the purposes of study. The study of purple bacterial molecular biology has enhanced and extended biochemical, biophysical and structural studies of photosynthesis by way of mutagenesis studies, which provide a way of probing the structure and function of proteins. Site directed mutagenesis of RC and LHC proteins are excellent examples of how molecular techniques can be used to analyse the function of proteins (Youvan *et al.*, 1985; Farchaus & Oesterhelt, 1989; Nagarajan *et al.*, 1990; Paddock *et al.*, 1989; Takahashi *et al.*, 1990; Lausscrmair & Oesterhelt, 1992; Jones *et al.*, 1992; Ermler *et al.*, 1994; Fowler *et al.*, 1992; Fowler *et al.*, 1994; Hess *et al.*, 1994).

Furthermore the strategies employed by photosynthetic bacteria for gene regulation appear to be similar to those found in non-photosynthetic species (see following sections), thus study of photosynthesis gene regulation increases our understanding of prokaryotic gene regulation in general.

Of further interest are the strategies which photosynthetic bacteria use to adapt optimise their photosynthetic capacity, and the way in which they adapt to a changing environment. Since many of the structural elements involved in photosynthesis have been conserved through the evolution of a wide range of photosynthetic organisms, from bacteria and algae through to higher plants, it seems possible that they will adopt similar patterns of gene regulation and expression. This seems especially likely when one considers the fact that the chloroplast organelles of higher plants are essentially symbiotic photosynthetic bacteria (Zablen *et al.*, 1975; Bonen & Doolittle, 1975).

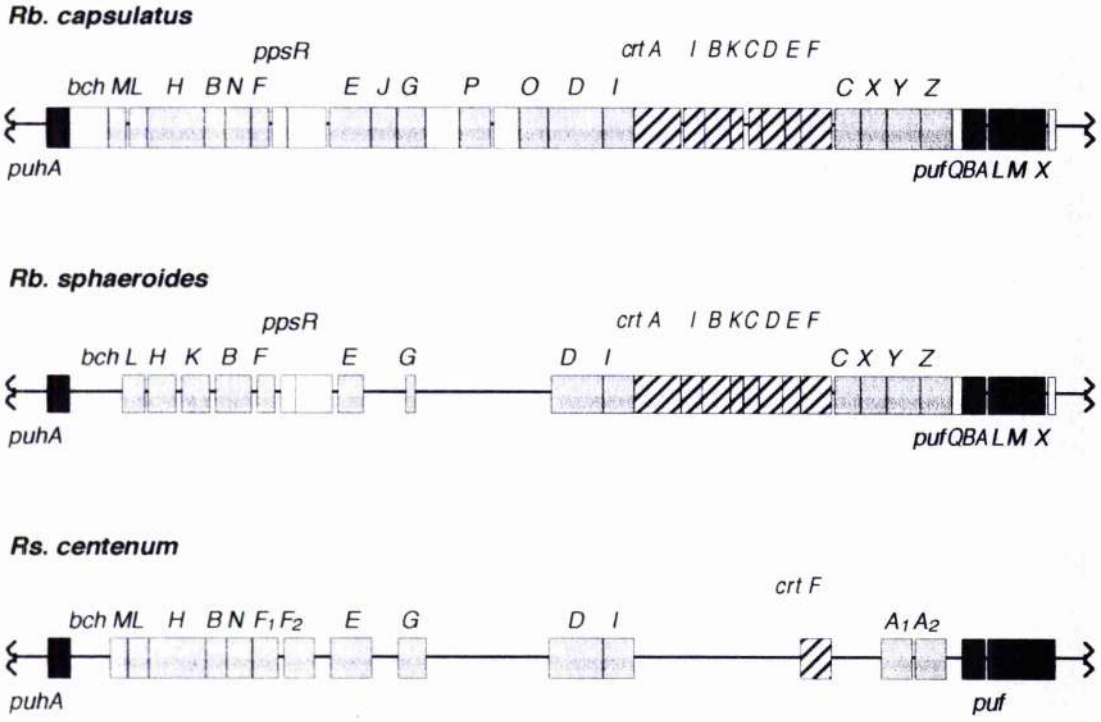
Two species of purple photosynthetic bacteria have been studied extensively these are the related *Rb. sphaeroides* and *Rb. capsulatus*. The latter species has been the subject not only of studies involving photosynthesis genes, but also of genes involved in other metabolic processes, such as nitrogen fixation (recent reports include: Hubner *et al.*, 1993; Schuddekopf *et al.*, 1993; FosterHartnett *et al.*, 1993; Cullen *et al.*, 1994; FosterHartnett *et al.*, 1994a ; FosterHartnett *et al.*, 1994b).

Genome size and organisation in purple bacteria

Genomic mapping studies have been carried out for both *Rb. sphaeroides* and *Rb. capsulatus*, yielding estimated genome sizes of 3.8 Mb for *Rb. capsulatus* (Fonstein *et al.*, 1992; Fonstein & Haselkorn, 1993, Fonstein *et al.*, 1995) and 4.5 Mb for *Rb. sphaeroides* (Suwanto & Kaplan, 1989a). This compares with an estimate of 4.7 Mb for the genome of *E. coli* (Smith *et al.*, 1987). Interestingly *Rb. sphaeroides* has been found to possess two circular chromosomes of 3 Mb and 0.9 Mb (Suwanto & Kaplan, 1989b; Suwanto & Kaplan, 1992), in contrast to the single chromosome of *Rb. capsulatus* (Fonstein *et al.*, 1992; Fonstein & Haselkorn, 1993; Fonstein *et al.*, 1995).

Figure 9

A schematic representation of the photosynthetic gene clusters of *Rb. sphaeroides*, *Rb. capsulatus* and *Rs. centenum*. Adapted from Yildiz *et al.*, 1992. The diagrams are based on genetic and physical mapping data from several studies: for *Rb. sphaeroides* (Coomber *et al.*, 1990; McGlynn & Hunter, 1993); for *Rb. capsulatus* (Taylor *et al.*, 1983; Armstrong *et al.*, 1989; Zsebo & Hearst, 1984; Yang & Bauer, 1990; Burke *et al.*, 1993a; Bollivar *et al.*, 1994); and for *Rs. centenum* (Yildiz *et al.*, 1992).



Key

- Genes coding for RC-LH1 structural polypeptides
- Genes involved in bacteriochlorophyll biosynthesis
- Genes involved in carotenoid biosynthesis
- Genes of indeterminate or possibly regulatory function

Note: The distance between the puhA gene and the puf operon in the *Rs. centenum* PGC is estimated to be between 38.4 and 46.3 kb, this compares with a distance of ~45 kb in *Rb. capsulatus* and *Rb. sphaeroides* (Yildiz *et al.*, 1992). The PGC from *Rb. capsulatus* has now been DNA sequenced in full (Genbank AC: Z11165), hence the greater detail of the *Rb. capsulatus* diagram. The unshaded ORFs have been given temporary designations (Genbank AC: Z11165), and several have been studied by directed mutagenesis (Bollivar *et al.*, 1994). The diagram of the PGC from *Rb. sphaeroides* is based on mixture of DNA sequence information and genetic mapping (for references see top of page), and that from *Rs. centenum* on complementation studies (Yildiz *et al.*, 1992), thus these maps are of a lower resolution than that from *Rb. capsulatus*.

Multiple chromosomes have now been identified in several other Proteobacteria, namely *Brucella melitensis* (Michaux *et al.*, 1993), *Leptospira interrogans* (Zuerner *et al.*, 1993), *Agrobacterium tumifaciens* (Allardet-Servent *et al.*, 1993), and *Pseudomonas cepacia* (Cheng & Lessie, 1994). The second (smaller) chromosome from *A. tumifaciens* is also notable because it is linear rather than circular (Allardet-Servent *et al.*, 1993).

The genetic and physical mapping studies carried out on *Rb. sphaeroides* (Suwanto & Kaplan, 1989a; Suwanto & Kaplan, 1989b; Coomber *et al.*, 1990; Wu *et al.*, 1991) and *Rb. capsulatus* (Taylor *et al.*, 1983; Zsebo & Hearst, 1984; Fonstein *et al.*, 1992; Fonstein & Haselkorn, 1993; Fonstein *et al.*, 1995) have identified a region of the genome approximately 50 kb in size within which the majority of structural genes required for the construction of a PSU are clustered. This is termed the "Photosynthesis Gene Cluster" (PGC) (Marrs, 1981; Taylor *et al.*, 1983; Zsebo & Hearst, 1984), and is described below.

Organisation of purple bacterial photosynthesis genes

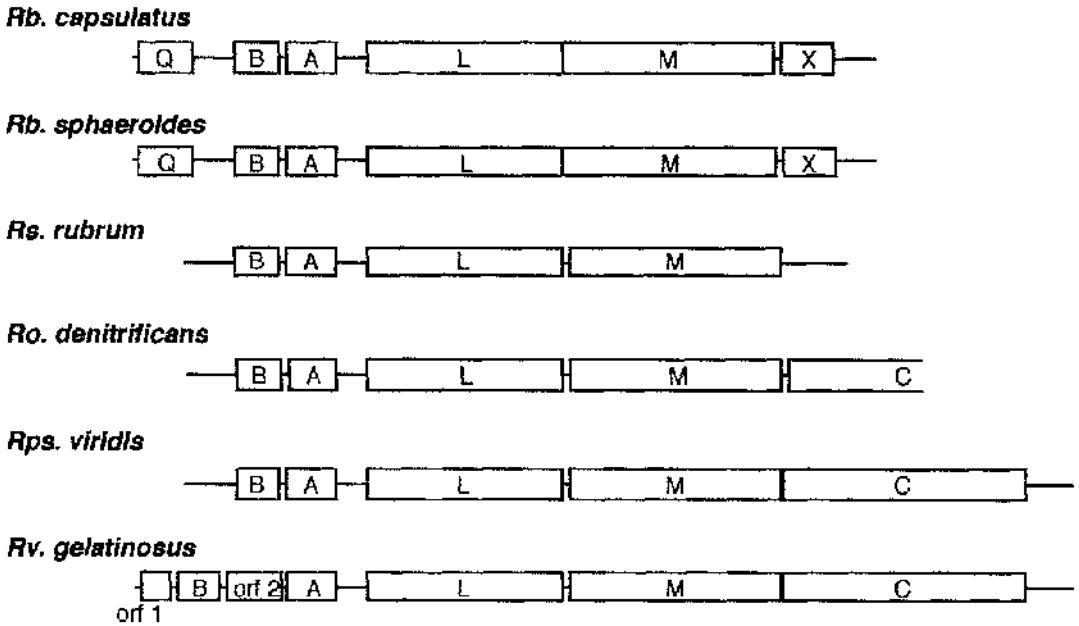
The PGC contains genes required for the Bchl and Car biosynthetic pathway flanked by the genes encoding the RC and LH1 polypeptides, and genes necessary for the assembly of the PSU. The PGC appears to be a highly conserved element in purple photosynthetic bacteria (Yildiz *et al.*, 1992). Figure 9 (opposite) illustrates the arrangement of genes within the PGC of *Rb. capsulatus*, *Rb. sphaeroides* and *Rs. centenum*. Although less information is available regarding the PGCs of *Rb. sphaeroides* and *Rs. centenum* it is clear that most of the genes are conserved and lie in similar positions. Many of the genes required for pigment biosynthesis are clustered together, flanked on one side by the puf operon containing the LH1 and RC structural genes, with the exception of puhA (the RC H subunit) which lies on the other side of the pigment biosynthesis genes. In *Rb. sphaeroides* it has been shown that genes required for pigment biosynthesis also lie outside of the PGC, such as hemT, hemA (Choudhary *et al.*, 1994) and hemF (Coomber *et al.*, 1992). However these genes appear to be involved in the earlier steps of Bchl and heme biosynthesis, thus it has been suggested that only those genes required for the latter stages of biosynthesis of PSU pigments are encoded within the PGC (Coomber *et al.*, 1992; Neidle & Kaplan, 1993).

The exact function of the gene products of the bch and crt genes has yet to be determined, a partial picture of the biosynthesis pathways and the role each protein plays in pigment biosynthesis has been outlined. For descriptions of Car and Bchl synthesis pathways in *Rhodobacter*, and the suggested function of the various crt and bch genes, the reader should consult the following (Armstrong *et al.*, 1989; Armstrong *et al.*, 1993; Armstrong, 1994; Lang *et al.*, 1994; Lang *et al.*, 1995; McGlynn & Hunter, 1993; Burke *et al.*, 1993a; Burke *et al.*, 1993b; Bollivar *et al.*, 1994).

The genes required for LH2 synthesis (excepting pigment biosynthesis genes) are clustered in the puc operon which is located ~18 kb away from the PGC in the direction of puhA, in *Rb. sphaeroides* (Suwanto & Kaplan, 1989a) and maps approximately 1.7 Mb away in *Rb. capsulatus* (Fonstein *et al.*, 1995).

Figure 10

Schematic representation of the puf operon from a range of purple bacteria. Adapted from Nagashima *et al.*, 1994. The original data come from several sequencing studies: *Rps. viridis* (Wiessner *et al.*, 1990), Rs. rubrum (Saegasser, 1992), *Ro. dinitrificans* (Liebetanz *et al.*, 1991), *Rv. gelatinosus* (Nagashima *et al.*, 1994), *Rb. capsulatus* (Youvan *et al.*, 1984; Bauer *et al.*, 1988; Bauer & Marrs, 1988) and *Rb. sphaeroides* (Williams *et al.*, 1983; Williams *et al.*, 1984; Williams *et al.*, 1986; Kiley *et al.*, 1987; Lee *et al.*, 1989).



Note : It can be seen that there is a conserved core of genes within the puf operons displayed above, the genes pufBALM being present in each species. These genes code for the LH1 β- and α-polypeptides and the RC L and M subunits. In addition to this core of genes those for pufC (RC cytochrome subunit), pufQ, pufX (for possible functions see main text), and genes whose function is at present unknown may also be found. A major deviation from this arrangement is seen in the green photosynthetic bacterium *Chloroflexus aurantiacus*, where the pufLM genes are found in one operon and the pufBA genes in a second operon (Shiozawa *et al.*, 1990; Watanabe *et al.*, 1995). In each species the genes have been shown to be transcribed from a promotor upstream of pufB (or pufQ where present), and thus transcribed from left to right in the diagram (Adams *et al.*, 1989; Hunter *et al.*, 1991; Wiessner *et al.*, 1990; Saegasser, 1992; Liebetanz *et al.*, 1991; Nagashima *et al.*, 1994). It should also be noted that *Ro. denitrificans* is a quasi-photosynthetic species which although capable of synthesising a photosynthetic apparatus is incapable of anaerobic photosynthesis (see the sections "Taxonomy of the purple photosynthetic bacteria" and "Energy transfer through the reaction centre and beyond" for a more detailed explanation).

It should be noted that the LH3 encoding genes have also been termed puc genes, since the amino acid sequences of LH3 apoproteins are extremely similar to those of LH2 (Tadros & Waterkamp, 1989; Tadros *et al.*, 1993; MacKenzie, 1990; Gardiner *et al.*, 1992). The puf and puc operons from a number of species have been cloned, and are described below.

The structure of the puf operon

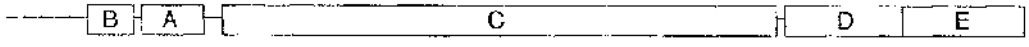
The arrangement of genes within the puf operon of the species *Rb. capsulatus*, *Rb. sphaeroides*, *Ro. denitrificans*, *Rps. viridis*, *Rs. rubrum* and *Rv. gelatinosus* is described in Figure 10 (opposite). In most species the genes pufBALM are present, coding for the LH1 α- and β-polypeptides and the RC L and M subunits. In those species which have a cytochrome associated with the RC the gene pufC is found to be present downstream of pufM (Wiessner *et al.*, 1990; Liebetanz *et al.*, 1991; Saegesser, 1992; Nagashima *et al.*, 1994). The two *Rhodobacter* species have two genes known as pufQ and pufX, flanking the basic pufBALM unit (Bauer *et al.*, 1988; Bauer & Marrs, 1988; Lee *et al.*, 1989a). The exact function of both these genes has yet to be elucidated, however pufQ appears to play a role in RC and LHC assembly and possibly acts as a carrier for Bchl (Bauer *et al.*, 1988; Hunter *et al.*, 1991; Gong *et al.*, 1994; Fidai *et al.*, 1994). The pufX gene product is a small protein with a possible membrane spanning helix (Lee *et al.* 1989a), that has been shown to be involved in cyclic electron transport (Lilburn *et al.*, 1992; Farchaus *et al.*, 1992). Thus it has been suggested that PufX plays a role in the transfer of ubiquinone and ubiquinol between the RC and the cytochrome bc_1 complex (Lilburn *et al.*, 1992; Farchaus *et al.*, 1992; Westerhuis *et al.*, 1993). A further study suggested that PufX was not necessary for cyclic electron transport in the absence of a LHC (McGlynn *et al.*, 1994), and thus may affect electron transport indirectly. However Karrasch *et al.* (1995) point out that since LH1 appears to completely encircle the RC it is not clear how the quinone molecules travel between the RC and the cytochrome bc_1 complex. It is clear therefore that PufX may well assist in the transfer of quinone to and from the RC 'through' LH1, possibly by preventing LH1 blocking the RC Q_B binding site (McGlynn *et al.*, 1994) - it would thus not be necessary in the absence of LHCs.

No pufQ or X genes have thus far been found associated with the puf operons of other purple bacteria, rather in the other species from which the puf operon has been cloned, that is: *Rps. viridis* (Wiessner *et al.*, 1990), *Rs. rubrum* (Saegasser, 1992), *Ro. dinitrificans* (Liebetanz *et al.*, 1991) and *Rv. gelatinosus* (Nagashima *et al.*, 1994); fragments of Bchl synthesis genes have been found to lie upstream of the puf operon - further evidence that the PGC may be a widely conserved structure in photosynthetic bacteria. As can be seen in Figure 10 the puf operon of *Rv. gelatinosus* does have extra genes within the puf operon, these two ORFs have no homology to either pufQ or X and their function is at present unknown (Nagashima *et al.*, 1994). Although ORF 2 shows some homology to pufB and may have a membrane spanning helix (Nagashima *et al.*, 1994).

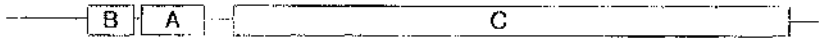
Figure 11

Schematic representation of the puc operon from a range of purple bacteria. The original data come from several sequencing studies: *Rb. capsulatus* (Youvan & Ismail, 1985; Tichy *et al.*, 1989); *Rb. sphaeroides* (Ashby *et al.*, 1987; Kiley & Kaplan, 1987; Lee *et al.*, 1989; Gibson *et al.*, 1992); *Rb. sulfidophilus* (Hagemann & Tadros, 1994); *Rps palustris* (Tadros & Waterkamp, 1989; Tadros *et al.*, 1993); *Rps. acidophila* (MacKenzie, 1990; Gardiner *et al.*, 1992; Barrett, 1995).

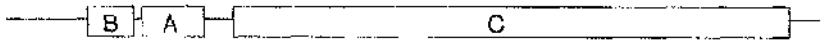
Rb. capsulatus



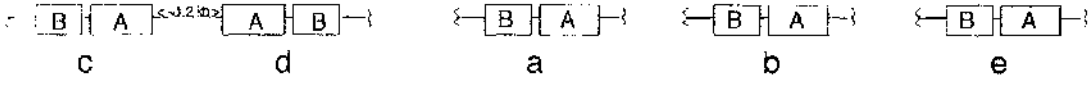
Rb. sphaeroides



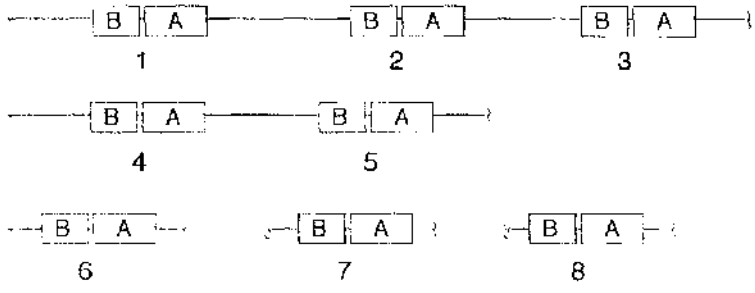
Rb. sulfidophilus



Rps. palustris



Rps. acidophila strain 7050



Note : Like the puf operon the puc operon has a conserved core of genes - the pucBA genes - these code for the LH2 β- and α-polypeptides. Those species which do not synthesise an LH3 complex appear to have a single copy of the puc operon, usually with the pucC gene lying downstream of pucA. The pucC gene has been shown to be essential for formation of LH2, PucC protein appears to be membrane bound (Tichy *et al.*, 1989), and it has been suggested that it plays a role in the assembly of LH2 complexes (Tichy *et al.*, 1991). In *Rps. acidophila* and *Rps. palustris* - species which synthesise an LH3 complex - multiple copies of the puc operon have been found, some copies of the operon are thought to code for LH3 β- and α-polypeptides in these species (Tadros & Waterkamp, 1989; Tadros *et al.*, 1993; MacKenzie, 1991; Gardiner *et al.*, 1992). *Rb. capsulatus* has further genes downstream of pucC, pucD and E. The pucE gene product is the LH2 γ polypeptide, which appears to stabilise the LH2 complex (Tichy *et al.*, 1991), the function of pucD is not clear at present.

The genes are transcribed from a promotor upstream of pucB, in the order pucBA (or pucBAC, pucBACDE where extra genes are present) (Tichy *et al.*, 1989; Lee *et al.*, 1989; Gibson *et al.*, 1992; Hagemann & Tadros, 1994; Tadros *et al.*, 1993; Barrett, 1995).

The only other example of a puf operon currently available is from the green photosynthetic bacterium *Chloroflexus aurantiacus*, which synthesises a RC and LH1 (B806-866) similar to those of purple bacteria (Feick *et al.*, 1982; Wechsler *et al.*, 1987). In this bacterium the pufLM genes are located in one operon (termed puf1) (Shiozawa *et al.*, 1990) whilst the pufBAC genes are located in a separate operon (puf2) (Watanabe *et al.*, 1995). This represents a deviation from the arrangement - described above - of genes found in puf operons of purple photosynthetic bacteria.

The structure of the puc operon

The puc operon encodes genes required for the synthesis of LH2 or LH3, and has been cloned and sequenced in a number of species, namely: *Rb. capsulatus* (Youvan & Ismail, 1985; Tichy *et al.*, 1989); *Rb. sphaeroides* (Ashby *et al.*, 1987; Kiley & Kaplan, 1987; Lee *et al.*, 1989b; Gibson *et al.*, 1992); *Rb. sulfidophilus* (Hagemann & Tadros, 1994); *Rps. palustris* (Tadros & Waterkamp, 1989; Tadros *et al.*, 1993); and *Rps. acidophila* (MacKenzie, 1990; Gardiner *et al.*, 1992; Barrett, 1995). The structure of the puc operon in these species is outlined in Figure 11 (opposite), it is apparent that there is much more variation in the structure of the puc operon than in the puf operon. The minimal structural unit of the puc operon can be considered to consist of the pucBA genes which encode the LH2 (or LH3) α- and β-apoproteins, in species which synthesise only LH1 and LH2, such as *Rb. sphaeroides*, the minimal unit is extended by the addition of the pucC gene downstream of pucA (Lee *et al.*, 1989b; Gibson *et al.*, 1992), and in *Rb. capsulatus* by the genes, pucCDE (Tichy *et al.*, 1989). In species which synthesise LH3 an unusual arrangement of puc genes is found, multiple copies of the puc operon are present. At present 5 pucBA gene pairs have been identified in *Rps. palustris* (Tadros & Waterkamp, 1989; Tadros *et al.*, 1993) and 8 in *Rps. acidophila* (MacKenzie, 1990; Gardiner *et al.*, 1992; Barrett, 1995), some of these pairs are linked to form larger clusters. Two pairs are known to be physically linked in the *Rps. palustris* genome (c and d), but the relationship between the other pairs is unknown (Tadros *et al.*, 1993). In *Rps. acidophila* two clusters have been sequenced, containing 3 (puc1, puc2 and puc3) and 2 (puc4 and puc5) pucBA gene pairs (MacKenzie, 1990; Gardiner *et al.*, 1992; Barrett, 1995). A further 3 gene pairs have been identified within a 20 kb fragment (puc6, puc7 and puc8) most likely forming another cluster, but the exact physical relationship between them is not known (MacKenzie, 1990)[4]. As with *Rps. palustris* the physical distance between the clusters is unknown (Gardiner *et al.*, 1992; Barrett, 1995).

At present there are no reports of a pucC gene associated with one or more of the puc clusters in *Rps. palustris*, however a gene homologous to pucC has recently been identified in *Rps. acidophila*, lying downstream of one of the pucBA gene clusters

[4] The nomenclature that has been use to describe the multiple gene copies differs in the two species. Workers on *Rps. palustris* have used a rather confusing nomenclature where the gene pairs are labelled as $\alpha\beta_a$, $\alpha\beta_b$ and so forth. A more acceptable nomenclature has been used for *Rps. acidophila*, following a scheme which has been used to describe the split puf operon of *C. aurantiacus* and multiple copies of operons in *E. coli*. Thus the pairs are labelled puc1BA, puc2BA, puc3BA and so forth. Since the relationship of between operons of the two species (for instance whether $\alpha\beta_a$ is equivalent to puc4BA) is not clear at present, I will use the described nomenclature of *Rps. acidophila*, and refer to the gene pairs of *Rps. palustris* as, by example, pucaBA.

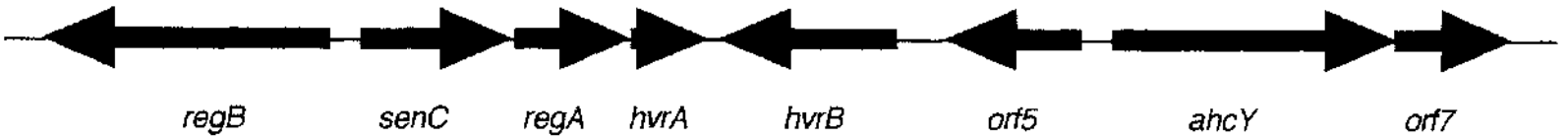
(Dr S.J. Barrett, personal communication). It seems likely that the pucC gene of other species will be identified in the future given the importance that it appears to have in the *Rhodobacter* species.

In both *Rb. capsulatus* and *Rb. sphaeroides* deletion of pucC results in LH2 being absent from photosynthetic membranes (Tichy *et al.*, 1989; Gibson *et al.*, 1992; LeBlanc & Beatty, 1993). Hydropathy analysis reveals that PucC is very hydrophobic, indicating that it is probably an integral membrane protein (Tichy *et al.*, 1989; Gibson *et al.*, 1992), and phoA fusion studies based on hydropathy analysis indicate that it has 12 membrane spanning segments (LeBlanc & Beatty, 1994). It has been suggested that PucC is involved in the assembly of LH2 (Tichy *et al.*, 1991). It is interesting to note that PucC has ~ 24% homology with two other ORFs in the PGC of *Rb. capsulatus* (Bollivar *et al.*, 1994), one of these, orf477 (previously known as F1696), is thought to have a role analogous to PucC (to which it is 47 % homologous), but be involved in the assembly of LH1 (Bauer *et al.*, 1991). Deletion of orf477 in *Rb. capsulatus* results in a 30% decrease in the levels of LH1 (Zsebo & Hearst, 1984; Bauer *et al.*, 1991), since deletion of pucC results in absence of LH2 it has been suggested that pucC is able to compensate in part for the loss of orf477 but not *vice versa* (Tichy *et al.*, 1991). The position of orf477 in the PGC of *Rb. capsulatus* is also notable, it lies upstream of puhA and is probably transcribed with it (Bollivar *et al.*, 1994). The other ORF, known as orf428 is situated within the Bchl synthesis genes between bchG and bchP (Bollivar *et al.*, 1994).

No reports of *Rb. capsulatus* pucDE gene homologs in other species have been published thus far. The pucE gene product is the *Rb. capsulatus* LH2 γ polypeptide, which appears to stabilise the LH2 complex (Tichy *et al.*, 1991), since no other isolated LH2 complexes seem to have a γ polypeptide it is perhaps not surprising that the pucE gene has not been identified elsewhere. The function of pucD is not clear at present, but may have some role associated with that of pucE since the coding frames of these two genes overlap by 4 nucleotides (Tichy *et al.*, 1989), and the genes may be translationally coupled (LeBlanc & Beatty, 1993).

Figure 12

Schematic representation of the photosynthesis regulatory gene cluster of *Rb. capsulatus*. Adapted from Sganga *et al.*, 1992 and Buggy *et al.*, 1994a.



Note: The cluster contains 8 open reading frames spanning a region of the genome approximately 7 kb in size. The genes are represented by arrows which point in the direction of transcription. The exact function of two genes is unclear at present, these are labelled orf5 and orf7. SenC has similarities to sensor kinase proteins but its function is also unclear (Mosley *et al.*, 1994). The regB and regA gene products are a sensor kinase/response regulator pair, probably involved in the regulation of photosynthesis genes in response to changes in oxygen tension (Sganga and Bauer, 1992; Mosley *et al.*, 1994; Inoue *et al.*, 1995). HvrA and HvrB and orf5 appear to be part of a light responsive regulatory system, and both HvrA and HvrB appear to be transcriptional activators (Buggy *et al.*, 1994a; Buggy *et al.*, 1994b). HvrA activates the transcription of puf and puh operons (Buggy *et al.*, 1994a). HvrB is capable of autoregulating its own expression as well as apparently activating the transcription of ahcY and orf5 (Buggy *et al.*, 1994b). The ahcY gene encodes the enzyme S-adenosyl-L-homocysteine hydrolase, which plays a key role in the metabolism of sulphur containing amino acids. It is thought that this enzyme also plays a role in the Bchl synthesis pathway (Sganga *et al.*, 1992). A full explanation of these gene products and their regulatory actions is given in the section "Photosynthesis gene regulation".

The regulatory gene cluster

In *Rb. capsulatus* a number of genes which appear to have a role in regulation of photosynthesis genes have also been found to be clustered together (Sganga & Bauer, 1992; Sganga *et al.*, 1992; Buggy *et al.*, 1994a; Buggy *et al.*, 1994b; Mosely *et al.*, 1994). Figure 12 (opposite) illustrates the arrangement of genes in the photosynthesis regulatory gene cluster of *Rb. capsulatus*. The cluster includes 8 genes spanning a region of the genome approximately 7 kb in size. The exact function of two genes is unclear at present, these are labelled orf5 and orf7.

The regB and regA gene products are a sensor kinase/response regulator pair, probably involved in the regulation of photosynthesis genes in response to changes in oxygen tension (Sganga and Bauer, 1992; Mosley *et al.*, 1994; Inoue *et al.*, 1995). The regA gene has also been cloned and sequenced from *Rb. sphaeroides*[5] (Eraso & Kaplan, 1994; Phillips-Jones & Hunter, 1994), as has the regB gene (Eraso & Kaplan, 1995). The predicted amino acid sequence of *Rb. sphaeroides* RegA exhibits 81% identity to that of *Rb. capsulatus* RegA (Phillips-Jones & Hunter, 1994), whilst the RegB amino acid sequences are 58% identical (Eraso & Kaplan, 1995), suggesting that the system of gene regulation in response to oxygen tension may be conserved between the two species. SenC has similarities to sensor kinase proteins but its function is also unclear (Mosley *et al.*, 1994). SenC has also been cloned and studied in *Rb. sphaeroides* (Eraso & Kaplan, 1995). The results of this study indicate that SenC is anchored in the membrane with its c-terminus in the periplasmic space and that like RegB it is involved in the positive regulation of photosynthesis gene expression in response to decreasing oxygen tension, although with a much more subtle effect (Eraso & Kaplan, 1995).

HvrA and HvrB and orf5 appear to be part of a light responsive regulatory system, and both HvrA and HvrB appear to be transcriptional activators (Buggy *et al.*, 1994a; Buggy *et al.*, 1994b). HvrA activates the transcription of puf and puh operons (Buggy *et al.*, 1994a). HvrB is capable of autoregulating its own expression as well as apparently activating the transcription of ahcY and orf5 (Buggy *et al.*, 1994b). The ahcY gene encodes the enzyme S-adenosyl-L-homocysteine hydrolase, which plays a key role in the metabolism of sulphur containing amino acids. It is thought that this enzyme also plays a role in the Bchl synthesis pathway (Sganga *et al.*, 1992).

There are few genes that have been found to play a part in the the regulation of photosynthesis gene expression but to be unlinked to both the PGC and the regulatory gene cluster. These include the previously mentioned hemT, hemA and hemF genes of *Rb. sphaeroides* (Coomber *et al.*, 1992; Choudhary *et al.*, 1994), which code for enzymes involved in heme biosynthesis and thus may influence the synthesis of Bchl , and a gene known as orf798 (Pollich *et al.*, 1993). Transposon insertion into orf798 (Pollich *et al.*, 1993) results in an inability of *Rb. capsulatus* to 'de-repress' the puf, puh and puc operons when the oxygen tension drops to a level suitable for photosynthesis.

[5] It should be noted that Phillips-Jones & Hunter, (1994) have applied the regA terminology first used in *Rb. capsulatus* to *Rb. sphaeroides*, whilst Eraso & Kaplan, (1994 & 1995) have used a different terminology. However it is clear that these genes are homologous in both sequence and function, thus I will use the *Rb. capsulatus* terminology in this discussion. The *Rb. sphaeroides* genes prrA, prrB and prrC will therefore be considered equivalent to regA, regB and senC in *Rb. capsulatus*.

Figure 13

Photosynthesis gene transcript levels measured in response to different oxygen tensions and irradiance levels in the species *Rb. capsulatus* and *Rb. sphaeroides*. Table A adapted from Zhu & Hearst, 1986. Table B from Zhu & Kaplan, 1985.

Table A: *Rb. capsulatus* transcript levels

Gene	mRNA level, %*				
	High Light	Dark (1hr)	Low Light	High Oxygen	Low Oxygen
pucBA	100	35	182	3	113
pufBA	100	51	173	28	110
pufL	100	25	130	27	100
pufM	100	28	139	32	96
puhA	100	28	176	30	106
bchJ, G, D	100	25	124	76	36
bchH, K, F	100	36	119	140	30
crtE, F	100	62	80	30	60
bchI, crtA, I, B	100	31	80	105	29
crtC, D	100	55	82	59	34
crtE	100	69	76	72	69

* The relative amounts of mRNA were measured by dot hybridization, quantified by densitometry of x-ray film, and expressed as the percentages of mRNA level under high light (set as 100).

Table B: *Rb. sphaeroides* transcript levels

Gene	mRNA level, %*			
	High Light	Dark (1hr)	Low Light	Semi Aerobic
pufBA	100	30	183	25
pufBA + pufL	100	28	152	28
pufM	100	25	147	25

* The relative amounts of mRNA were measured by dot hybridization, quantified by scintillation counting of the filters, and expressed as the percentages of mRNA level under high light (set as 100).

Note: The two tables above illustrate the relative transcript levels of a number of photosynthesis genes when the bacteria *Rb. capsulatus* and *Rb. sphaeroides* are grown under various conditions. It should be noted that since these values are relative rather than absolute it is only possible to compare data between columns in a single row, and not between rows or between the two tables. The high light treatment of Table A was 30 Wm^{-2} and the low light treatment 6 Wm^{-2} (Zhu & Hearst, 1986). This compares with a high light irradiance of 10 Wm^{-2} and low light of 3 Wm^{-2} for Table B (Zhu & Kaplan, 1985). In both studies the dark treatment was carried out by placing high light grown cultures in the dark for 1 hr. The semi-aerobic condition of the *Rb. sphaeroides* study is equivalent to the low oxygen condition of the *Rb. capsulatus* study.

However it is not yet clear whether orf798 has a direct effect on oxygen repression, or an indirect effect via the porphyrin biosynthesis pathway (Pollich *et al.*, 1993).

A more detailed description of the state of knowledge as regards these gene products and their regulatory actions is given in the following sections.

Photosynthesis Gene Expression

Early studies on photosynthetic bacteria indicated that synthesis of pigment-protein complexes was inhibited by oxygen, and suppressed by high irradiance under anaerobic conditions (Cohen-Bazire & Kunisawa, 1960; Aagard & Sistrom, 1972; Oelze & Drews, 1972; Takemoto & Huang Kao, 1977; Drews & Oelze, 1981; Ohad & Drews, 1982; Chory & Kaplan,1983). Following the cloning of many photosynthesis genes from *Rb capsulatus* and *Rb sphaeroides* studies on their expression showed that the pattern of gene expression closely matches the pattern of pigment and protein synthesis identified in early studies on photosynthetic bacteria (Clark *et al.*, 1984; Zhu & Kaplan, 1985; Zhu & Hearst, 1986). Both light and oxygen were shown to repress transcription, in the same manner that they were shown to repress the levels of protein and pigments in the ICM.

Oxygen response

Oxygen appears to have a much stronger repressive effect on photosynthesis gene expression than light, however the scale of repression varies depending on the gene product encoded. Initial studies in *Rb. capsulatus* indicated that crt gene transcript levels vary little in response to oxygen, whilst bch genes are more firmly repressed and respond in a manner similar to the puf and puh operons (Clark *et al.*, 1984; Klug *et al.*, 1985; Zhu & Hearst, 1986). The puc operon appeared most sensitive of all to oxygen tension (Clark *et al.*, 1984; Zhu & Hearst, 1986). These responses to oxygen are described in Figure 13 (opposite), where the results of two studies are displayed in table form. It is clear that certain transcripts vary more widely than others, but that all except those hybridising with the *Rb. capsulatus* bchH,K,F probe are repressed by oxygen. It now seems clear however that the latter result was due to the fact that the probe carried part of the ppsR regulatory gene which is expressed under high oxygen tension (Ponnampalam *et al.*, 1995).

However detailed studies on the crt, bch, puf, puh and puc gene expression of *Rb. capsulatus* during the transition to anaerobiosis, have provided a slightly more complicated picture (Cook *et al.*, 1989; Armstrong *et al.*, 1993). These studies indicated that whilst the level of crtI and B mRNA remained constant during the transition, levels of crtA, C, D, E, F, K, bchC and bchD increased 2- to 12-fold (Armstrong *et al.*, 1993). The levels of puf and puh mRNA increased in a similar manner, 6 to 8 fold (Cook *et al.*, 1989), whilst puc mRNA increased 25-fold, but the puc increase lagged ~20 minutes behind other genes (Zhu & Hearst, 1986). This lag has been observed in several other studies of *Rb. capsulatus* (Klug *et al.*, 1985) and *Rb. sphaeroides* (Hunter *et al.*, 1987; Kiley & Kaplan, 1987). The response of *Rb. sphaeroides* is clearly different in a least one respect, since the levels of crtI mRNA decrease under aerobic conditions, rather than remaining constant as they appear to do

in *Rb. capsulatus* (Lang *et al.*, 1994).

Light response

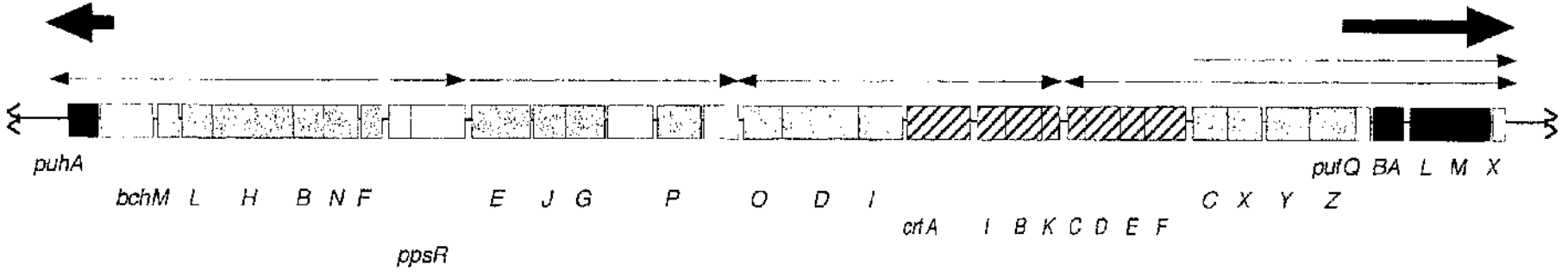
Whilst puf, puh, puc and bch transcript levels are reduced under high light intensity, crt expression reacts in the opposite way, increasing with irradiance levels (Zhu & Hearst, 1986). This is to be expected given the role that Car plays in photoprotection (Cogdell, 1985). Figure 13 illustrates the effect of light-intensity on photosynthesis gene expression. It can be seen that the puc operon is most responsive to changes in irradiance in both *Rb. sphaeroides* and *Rb. capsulatus*, this is also to be expected since the puc operon is regarded as the variable LHC, increasing in quantity in response to decreasing irradiance.

Rps. palustris and *Rps. acidophila* show a more complex pattern of gene expression in response to changing irradiance as befits their more complex photosynthetic units. Three of the pucBA gene clusters of *Rps. palustris* appear to be expressed only under low-light growth, whilst at high light all five are expressed (Tadros *et al.*, 1993). Surprisingly northern blot hybridisation indicated that the abundance of all these transcripts increased as the light intensity is raised, this is the opposite to the response generally observed in other species (Tadros *et al.*, 1993). Whilst one would expect expression of genes encoding LH3 to change in this way, LH2 gene expression would be expected to react in the opposite manner, however it is not known at present which of the gene clusters encodes which LHC polypeptides (Tadros *et al.*, 1993). There is also a study on puc operon expression in *Rb. capsulatus* in which a similar response is observed (Zucconi & Beatty, 1988). Although northern blot hybridisation using RNA from high and low light grown cultures showed greater abundance of puc mRNA at low light, an additional approach using S1 nuclease protection appeared to show the opposite, that there was more puc mRNA at high light than low (Zucconi & Beatty, 1988). This discrepancy suggests that interpretation of the early blot hybridisation studies (mentioned above) may be somewhat more complex than was initially assumed, however further studies using techniques which allow measurement of promotor strength under various conditions should clarify this area.

Of further interest in regard to light regulation of photosynthesis genes is the discovery that the expression of puc and puf operons in *Rb. sphaeroides* is repressed to the greatest extent by blue light at a wavelength of 450 nm, and furthermore that greater repression of the puc operon occurs than of the puf operon (Iba & Takamiya, 1989; Shimada *et al.*, 1992). This suggests that light induced repression of photosynthesis gene expression may be mediated in part by a blue pigment (Shimada *et al.*, 1992).

Figure 14

A schematic view of proposed transcription units in the *Rb. capsulatus* PGC. Diagram adapted from Bauer *et al.*, (1991) and Wellington *et al.*, (1992). This diagram illustrates the overlapping operonal and superoperonal transcripts which appear to be synthesised in *Rb. capsulatus*.



Key

- Genes coding for RC-LH1 structural polypeptides
- Genes involved in bacteriochlorophyll biosynthesis
- Genes involved in carotenoid biosynthesis
- Genes of indeterminate or possibly regulatory function

Note: Arrows represent proposed unprocessed transcripts, and point in the direction of transcription. The thick arrows indicate the highly expressed (and oxygen regulated) transcripts of the puf and puh operons. The smaller arrows represent transcripts expressed at a lower level, often under aerobic as well as anaerobic conditions. This arrangement of genes into 'superoperons' is thought to facilitate a rapid transition from aerobic respiratory to anaerobic photosynthetic growth, by ensuring coordinated expression of photosynthesis genes at low levels under aerobic conditions (Wellington *et al.*, 1992).

Regulation of photosynthesis gene expression

In the previous sections I have summarised the physical arrangement of genes thought to be involved in the construction of the photosynthetic apparatus in purple photosynthetic bacteria and outlined the changing gene expression in response to environmental changes. Regulation of the expression of these genes can occur at 4 conceptual levels: at the level of transcription; post transcription; at translation; and post-translationally. Evidence of regulatory mechanisms has already been detected at the transcriptional, post-transcriptional and post-translational levels, and the current state of knowledge regarding these mechanisms will be presented in the sections following. As in previous sections extensive reference will be made to the species *Rb. capsulatus* and *Rb. sphaeroides*, since much more is known regarding the molecular biology of photosynthesis in these species than in less studied species. Where information from other species is known comparisons will of course be made.

Transcriptional control mechanisms

The control of transcription from photosynthesis genes has been shown to be relatively complex in *Rb. capsulatus* and *Rb. sphaeroides*, this is perhaps not surprising given the large number of genes (42 in the *Rb. capsulatus* PGC and puc operon) which have to be expressed coordinately in response to changing irradiance levels and oxygen tension. To achieve this control photosynthetic bacteria appear to utilise transcriptional control elements similar to those used by other Proteobacteria, such as *E. coli* as well as novel elements. The promotors, activators, repressors, sensors and response regulators of bacterial gene expression all appear to be applied to photosynthesis genes. However a further method of coordinating gene expression appears to be used by *Rb. capsulatus* and *Rb. sphaeroides*, both this 'superoperonal' control mechanism and the more traditional control elements thus far defined for photosynthesis genes are described in the following sections.

Superoperonal transcription

The genes constituting the photosynthesis gene cluster of *Rb. capsulatus* are arranged into several operons[6], each operon appears to have its own promotor from which polycistronic transcripts may be transcribed (Armstrong *et al.*, 1989; Young *et al.*, 1992; Alberti, 1991). However it has also been suggested that 'superoperons' exist in the PGC of *Rb. capsulatus* (Young *et al.*, 1989), these result from transcriptional read-through across two or more operons. Figure 14 (opposite) illustrates the initial transcripts proposed to be produced from the PGC of *Rb. capsulatus* (Bauer *et al.*, 1991; Wellington *et al.*, 1992). These include transcripts covering the operons puh, bchFNBHLMorf427 , ppa-pps, bchEJGorf428Porf176, crtAbchIDO, crtIBK, crtDC, crtEF, bchCXYZ and pufQBALMX (Bauer *et al.*, 1988; Giuliano *et al.*, 1988; Young *et al.*, 1989; Adams *et al.*, 1989; Wellington and Beatty, 1989; Burke *et al.*, 1993b). And superoperonal transcripts covering crtEF-bchCXYZ-pufQBALMX (Young *et al.*, 1989; Wellington and Beatty, 1991), bchCXYZ-

[6] An 'operon' is defined as two or more adjacent genes, transcribed into a single mRNA, and the adjacent transcriptional control sites required for expression.

pufQBALMX (Young *et al.*, 1989; Wellington and Beatty, 1991), and bchFNBHLMorf427-puhA (Bauer *et al.*, 1991). The multi-operonal messages produced from such read-through allow the transcriptional linkage of pigment biosynthesis and structural polypeptide genes. This superoperonal linkage has been shown to be essential for wild-type transcriptional levels of the bchCXYZ and puf operons (readthrough from the crtEF and bchCXYZ promotors) (Young *et al.*, 1989; Wellington and Beatty, 1991; Wellington *et al.*, 1991), and the puh operon (readthrough from bchFNBHLMorf427 promotor) (Yang and Bauer, 1990; Bauer *et al.*, 1991). This is thought to confer two advantages to *Rb. capsulatus* : firstly it ensures that pigment biosynthesis is linked to structural polypeptide production, thus preventing a potentially dangerous accumulation of unbound pigment (from which free radicals could be produced in the presence of light and oxygen); secondly it facilitates the rapid transfer from aerobic respiratory to anaerobic photosynthetic growth, since there will be a low level of photosynthetic apparatus and its precursors synthesised under aerobic conditions, which can be utilised as soon as anaerobic conditions are established (Young *et al.*, 1989; Yang and Bauer, 1990; Wellington and Beatty, 1991; Wellington *et al.*, 1991; Bauer *et al.*, 1991; reviewed by Wellington *et al.*, 1992).

The apparent conservation of the PGC in purple photosynthetic bacteria suggests that superoperonal control mechanisms may also be widespread. However the large superoperonal transcripts detected in *Rb. capsulatus* have rarely been detected in other species. One study involving analysis of the crtEF-bchCXYZ-puf region of the *Rb. sphaeroides* PGC, has revealed the presence of a potential promotor between crtE and crtF, and that the promotors of these and the bchCXYZ operons are likely to be oxygen regulated unlike those in *Rb. capsulatus* (Lang *et al.*, 1995). However a related study on the bchCXYZ region revealed the presence a large (>9.5 kb) transcript, as well as a 2 smaller transcripts, suggesting that a superoperonal transcript possibly covering bchC to pufX is produced by *Rb. sphaeroides* (McGlynn & Hunter, 1993). This suggests that superoperonal transcription is a conserved feature of the PGC, but that the distribution of transcription initiation points and the regions spanned by transcripts follows a species specific pattern.

The other photosynthesis genes identified in the previous sections fall into more typical operon structures, the puc operon from various species of photosynthetic bacteria appears to be expressed in the form of a single polycistronic transcript (see Figure 11), and the *Rb. capsulatus* regulatory gene cluster has several small operons (Sganga *et al.*, 1992 and Buggy *et al.*, 1994a).

Cis acting regulatory elements

A number of elements have been identified which lie up and downstream of photosynthesis genes and play a role in the regulation of transcription from them. Whilst many of these elements have been identified by sequence homology and thus remain potential sites of action, several have been identified and analysed by the use of assays for DNA binding proteins and site directed mutagenesis. These elements include promotors, many of which show homology to the consensus σ^{70} and σ^{60} promotor recognition sequences of *E. coli* (Young *et al.*, 1989; Armstrong *et al.*, 1989).

Sequences which are thought to represent the recognition sites of repressor and activator proteins have also been identified, again these often show homology to known protein binding sites. A number of sites which bear no resemblance to known protein binding elements have also been revealed, though it is not yet clear what role these elements play in the regulation of photosynthesis genes.

Transcription termination occurs at sites in cis to photosynthesis genes and has been shown to be an important factor in the regulation of several photosynthesis operons, however since the cis elements involved often have secondary roles in post-transcriptional control mechanisms, I will first discuss those cis sites and trans acting factors which influence transcription initiation before moving onto transcription termination.

Promotors

Initial attempts to identify promotors of photosynthesis genes involved mapping of the 5′ end of a detected mRNA species, which should of course represent the site of transcription initiation. However this was often complicated by the superoperonal organisation of photosynthesis genes, which potentially allows transcription of genes from several upstream promotors and by the post-transcriptional processing of transcripts (described in the section "Post-transcriptional control" below). Difficulties with sequence analysis were encountered because several promotors (notably those of puf and puh operons) lie within the ORFs of upstream genes, masking any homology between promotors. Thus site directed mutagenesis and the use of transcriptional fusions to reporter genes have proved essential for delineating promotor regions.

The promotors identified upstream of photosynthesis genes appear to fall into 3 generalised groups, the first of these includes the promotors of most crt and bch encoding operons which show some homology to the *E. coli* σ^{70} promotor consensus. The *E. coli* σ^{70} consensus , TTGACA$(N)_{15\text{-}19}$TATAAT (McClure, 1985), is usually best matched in the more distal -35 (to transcription initiation) region, with more variation occurring in the -10 region. It should be noted that even *E. coli* σ^{70} promotors vary from this consensus, such variation actually seems to determine the strength of the promotor (McClure, 1985; Reznikoff *et al.*, 1985), it is also important to note that variation away from the consensus in one region can be compensated by increasing consensus to the other regions of the promotor (including the spacer region). Thus far sequences matching the σ^{70} consensus have been identified upstream of the *Rb. capsulatus* genes bchF (Burke *et al.*, 1993b), bchC (Ma *et al.*, 1993), crtI and crtD (Armstrong *et al.*, 1989). Similarly the *Rb. sphaeroides* genes bchC (Penfold & Pemberton, 1994); bchF and bchE (Gomelsky & Kaplan, 1995); crtA, crtD and crtE (Lang *et al.*, 1995) have potential σ^{70} like promotors. It should be noted however that most of these promotors have been identified by sequence homology, only the element upstream of *Rb. capsulatus* bchC has been studied by site directed mutagenesis (Ma *et al.*, 1993).

The second group of promotors includes the puf and puh promotors, which appear to have a somewhat unique structure. These have been sequenced in a wider range of species and studied intensively in the two *Rhodobacter* species. In both puf

and puh operons the promotors lie within upstream genes, the puf promotor within the bchZ (Bauer *et al.*, 1988; Young *et al.*, 1989; Hunter *et al.*, 1991)[7] gene and the puh promotor within orf477 (F1696) (Berard *et al.*, 1989; Bauer *et al.*, 1991). It has been suggested that puf and puh promotors resemble the consensus of σ^{54} (NtrA) promotors involved in transcription of nitrogen fixation genes in *Rb. capsulatus* (Bauer *et al.*, 1991). This is of interest because nitrogen fixation, like photosynthesis, is only carried out under anoxygenic conditions. However NtrA is not itself involved with transcription of photosynthesis genes since ntrA$^-$ strains are still capable of photosynthesis (Jones & Haselkorn, 1989). Such homology has led to the suggestion that the puf and puh promotors utilise an alternative σ factor specific to photosynthesis (designated σ^P), as yet no other evidence for such a σ factor has been found (Bauer *et al.*, 1991). As well as homology between the puf promotors of *Rb. capsulatus*, *Rb. sphaeroides* and *Rps. viridis*, homology also exists between these promotors and the puh promotors of *Rs. rubrum* and *Rb capsulatus* (Bauer *et al.*, 1993). This suggests that similar elements are used to control expression of puf and puh operons and that these elements are conserved to a limited extent between species of photosynthetic bacteria.

The last group of promotors contains only that of the puc operon, in the *Rhodobacter* species this shows some homology to the σ^{70} like crt promotors, particularly in the -35 region, but clearly differs more around the -10 region. The *Rb. capsulatus* puc operon is differentiated from that of *Rb. sphaeroides* however because two transcription start sites have been mapped which occur within 15 bp of each other (Zucconi & Beatty). The *Rps. palustris* puc gene clusters have been shown (by the use of lacZ gene fusions to the puc upstream regions) to have different promotor strengths, and there is clear variation between their sequences (Tadros *et al.*, 1993). Only the puc2BA and puc5BA clusters have potential σ^{70} recognition sequences, it is not clear at present what promotors are utilised upstream of the other *Rps. palustris* puc gene clusters (Tadros *et al.*, 1993). The puc operon promotors thus far cloned differ further from σ^{70} type promotors of crt and bch genes in their use of certain other cis and trans elements, these are described in the following sections.

Repressors and activators of transcription

A number of interesting sites adjacent to mapped and suggested promotor elements have been found. Probably the most widespread and highly conserved cis acting element thus far identified in photosynthetic bacteria is the palindrome $TGT(N)_{12}ACA$. This sequence has been identified upstream of the *Rb. sphaeroides* bchF, bchE (Gomelsky & Kaplan, 1995); crtD (Garí *et al.*, 1992); crtI (Lang *et al.*, 1994) and occurs in tandem upstream of the puc operon (Lee & Kaplan, 1992). Similarly *Rb. capsulatus* has this element upstream of the genes bchF (Burke *et al.*, 1993b); bchC (Ma *et al.*, 1993); pucB (Tichy *et al.*, 1989); crtD and crtE (Armstrong *et al.*, 1989); and in tandem upstream of crtA and crtI (Armstrong *et al.*, 1989). In

[7] In contrast to the deletion analysis carried out by Hunter *et al.*, (1991), which indicated that the *Rb. sphaeroides* O_2 regulated puf promotor lay upstream of pufQ, transcript end-mapping experiments (Zhu *et al.*, 1986; DeHoff *et al.*, 1988; Lee *et al.*, 1989a) suggested two puf transcription start sites upstream of pufB (downstream of pufQ), thus some confusion remains over the location of *Rb. sphaeroides* puf operon promotors although it seems likely that the oxygen regulated promotor lies upstream of pufQ.

addition four of the five sequenced puc gene clusters from *Rps. palustris* also have this palindrome lying upstream, it occurs once 5′ to puc1BA, puc2BA and puc4BA, and twice upstream of puc5BA (Tadros *et al.*, 1993).

These sites are known to bind the protein PpsR, which is responsible for oxygen stimulated repression of many photosynthesis genes in *Rb. capsulatus* (Ponnampalam *et al.*, 1995) and *Rb. sphaeroides* (Gomelsky & Kaplan, 1995). Most of the sites overlap the putative σ^{70} promotor elements described in the previous section, this is a typical position of *E. coli* repressor binding at σ^{70} promotors and probably reflects a need for physical blocking of the promotor sequence (Collado-Vides *et al.*, 1991). The palindrome is also homologous to a consensus binding site – TGTGT$(N)_{6\text{-}10}$ACACA – derived from a range of prokaryotic DNA binding regulatory proteins which act as repressors or activators (Gicquel-Sanzey & Cossart, 1982; Buck *et al.*, 1986). The PpsR protein and its action are discussed further in the section "Trans acting regulatory elements" below.

Cis elements less typical of σ^{70} promotors are found upstream of the *Rb. sphaeroides* puc promotor, in particular a sequence 120 bp upstream of the transcription start – AAAATTGTCCCTTTT – has been shown to bind a protein (designated PORP – puc oxygen responsive protein) that appears to bind more strongly when cells are cultured under aerobic conditions (McGlynn & Hunter, 1992). Thus this site may represent the point of action of an oxygen responsive repressor protein (McGlynn & Hunter, 1992). Of further interest regarding this site is that it bears a resemblance to the consensus FNR regulator binding site (Colonna-Romano *et al.*, 1990; Lee & Kaplan, 1992) and is overlapped by a putative IHF binding site that is protected by *E. coli* IHF in DNase I footprinting analysis (Lee & Kaplan, 1992). This site also marks the boundary of the puc operon DRS (downstream regulatory region, -150 to -1 of the transcription start and involved in regulation of expression by oxygen) and URS (upstream regulatory region, -629 to 150, involved in regulation by light and oxygen), identified by Lee & Kaplan, (1992). Since IHF belongs to a family of proteins involved in DNA binding and bending (Oberto *et al.*, 1994), it was suggested that this site functions in the bending of the DNA and thus facilitates interaction of the URS with the DRS, to allow activated transcription of the *Rb. sphaeroides* puc operon as well as serving as a repressor binding site in the presence of oxygen (Lee & Kaplan, 1992; Lee *et al.*, 1993).

Between the putative IHF binding site and the promotor/PpsR binding region or the *Rb. sphaeroides* puc operon lies a sequence – TGGC$(N)_{10}$TCGCA – which is similar to a sequence – TGGC$(N)_8$CCGCA – found in a comparable position in the puc operon of *Rb. capsulatus*, interestingly a similar sequence lies upstream of the pufQ gene in the *Rb. sphaeroides* – CGGC$(N)_{12}$TCGCT – and *Rb. capsulatus* – CGGC$(N)_8$TCGCT – puf operons (Lee & Kaplan, 1992). Although this sequence has no known function at present its conservation and positioning (close to and upstream of the promotor) suggest that it may have a role to play in transcriptional regulation, possibly as an activator binding site (Collado-Vides *et al.*, 1991; Lee & Kaplan, 1992).

A further cis acting site has been identified in the puf operons of *Rb. capsulatus* and *Rb. sphaeroides*, this element is a represented by a region of partial dyad symmetry

and approximately 10 bp 3′ to it located immediately 5′ to the puf promotor region (Klug, 1991b). Initial gel retardation studies in *Rb. capsulatus* indicated that the element represented a protein binding site, and that different DNA-protein complexes formed with extracts of cells cultured under different conditions (Klug, 1991b). Specifically, extracts from cells grown under aerobic conditions were seen to form 2 different DNA-protein complexes (I and II), when semi-aerobic conditions were used only complex I was formed (Klug, 1991b). Thus it was suggested that the site functioned in oxygen regulated repression of the puf operon, and also the puc operon, since this competed with the puf operon to bind the putative repressor (Klug, 1991b). In *Rb. sphaeroides* the homologous element has also been shown to be a protein binding site but to be involved in light regulated repression of the operon as well as oxygen stimulated repression (Shimada *et al.*, 1993).

Trans acting regulatory elements

Trans acting elements can be conceptually divided into two groups, that is those which interact with cis regulatory sites, often binding DNA, and those which act further up the regulatory pathway, sensing and responding to environmental conditions such as light and oxygen. Those binding to the cis elements described in the previous section will be described first, before a description of the remaining trans acting elements is given.

The regions of partial dyad symmetry in the puf operons of *Rb. capsulatus* and *Rb. sphaeroides* described above bind a protein or proteins in a manner which suggests that the protein acts to repress transcription when oxygen is present (Klug, 1991b). The observation of two DNA-protein complexes, only one of which appeared to form under semi-aerobic conditions, has led to the suggestion the protein may dimerise (or that a second protein may also bind) under fully aerobic conditions (Klug, 1991b). This would thus allow a two step repression process, with complete repression at high oxygen tension, and partial repression at lower oxygen tension (at which some puf transcript is usually formed) (Klug, 1991b). The protein was also shown to bind more strongly when dephosphorylated (Klug, 1991b). The comparable element in *Rb. sphaeroides* has also been shown to bind a repressor-like protein which binds more strongly when dephosphorylated (Shimada *et al.*, 1993). Thus far however it has only been seen to form a single DNA-protein complex, covering only one arm of the dyad symmetry element (like the *Rb. capsulatus* complex I), although this difference may the result of variation in the assay conditions used (Shimada *et al.*, 1993). In addition the protein was shown to bind more strongly using extracts from cells cultured semi-aerobically under blue light, than when red light extracts were used (Shimada *et al.*, 1993). This suggests that under semi-aerobic conditions the protein also mediates light-regulated repression of the puf operon in *Rb. sphaeroides* (the effect of light was not measured in *Rb. capsulatus*) (Shimada *et al.*, 1993). Thus although this protein has been identified in *Rhodobacter* its exact role, and the method by which it is phosphorylated/dephosphorylated *in vivo* have yet to be determined.

The trans acting factor PORP which binds upstream of the puc promotor of *Rb. sphaeroides* has received little more attention as yet, as mentioned previously PORP

appears to be a monomeric oxygen responsive repressor protein, and has an apparent molecular weight of 22500 (McGlynn & Hunter, 1992).

IHF can be considered as a putative trans acting factor binding in cis to the puc operon, although it has not been shown that native IHF binds to the puc operon site protected by *E. coli* IHF (Lee & Kaplan, 1992; Lee *et al.*, 1993), both the himA and hip genes (encoding the α and β subunits of IHF) have been sequenced in *Rb. capsulatus* (Toussaint *et al.*, 1991; Toussaint *et al.*, 1993), and himA in *Rb. sphaeroides* (Lee *et al.*, 1993). This suggests that a native IHF is present and may bind to these sites *in vivo*. The binding of IHF from *Rb. capsulatus* to DNA has also been modelled (Toussaint *et al.*, 1994).

The best studied trans acting factor at present is the protein PpsR. This is encoded by the ppsR gene, located between the bchF and bchE genes in the PGC of both *Rb. capsulatus* and *Rb. sphaeroides* (see the Figure 9) (Penfold & Pemberton, 1991). PpsR is a repressor protein which binds to the palindrome $TGT(N)_{12}ACA$ (Gomelsky & Kaplan, 1995) by means of a helix-turn-helix (HTH) motif in its C-terminal domain which shows homology to many other DNA binding proteins, including several response regulator proteins (Penfold & Pemberton, 1994). However the N-terminal domain has little homology to these response regulator proteins (Penfold & Pemberton, 1994). Interruption of ppsR induces over-production of Bchl and Car, whilst over-expression of ppsR leads to a complete suppression of photopigment biosynthesis (Penfold & Pemberton, 1991; Penfold & Pemberton, 1994). PpsR appears to represses crt, bch and puc gene expression in response to the presence of oxygen (Zhu & Hearst, 1986; Ma *et al.*, 1993; Ponnampalam *et al.*, 1995). It is not known at present how oxygen tension is sensed and the information passed onto PpsR.

Genes coding for several trans acting factors are located in the previously described regulatory gene cluster of *Rb. capsulatus* and several are also linked in *Rb. sphaeroides*. These include the RegB and RegA proteins which operate as a sensor kinase/response regulator pair, and required for full anaerobic induction of puf, puh and puc operons (Sganga and Bauer, 1992; Mosley *et al.*, 1994; Inoue *et al.*, 1995; Phillips-Jones & Hunter, 1994; Eraso & Kaplan, 1994; Eraso & Kaplan, 1995). RegB is a sensor histidine kinase with a transmembrane domain and a cytoplasmically located domain, the transmembrane domain serves to anchor the protein in the membrane and may also allow it to sense the redox state of components of the electron transport chain (Mosley *et al.*, 1994). The cytoplasmic domain is able to autophosphorylate at a conserved histidine residue and transfer the phosphate to a conserved aspartate residue on RegA (Inoue *et al.*, 1995), these are typical features of sensor kinase/response regulator two component regulation systems (Ronson *et al.*, 1987; Parkinson, 1993). RegA shares homology with several response regulator proteins, but has no apparent DNA binding domain (such as an HTH motif), this indicates that RegA functions as an intermediate in a more complex regulatory cascade, most likely passing on the phosphate to further regulatory factors (Sganga and Bauer, 1992; Phillips-Jones & Hunter, 1994; Eraso & Kaplan, 1994). SenC which is encoded by the senC gene lying between regA and regB may also interact with RegA (Mosley *et al.*, 1994; Eraso & Kaplan, 1995). SenC like RegB has similarities to sensor kinase proteins and is

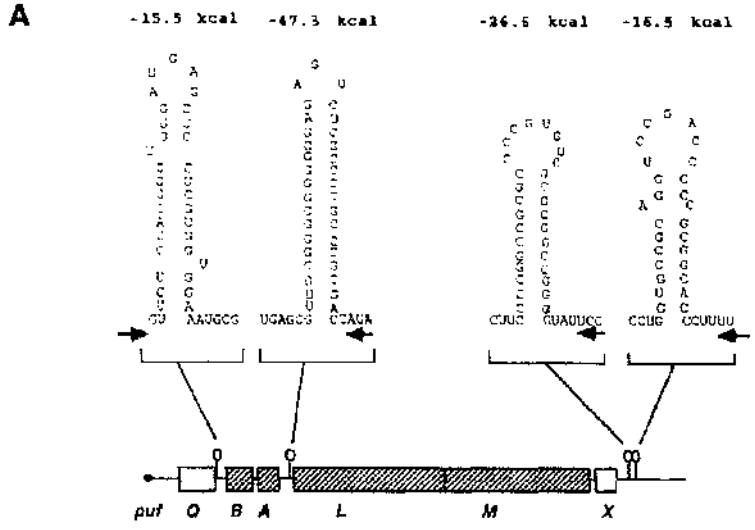
anchored in the membrane, but in contrast its C-terminus in the periplasmic space, like RegB it is involved in the positive regulation of photosynthesis gene expression in response to decreasing oxygen tension, although with a much more subtle effect (Mosley *et al.*, 1994; Eraso & Kaplan, 1995). It should be noted that RegB may also have phosphatase activity and thus negatively affect photosynthesis gene expression under aerobic conditions (Eraso & Kaplan, 1995). Thus the RegB/RegA/SenC system appears to play a role in global activation of photosynthesis genes in response to a drop in low oxygen tension.

Two light responsive trans acting factors have been identified in *Rb. capsulatus*, encoded by the genes hvrA and hvrB (see the section "The regulatory gene cluster"). HvrA and HvrB and orf5 appear to be part of a light responsive regulatory system, and both HvrA and HvrB appear to be transcriptional activators (Buggy *et al.*, 1994a; Buggy *et al.*, 1994b). HvrA has a potential HTH motif and may to bind to the puf and puh promotor regions, activating transcription under low light conditions (Buggy *et al.*, 1994a). HvrA is in fact required for the WT response to low light conditions, HvrA$^-$ mutants have slower growth under these conditions (Buggy *et al.*, 1994a). HvrB shows homology to the LysR family of transcriptional regulators, and like these regulators contains an HTH motif in its N-terminal domain (Henikoff *et al.*, 1988; Buggy *et al.*, 1994b). HvrB also shares several other features with members of this group: it has a similar mass (~32 kDa); it is capable of autoregulating its own expression; and it discoordinately activates the transcription of the nearby ahcY and orf5 genes (Buggy *et al.*, 1994b). HvrB appears to activate the expression of orf5 only under high light and the expression of ahcY only under low light (Buggy *et al.*, 1994b). HvrB may differ from other members of the LysR family however in that these usually have metabolic compounds as coregulators which are involved with the activities of the target genes, thus far HvrB seems only to respond to light intensity (Buggy *et al.*, 1994b). It is not known at present whether HvrB directly senses light intensity or whether further trans acting factors are involved. Although the role of the orf5 gene is unknown, ahcY gene is known to encode the enzyme S-adenosyl-L-homocysteine hydrolase, which plays a key role in the metabolism of sulphur containing amino acids (Sganga *et al.*, 1992). It is thought that this enzyme also plays a role in the early stages of the Bchl synthesis pathway, and that *Rb. capsulatus* may thus partly regulate Bchl synthesis by controlling the level of Bchl precursors (Sganga *et al.*, 1992).

DNA gyrase, a more generalised trans acting factor of gene regulation may also affect photosynthesis gene expression, a study using gyrase inhibitors has indicated that transcription of puf, puc and bch genes is inhibited by gyrase inhibitors (Zhu & Hearst, 1988). This suggests that DNA supercoiling is necessary for effective transcription of photosynthesis genes (Zhu & Hearst, 1988).

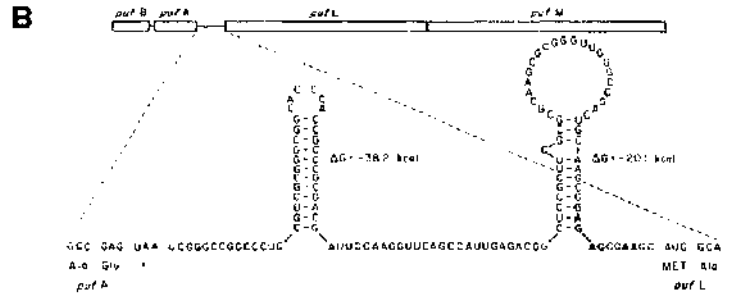
Figure 15

Diagrams illustrating the potential secondary structure elements (stem-loops) in the puf operon transcript of *Rb. capsulatus* (**A**) and the puf intercistronic stem-loop structures from *Rb. sphaeroides* (**B**). Adapted from Klug, (1993) and Del Ioff *et al.*, (1988).



Note : The predicted secondary structures present in the initial transcript from the *Rb. capsulatus* puf operon are illustrated in **A**, the lower part of the diagram illustrates their position with respect to the puf genes. The solid arrows point in the direction of the bulk of the transcript and away from the end, and the calculated free energy values for the structures are given above. The two stem-loops to the right lie at the extreme 5′ and 3′ ends of the pufBA transcript segment which is produced by degradation of the full transcript, and serve to protect the fragment from further degradation. The two loops to the right serve as transcription terminators and lie at the end of the pufLMX transcript segment, they also serve to protect the fragment from further degradation.

In **B** below the intercistronic stem-loop structures of the *Rb. sphaeroides* puf transcript are displayed, the upper part of the diagram illustrates the position in which they are found in the puf operon. The calculated free energy values for the structures are given at the side of each structure. It is suggested that the right hand structure acts as an attenuator, terminating transcription of certain transcripts. A more detailed description of the proposed attenuation mechanism is given in the main text.



Termination of transcription

Downstream of photosynthesis genes sequences resembling *E. coli* transcription terminators and attenuators (Platt, 1986) are often present (Chen *et al.*, 1988). Usually these elements are comprised of inverted repeats (palindromes) which are capable of forming secondary structure (stem & loop) in an RNA transcript once they have been transcribed[8], such structures are also important for the main post-transcriptional control mechanism identified thus far for photosynthesis genes (Fritsch *et al.*, 1995). The downstream stem-loops prevent degradation of the upstream transcript by RNase enzymes, thus increasing the lifetime of certain photosynthesis gene transcripts (Fritsch *et al.*, 1995), a more detailed explanation of this stabilisation role is given in the section "Post-transcriptional control" below.

The proposed mechanism of transcription termination by stem-loop forming dyad symmetry elements involves the formation of the RNA secondary structure as the DNA is transcribed, followed by termination of transcription at a run of Thymidine residues lying immediately downstream of the symmetry element (Platt, 1981). It has been suggested that the dyad symmetry element acts to slow down the polymerase (Gilbert, 1976), and that the relative instability of Adenine-Uridine bonds facilitates the release of the transcript (Martin & Tinoco, 1980).

Analysis of potential terminators in *Rb. capsulatus* has revealed that a run of as few as 4 Thymidine residues may be sufficient for termination, but that deviations below this level result in extensive transcriptional readthrough (Chen *et al.*, 1988). For example the puf intercistronic stem-loop (between pufA and L) is followed by the sequence CATA which appears to terminate (at most) 25% of transcripts (Belasco *et al.*, 1985). The 2 stem-loop structures lying downstream of *Rb. capsulatus* pufX more closely resemble transcription terminators, being followed by the sequences TATT and TTTT, and clearly function as such (Chen *et al.*, 1988). These structures are illustrated in part A of Figure 15 (opposite).

A comparable stem-loop found downstream of the pufA gene of *Rb. sphaeroides* appears to be far more effective at terminating transcription (DeHoff *et al.*, 1988; Lee *et al.*, 1989a). This structure is illustrated in part B of Figure 15 (opposite). Unlike *Rb. capsulatus* it has been suggested that the two major *Rb. sphaeroides* puf transcripts (pufBA and pufBALMX) may be transcribed from different promotors (Zhu *et al.*, 1986)[9]. Furthermore it has been proposed that the pufBA transcript is transcribed from the more distal of the two promotors and terminates at the stem-loop downstream of pufA, whilst transcription from the closer promotor appears to read through the intercistronic stem-loop and give rise to the larger pufBALMX transcript (Zhu *et al.*, 1986; DeHoff *et al.*, 1988). A mechanism involving ribosome stalling at the ribosome binding site of an untranslated leader ORF (termed orfK) upstream of pufB has been suggested (DeHoff *et al.*, 1988). In the proposed model the full length puf transcripts initiate too close to this ribosome binding sequence for ribosomes to bind, and thus

[8] These elements are also known as Repetitive Extragenic Palindromic sequences (REP's) (Stern *et al.*, 1984).

[9] Although a different report suggested only a single promotor further upstream (Hunter *et al.*, 1991), the model described may still be compatible, but with the two alternative mRNA 5′ ends considered as processing sites rather than transcription initiation sites.

translation begins at the start of pufB and remains coupled with transcription. The close proximity of the ribosome to RNA polymerase prevents the RNA from folding into a stem-loop as it is transcribed, allowing the polymerase to read through the terminator sequence. In contrast the transcripts initiated at the more distal promotor contain the upstream ribosome binding site, ribosomes bind to this site and stall, uncoupling transcription from translation. In this case the terminator is able to fold into its secondary structure following transcription and thus transcription terminates downstream of pufA. This mechanism is similar to the model for attenuation at secondary structure elements which lie upstream of *E. coli* amino acid biosynthesis operons (Oxender *et al.*, 1979), and the *Rb. sphaeroides* puf intercistronic terminator has thus also been viewed as an attenuator. Thus in the *Rb. sphaeroides* puf operon attenuation may in part give rise to the differential expression of the genes contained within the puf operon.

It should be noted that similar stem-loops have been identified in the puf operons of *Rs. rubrum* (Belanger & Gingras, 1988), *Rps. viridis* (Wiessner *et al.*, 1990), *Ro. denitrificans* (Liebetanz *et al.*, 1991), and *Rv. gelatinosus* (Nagashima *et al.*, 1994). Also in the *Rb. capsulatus* (Tichy *et al.*, 1991) and *Rb. sphaeroides* (Lee *et al.*, 1989b) puc operons; downstream of the crtK, crtC, crtI, crtB and crtF genes of *Rb. capsulatus* (Armstrong *et al.*, 1989), and downstream of the *Rb. sphaeroides* crtK, crtC, crtB and crtE (Lang *et al.*, 1994; Lang *et al.*, 1995). Even the multiple puc operons of *Rps. palustris* have dyad symmetry elements lying downstream (Tadros *et al.*, 1993). However it is not clear at present to what extent these structures act as transcription terminators, attenuators or transcript stabilisers.

Post-transcriptional control of photosynthesis gene expression

The main method of post-transcriptional control that has been identified thus far involves the degradation of mRNA transcripts by ribonucleases. Stem loop structures identified in *Rb. capsulatus* which lie between the pufA and L genes and downstream of pufX have been shown to stabilise puf mRNA, and to be essential for wild-type expression of puf genes (Klug *et al.*, 1987; Chen *et al.*, 1988; Belasco & Chen, 1988; Klug & Cohen, 1990). The secondary structures form part of the recognition site for endoribonuclease cleavage (Ehretsmann *et al.*, 1992; Mackie, 1992), and also protect the 5′ region from 3′-5′ exoribonuclease degradation (Chen *et al.*, 1988; Klug & Cohen, 1990; Fritsch *et al.*, 1995). Such control of transcript degradation appears to be responsible (at least in part) for the difference in abundance of the puf transcript segments, the 0.5 kb pufBA fragment (half life ~20 min.) being nine times more abundant than the initial 2.7 kb pufBALMX transcript from which it is derived (half life 4-5 min.) (Belasco *et al.*, 1985). This differential transcript stability clearly plays a role in determining the eventual molar ratio of LH1 to RC proteins (~ 16 :1) necessary for the correct RC-LH1 structure in *Rb. capsulatus* (see the section "Light harvesting complexes" for more information). A similar pattern is seen in *Rb. sphaeroides* where the 0.5 kb pufBA segment (half life ~20 mins.) is 10-15 times more abundant than the initial 2.6 kb pufBALMX transcript (half life ~9 mins.), *Rb. sphaeroides* also has a stem-loop structure 3′ to pufA which acts to stabilise the transcript (Zhu *et al.*, 1986).

Similar RNA secondary structure elements have been found in the *Rb. capsulatus* puc operon (Tichy *et al.*, 1991), suggesting that a similar mechanism is used to achieve the differing abundance of the pucBA, pucC and pucDE transcript segments. These elements are also present in the puf (DeHoff *et al.*, 1988) and puc (Lee *et al.*, 1989a) operons of *Rb. sphaeroides*, in the puf operon they have been shown to stabilise transcript segments in a similar manner to *Rb. capsulatus* (DeHoff *et al.*, 1988). In addition similar stem-loops have been identified in the puf operons of *Rs. rubrum* (Belanger & Gingras, 1988), *Rps. viridis* (Wiessner *et al.*, 1990), *Ro. denitrificans* (Liebetanz *et al.*, 1991), and *Rv. gelatinosus* (Nagashima *et al.*, 1994), and similar (long lived pufBA, short lived pufBALMX) transcripts identified in *Rs. rubrum* (Belanger & Gingras, 1988), *Rps. viridis* (Wiessner *et al.*, 1990) and *Rv. gelatinosus* (Nagashima *et al.*, 1994) - although in the latter case the transcripts are larger owing to the presence of extra ORFs next to pufB and A (see the section "The structure of the puf operon"). Additionally the Rb. capsulatus puc operon has a similar array of stem loop forming structures and a polycistronic transcript with segments of differential stability (Tichy *et al.*, 1991; LeBlanc & Beatty, 1993). The conservation of these elements suggests that the methods of puf transcript degradation, as well as the overall structure of the puf operon may have been conserved in many photosynthetic bacteria.

A more recent study has identified an *E. coli* like RNase E in *Rb. capsulatus* which may influence the rate of decay of photosynthesis gene transcripts (Fritsch *et al.*, 1995). In *E. coli* this enzyme cleaves mRNA transcripts at specific sites close to stem-loop structures (Ehretsmann *et al.*, 1992; Mackie, 1992). A consensus sequence for the RNase E cleavage site in *E. coli* has been deduced – (G/A)AUU(A/U) – with one or more cleavages occurring in the central AUU section (Ehretsmann *et al.*, 1992). A second consensus – GNYUUU – represents the cleavage site of *E. coli* RNase K (Lundberg *et al.*, 1990), which is a proteolytic fragment of RNase E (Carpousis *et al.*, 1994). Furthermore RNase E appears to initiate the decay of the bulk of transcripts in *E. coli* (Babitzke & Kushner, 1991; Melefors & von Gabain, 1991; Taraseviciene *et al.*, 1991), indicating that it influences the rate of decay of many transcripts.

It has been found that a sequence element in the pufLMX segment of *Rb. capsulatus* which influences rate limiting cleavage of the transcript segment shows homology to the RNase E cleavage sites of *E. coli*, and is cleaved by RNase *in vivo* in *E. coli* as well as in Rb. capsulatus (Fritsch *et al.*, 1995). This element consists of an RNA sequence – GGC´UUUU – and a downstream stem-loop, cleavage occurring between the C and the U (Fritsch *et al.*, 1995). On this basis a model of puf transcript decay has also been outlined (Klug, 1993; Fritsch *et al.*, 1995), the general principles of which should apply to other photosynthesis genes and to many other genes from a wide range of species. In this model the stem-loop structures serve to prevent degradation of the transcript by 3´ to 5´ acting exoribonucleases. The rate limiting step in puf transcript degradation must then involve endonucleolytic cleavage by RNase E at specific sites within the transcript. Following this initial cleavage secondary degradation processes occur very rapidly, probably because RNase E is associated with PNPase (phosphorolytic 3´ exonuclease polynucleotide phosphorylase) as has been found in *E. coli* (Carpousis *et al.*, 1994). The distribution of target sites and stem-loop

structures thus determines the stability of the various transcript segments.

The conservation of this degradation mechanism between two such diverse bacterial species and the fact that similar RNA secondary structures are evident downstream of genes in other photosynthetic bacteria (see above) and in bacteria in general (Stern *et al.*, 1984; Wong & Chang, 1986; Newbury *et al.*, 1987), indicates that differential transcript segment stability may be a widespread mechanism for differential gene expression (Fritsch *et al.*, 1995).

Of further interest is the discovery that stability of certain photosynthesis gene transcripts is affected by oxygen tension in both *Rb. capsulatus* (Klug, 1991a) and *Rb. sphaeroides* (Zhu *et al.*, 1986), indicating that transcript stability in the face of ribonuclease digestion has a role to play not only in ensuring the correct stoichiometry of proteins in the photosynthetic apparatus, but also in the oxygen regulation of photosynthesis gene expression. That exposure to oxygen accelerates the degradation of puf transcripts (Klug, 1991a; Zhu *et al.*, 1986) is evidence of another method by which photosynthetic bacteria respond to changes in oxygen tension, and provides a rapid mechanism by which synthesis of the photosynthetic apparatus can be shut down. This response also indicates that post-transcriptional control mechanisms extend further than differential transcript stability based on cis acting regulatory sequence elements, to include trans acting elements capable of responding to environmental stimuli.

Post-translational control of photosynthesis gene expression

Little hard evidence of post translational control mechanisms is available at present. However several observations indicate that post-translational control mechanisms are likely to exist. For instance, it has been found in that in mutant *Rb. sphaeroides* strains which do not contain coloured Car, the pucBA genes are transcribed and translated, but their products are not incorporated into the membrane, Car is essential for stability of and assembly of LH2 (Lang & Hunter, 1994). The assembly process appears to be very important, RC L, M and LH1 polypeptides are similarly unstable in the absence of photopigments or one of the constituent polypeptides (Dierstein *et al.*, 1984; Sockett *et al.*, 1989; Varga & Kaplan, 1993). Also LHC^- or Car^- strains often show altered membrane morphology (Hunter *et al.*, 1988; Kiley *et al.*, 1988; Golecki *et al.*, 1991; Gibson *et al.*, 1992; Lang & Hunter, 1994). All these observations underline the necessity for coordinated expression of photosynthesis genes and assembly of the various products into a viable photosynthetic apparatus. When these processes break down the gene products (with some exceptions) appear unstable and are rapidly degraded, thus post-translational mechanisms must exist to coordinate the assembly of the photosynthetic apparatus, and prevent unnecessary use of cellular resources when this is not possible.

Rubrivivax gelatinosus

In the previous sections I have described the present state of knowledge regarding the synthesis and structure of the purple bacterial photosynthetic apparatus. The work within this thesis was carried out using the purple bacterium *Rubrivivax gelatinosus*, initially this organism was known as *Rhodopseudomonas gelatinosa*, later

renamed and reclassified as *Rhodocyclus gelatinosus* (Imhoff *et al.*, 1984), and more recently as *Rubrivivax gelatinosus* (Willems *et al.*, 1991). *Rv. gelatinosus* is one of the few species of photosynthetic bacteria identified thus far which belongs to the β subclass of the Proteobacteria, most species belong to the α subclass (see the section "Taxonomy of the purple photosynthetic bacteria" for more details). This separate evolutionary lineage is also reflected in the structure and functioning of the *Rv. gelatinosus* photosynthetic apparatus. Protein sequencing of the LHC polypeptides and spectroscopic analysis indicate an overall similarity to species in the α subclass such as the *Rhodobacter* species, but also some interesting differences, most notably in the size of the LH2 α-polypeptide which appears to have two membrane spanning segments – a feature not observed in any other species (Zuber & Brunisholz, 1991; Brunisholz *et al.*, 1994). The puf operon has been cloned from *Rv. gelatinosus* strain IL 144 (see the section "Structure of the puf operon"), and also shows similarity to other purple bacterial species, but has additional genes (Nagashima *et al.*, 1994). Thus *Rv. gelatinosus* provides an interesting comparison to species of the α Proteobacteria.

The study detailed within this thesis was based on a difference noted between two strains of *Rv. gelatinosus* and described by Brunisholz *et al.*, (1994). The two strains DSM 149 and DSM 151 appear to differ in the ratio of RC-LH1:LH2 that they synthesise when cultured under identical conditions, strain 151 producing more LH2 than 149. Since this difference must be due to a difference (or differences) in the regulation of LHC synthesis it was thought that these two strains would provide an interesting comparison in terms of photosynthesis gene expression. Thus the initial aims were to clone and sequence the puc operon from both strains, and to delineate more precisely the physiological difference between the two strains. Sequencing would allow a comparison of cis elements surrounding the puc genes, and may indicate a simple difference between the two strains. The culture of the two strains under different regimes coupled with spectroscopic and northern hybridisation studies would allow the difference to be defined on a more precise basis, and may indicate whether any differences occur at the transcriptional level.

Materials and Methods

General laboratory reagents were purchased from the Sigma Chemical Company, Fisons Ltd. or BDH Ltd. unless otherwise stated. Company names are given in square brackets.

Culture of photosynthetic bacteria

Strains 149 and 151 of *Rv. gelatinosus* were cultured under various conditions, in all cases medium with Succinate as a carbon source was used, and a temperature of 28°C±2°C was maintained. A recipe for c-succinate medium can be found in Appendix 1.1. At the beginning of the project a glycerol stock (30% glycerol/70% liquid culture, mixed and stored at -70°C) was made from a culture obtained from a single colony, and this used as a source for the isolation of single colonies each time material was required for experimental work. All plating and culture transfer was carried out in a laminar air-flow hood using sterile media and equipment.

Anaerobic solid culture

Initial cultures were grown in sterile plastic petri dishes on 1.5% agar supplemented with yeast extract (0.3%) and casamino acids (0.2%). These plates were placed in GasPaK® containers [BBL Microbiology Systems] to provide an anaerobic environment, and illuminated with incandescent bulbs. Following growth a single colony was selected, streaked onto a second agar plate, and a single colony from this plate used to inoculate stab cultures (in 25 ml McCartney bottles) half filled with the same media. Following a minimum of 48 hours growth succinate media was used to fill the bottle and the culture incubated further 48 hours, the resulting liquid culture was then used to inoculate larger liquid cultures as described below.

Anaerobic liquid culture

100 or 500 ml flat sided glass bottles containing sterile c-succinate medium were inoculated with 5 to 20 ml of suspended *Rv. gelatinosus* from a stab culture or from a previous liquid culture. Culture then occurred at 28°C±2°C and at the irradiance level stated for each experiment. Different irradiance levels were achieved by placing the 100 ml culture bottles at different distances from the light source, and by using different numbers of incandescent bulbs for the light source. High irradiance cultures were centred between two banks of three 100 W household light bulbs (total six), medium irradiance cultures were placed to one side of a bank of three bulbs, whilst the lowest irradiance level was achieved by placing the culture bottle in front of a single bulb (100 W). The flat side of the bottle was always facing the light source. The irradiance level was measured using a UDT 40X Opto-meter [United Detector Technologies], which measures light energy over the range 400 to 800 nm. The three irradiance levels used were measured to be 100, 40 and 18 W m^{-2}, and will be referred to as high-light (HL), medium-light (ML) and low-light (LL) throughout the thesis.

Aerobic and semi-aerobic liquid culture

These conditions were achieved by dark culture in 250 ml conical flasks, the semi-aerobic cultures contained 200 ml of c-succinate medium and the aerobic cultures 100 ml. Both types were incubated in a orbital shaker at 140 rpm, 28°C, the difference in volume and therefore aeration gives rise to a difference in oxygen tension.

Cell counts

Cells were counted in a number of samples from a growth curve experiment as follows. The samples were diluted in iodine/glycerol (50% aqueous Iodine/50% Glycerol) to both stain and immobilise them, dilution was to a level which would allow easy counting of the cells. The solution was then placed on a microscope slide and cells counted using a haemocytometer grid. The grid had a volume of 1 mm^3 and was divided into 25 squares, the bacteria in 5 of these squares were counted, then the mean calculated multiplied by 25000 and by the dilution factor to give the number of cells per ml. The cells were visible as darkly staining rods (or spheres when viewed head on), and were often moving slowly.

Absorption spectroscopy

Whole cell absorption spectra were measured on a Shimadzu Corporation UV-2101PC recording spectrophotometer fitted with an integrating sphere attachment to reduce the effect of scattering in particulate samples. The 3 ml quartz reference cuvette was filled with sterile c-succinate media from the same batch used to culture the bacteria, and the spectrophotometer parameters set as follows: Recoding range - 0 to 0.8 absorption units cm^{-1}; Wavelength range - 900 to 350 nm; Scanning speed - very slow; slit width - 5.0 nm; Sampling interval - 1 nm. Absorption was always measured across a 1 cm path length.

A Shimadzu UV-160A recording spectrophotometer was used for the assay of specific absorbance values where necessary, in particular for the measurement of $A_{260\,nm}$ / $A_{280\,nm}$ ratios used in the determination of nucleic acid concentration and purity. A three ml quartz cuvette was used to measure absorbance at 260 and 280 nm and the concentration calculated as described below.

Molecular biological techniques

Determination of Nucleic acid concentration

Following the measurement of $A_{260\,nm}$ as described above the calculation of DNA, RNA and oligonucleotide concentration and purity was made using the following equation:

$$(A_{260\,nm} - A_{280\,nm}) \times 2 = \text{Adjusted } A_{260\,nm}$$

This adjusted $A_{260\,nm}$ value was then multiplied by: 50 μg ml^{-1} for DNA, 40 μg ml^{-1} for RNA, and 20 μg ml^{-1} for oligonucleotides as described in Appendix E5 of the

Handbook of molecular cloning (Sambrook *et al.*, 1989).

When too little sample was available for spectroscopic determination, the Ethidium Bromide plate method was used. This is also described in the Handbook of molecular cloning, Appendix E5 (Sambrook *et al.*, 1989), the concentration was estimated by spotting 1 or 0.5 µl of the sample onto a thin 1% agarose gel containing Ethidium Bromide, illuminating using a UV transilluminator and comparing the intensity of the dot with those of lambda DNA of a known concentration.

Restriction endonuclease digests

Restriction enzymes were purchased either from Gibco-BRL Ltd., Pharmacia Ltd., or Promega Corporation. Digests were carried out in the buffer supplied by the manufacturer and at the specified temperature (usually 37°C). In the case of multiple digests (using more than one enzyme and digesting in the same buffer), a compatible buffer was used if specified by the manufacturer or a medium salt buffer such as 1-4-all buffer (Pharmacia Ltd.). Incubation was for 1 hour unless otherwise specified.

Agarose gel electrophoresis

Agarose gel electrophoresis was used to size-separate DNA and RNA fragments. DNA (or RNA) in solution is negatively charged at neutral pH, thus when loaded into a gel near the cathode and an electric current applied down the gel it moves towards the anode. The agarose gel fibres slow the progress of the DNA but differentially slow movement of fragments of different sizes. Thus electrophoresis separates different sized nucleic acid fragments. A full explanation of this technique in both practical and theoretical terms is given in Section 6.3 of the Handbook of molecular cloning (Sambrook *et al.*, 1989) and by Ogden & Adams, (1989).

Electrophoresis tanks of various sizes were purchased from Gibco-BRL Ltd. and Anachem Ltd, a model 400H power supply [Gibco-BRL Ltd] was used to run the gels. Agarose was purchased from Gibco–BRL Ltd and melted using a microwave oven. For DNA Tris-Borate-EDTA (TBE) gel buffer was used (Ingredients given in Appendix B23 of the Handbook of molecular cloning (Sambrook *et al.*, 1989), whilst RNA gels used a denaturing buffer (formaldehyde gel running buffer) containing formaldehyde and formamide, as described by Ogden & Adams, (1987). Additionally RNA samples contained formaldehyde and formamide and were heated to 65°C to denature before loading. Samples were also mixed with loading buffer before loading (15% Ficoll 400, 0.25% Bromophenol Blue and 0.25% Xylene cyanol in water), and were stained within the gel by immersing the gel in a solution of Ethidium Bromide at a concentration of 0.5 µg ml^{-1} for 15-30 minutes, followed by immersion in deionised water for approximately 15 minutes. Agarose concentration, voltage applied to the gel and staining/destaining times were varied according to experimental requirements, exact parameters are given with particular experiments in the Results sections.

Photography of agarose gels

Following staining (and destaining where necessary) gels were placed on a UV transilluminator [UVP, inc.] and photographed with a 35 mm camera [Pentax] fitted with a x2 close up filter [Jessops] and a red(25A) filter [Hoya Corp.]. Exposure times of 30 s, 1 minute, and 1 minute 30 s were used and the best exposure for a particular gel was developed.

Southern and Northern transfer

Both RNA and DNA were transferred to Hybond-N nylon membrane [Amersham International plc] by capillary transfer, using modifications of the techniques described by Southern, (1975) and Meinkoth & Wahl, (1984).

Southern (DNA) transfer to nylon membrane was carried out in accordance with the manufacturers instructions, and essentially as described in section 9.42 of the Handbook of molecular cloning (Sambrook *et al.*, 1989), where details of buffer preparation and equipment setup can be found. Where fragments larger than 10 kb were to be transferred the gel was first soaked in 0.25% HCl to partially hydrolyse the DNA and thus ensure the transfer of such fragments. Soaking time was in the region of 20 minutes, or 10 minutes after the dyes in the loading buffer changed colour. If all fragments to be transferred were smaller than 10 kb this step was omitted and the gel soaked in denaturing solution (1.5M NaCl, 0.5M NaOH) for 45 minutes. The gel was then soaked in neutralising solution (1.5M NaCl, 0.5M Tris-HCl pH 8.0) for 30 minutes. Transfer was carried out in 6x SSC buffer diluted from a 20x stock solution (3M NaCl, 0.3M Na_3 citrate) using 3MM paper [Whatman Paper Ltd.] as a wick. Following the transfer the membranes were washed in 2x SSC and allowed to air dry, DNA side up, before fixing.

Several Southern blots were carried out using the Alkaline transfer protocol described in Section 9.45 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). In this case the denaturation step was shortened to 30 minutes and the neutralising step omitted. Transfer was carried out using alkaline transfer buffer (0.25M NaOH, 1.5M NaCl). This protocol has the advantage of speed, reducing the amount of time that DNA has to diffuse through the gel before transfer begins.

RNA was transferred to Hybond-N using 20x SSC buffer, the gel being given no pre-treatments except for destaining (30 minutes). Following transfer the filter was rinsed with 2x SSC, then dried and fixed as for DNA containing gels.

Following blotting both DNA and RNA were fixed to the membrane by UV crosslinking, filters were covered with catering grade 'cling-film', placed DNA side down on a UV transilluminator, and illuminated for 2 minutes.

Preparation of nucleic acid probes

Three basic types of DNA probes were used in the analysis of Northern and Southern blots, that is fragments of DNA obtained by restriction enzyme digestion and purified from agarose gels, fragments of DNA obtained by PCR, and synthesised oligonucleotide probes.

A 29 base oligonucleotide homologous to the 5′ end of the pucA gene was used

for initial southern hybridisations and during the construction of the strain 151 genomic library. This oligonucleotide was designed by Dr. Stuart Barrett of the University of Glasgow for the analysis of puc genes in *Rps. acidophila*, and is homologous to the 5′ end of the *Rps. acidophila* pucA genes (see the Introduction section "Structure of the puc operon"), and is known as the 'α-consensus oligonucleotide'. Since the protein sequence of the *Rv. gelatinosus* LH2 α-polypeptide had been published (Zuber & Brunisholz, 1991), it was possible to calculate that the oligonucleotide would be between 72% and 93% homologous to the pucA gene of strain 151, and thus be suitable for use in isolation of that gene. The α-consensus oligonucleotide is degenerate, its sequence is given here:

3′ TAC TTG GT(T/C) CCG TT(T/C) TAG ACC TGG CAG CA 5′

the residues in brackets are degenerate, and represent the two possibilities at each point (thus the α-consensus oligonucleotide is actually a mixture of 4 different molecules). The most leftward triplet is the antisense strand of the ATG start codon (Methionine) of pucA.

This oligonucleotide was labelled with ^{32}P using T4 polynucleotide kinase [Boehringer-Mannheim] according to the protocol given in Appendix 3.1 and described in Section 11.31 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). T4 polynucleotide kinase phosphorylates a free 5′-OH group of a DNA molecule by transfer of the γ-^{32}P phosphate from [γ-^{32}P]ATP. Following the labelling reaction the oligonucleotide was purified away from free radiolabel by passage through a Sephedex G-25 chromatography column as described in Section 11.37 of the Handbook of molecular cloning (Sambrook *et al.*, 1989), once eluted it was added to the hybridisation mixture.

A fragment encoding part of the *Rb. sphaeroides* pucC gene was also used as a probe, this was cut from a plasmid construct known as pUBS+pPUCF$_2$BamEco3, kindly supplied by Dr. Lucien Gibson of the University of Sheffield. Double digestion of this plasmid with BamHI and PstI yields a 1.45 kb fragment carrying most of the pucC gene (the gene is truncated at its 3′ end), this fragment was separated from such a digest by electrophoresis in a low melting point agarose gel run with TAE (Tris-Acetate-EDTA) buffer as described in Appendix B23 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). A slice of gel carrying the 1.45 kb fragment was excised from the gel and the fragment extracted using a Geneclean II kit [Bio 101 Inc.] according to the manufacturers instructions. This kit uses a silica matrix which binds DNA allowing contaminants to be removed, once the fragment has been eluted from the matrix it is pure enough for labelling.

Probes were also produced using the polymerase chain reaction (also see section on PCR). These include a fragment carrying the pucBA genes from strain 151 and a second fragment carrying most of the pucC gene from strain 151, these were produced as described in the section on PCR below.

Both PCR produced probes and probes excised from gels were labelled by Random priming. A protocol for the Random prime labelling of DNA is given in Appendix 3.2, based on that described by Feinberg & Volgelstein, (1983). This involves the binding of random hexanucleotides to denatured double-stranded DNA,

followed by a 'fill in' reaction using the Klenow fragment of *E. coli* DNA polymerase I, and radiolabelled nucleotides. Samples were cleaned up after the labelling reactions using Magic™ DNA clean-up minicolumns [Promega Corp.], as specified in the manufacturers instructions. This is a resin based system which reversibly binds DNA, such that contaminants are either removed during a wash step (80% isopropanol) or remain in the column when DNA is eluted. The probes were eluted from the Magic columns in 50 µl of sterile analar water and denatured using the alkali denaturation method as described by Mason & Williams, (1987). Alkali denaturation was achieved by adding one tenth volume of 3M sodium hydroxide and leaving for 5 minutes at room temperature. The sample was then neutralised by the addition of one fifth of the original volume of 1M Tris-HCl pH 7.0, and one tenth of the original volume of 3M HCl. The probes were then ready for addition to the hybridisation solution.

Nucleic acid hybridisation

A full discussion of the kinetics and theory of nucleic acid hybridisation can be found in Britten & Davidson, (1987), as can practical advice on the techniques used in hybridisation.

Two slightly different hybridisation protocols were used depending on whether the probe being used was an oligonucleotide or a double-stranded DNA probe. These are described in Section 9.47 of the Handbook of molecular cloning (Sambrook *et al.*, 1989), and recipes for solutions given in Appendix B4 of the same handbook, the protocols are outlined below.

All hybridisations were carried out in sealable plastic containers or in sealed plastic bags, and incubated in a shaking waterbath. The volume of hybridisation solution added was approximately 1 ml for every 5 cm^2 of filter being probed. For hybridisation utilising oligonucleotide probes filters were first placed directly into prehybridisation solution (4x SET buffer, 10x Denhardts solution, 0.1% SDS and 0.1% sodium pyrophosphate) and prehybridised at 48°C for a minimum of 2 hours. This step saturates binding sites on the membrane which would otherwise give rise to high background levels caused by non-specific binding of the probe. Hybridisation was carried out by raising the temperature to the desired level, in the case of the α-consensus oligonucleotide this was 56°C, and adding the labelled oligonucleotide to the prehybridisation solution. Following hybridisation, which was usually carried out overnight, the filters were washed. The hybridisation solution was poured off and prewarmed wash solution 1 (3x SET buffer, 10x Denhardts, 2% SDS and 1% sodium pyrophosphate) immediately added. This was then incubated in a shaking water bath (at the hybridisation temperature) for 30 minutes, following which the wash solution was replaced with prewarmed wash solution 2 (1x SET buffer, 1% SDS and 0.1% sodium pyrophosphate), and incubated with shaking as before. After 30 minutes the second wash solution was poured off and the filters treated as described at the end of this section.

Filters being probed with random prime labelled DNA probes were hybridised in a different manner. Prehybridisation solution (4x SET buffer, 10x Denhardts solution, 0.1% SDS and 0.1% sodium pyrophosphate) was used as for the

oligonucleotide probed filters, however for the *Rv. gelatinosus* strain 151 pucBA probe and pucC probe an incubation temperature of 65°C was used, whilst the *Rb. sphaeroides* pucC probe was prehybridised and hybridised at 55°C. Prehybridisation was for a minimum of 2 hours, hybridisation was carried out overnight. Two washes of 20 minutes duration with wash solution 1 (3x SET buffer, 0.1% SDS and 0.1% sodium pyrophosphate) at the hybridisation temperature, were followed by a single wash of 20 minutes in wash solution 2 (1x SET buffer, 0.1% SDS and 0.1% sodium pyrophosphate). When a more stringent wash was required a 0.1% SET containing wash solution was used as the second wash.

Following washes filters from both types of hybridisation were wrapped in cling film (without drying, to facilitate reprobing where necessary), and autoradiographed as described below.

Autoradiography

Autoradiography of radiolabelled samples was carried out in cassettes [GRI Ltd.]. For Northern and Southern blots, which were investigated using probes labelled with the isotope ^{32}P, the cassettes were fitted with High-Speed-X intensifying screens [GRI Ltd.] and autoradiography was carried out at -70°C. For DNA sequencing the samples were labelled with ^{35}S, in this case the intensifying screens were removed and autoradiography was carried out at room temperature. In all cases Fuji RX medical X-ray film [Fuji Photo Film Co., Ltd.] was used, and was developed in an X-Omat automatic film processor [Kodak Ltd.]. Exposure times were varied according to the requirements of individual experiments.

Oligonucleotide design and preparation.

Oligonucleotides for use in PCR, DNA sequencing and as probes were designed avoiding sequences which could potentially dimerise or form loops. A GC content of 50% was attempted, but this proved difficult owing to the high GC content of the *Rv. gelatinosus* genome (~72%) (Willems *et al.*, 1991). A check was also made for other potential binding sites in the target sequence (when target sequence information was available). Primers were synthesised in house by V.B. Math of the University of Glasgow. They were prepared by ethanol precipitation and resuspended in sterile analar water, the concentration was then calculated and the primers diluted to a useful concentration before storage at 4°C.

Culture of *E. coli*

Two strains of *E. coli* were used, DH5α was used for the transformation and propagation of plasmid DNA, whilst KW251 was used in the propagation of phage λ DNA. The strain genotypes are given in Appendix 4. Glycerol stocks of these strains were maintained at -70°C and streaked out onto LB agar when required, isolated single colonies were then used for the production of larger cultures. Details of LB agar and LB medium composition can be found in Appendix A1 of the Handbook of molecular cloning (Sambrook *et al.*, 1989).

Small scale cultures (~3 ml) of *E. coli* were grown in 50 ml falcon tubes

containing sterile LB medium and stoppered with a cotton wool plug, these were incubated in a orbital shaker at 37°C and 140 rpm overnight. Where necessary the appropriate antibiotic was added to the medium. Large scale cultures were grown in one litre glass conical flasks containing 250 ml of LB and incubated under the same conditions as the small scale cultures. Again, antibiotics were added where necessary. For medium scale cultures a 100 ml conical flask with 10 or 20 ml of LB was used.

Culture was generally carried out overnight, with 2 notable exceptions. When DH5α cells were required for transformation their A_{600} was monitored and cells harvested at or near an A_{600} of 0.125 cm^{-1}. This is equivalent to 1 x 10^8 cells ml^{-1}, such a culture consists almost entirely of young cells, which are required for efficient transformation. The second exception was during the culture of phage λ, which were harvested when lysis occurred, this is described in more detail in the section "Methods used in genomic library construction" below.

Plasmid isolation and purification

Small scale preparations of plasmid DNA were isolated from 3 ml cultures. 1.5 ml of cells were harvested at 13000 rpm for 5 minutes in a microcentrifuge, and DNA isolated according to the protocol described in "Lysis by boiling" Section 1.29 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). During this protocol genomic DNA, protein and ribosomal RNA become congealed into a lump which can be removed from the tube, leaving behind plasmid DNA, tRNA and mRNA. A simple RNase treatment followed by precipitation with isopropanol and resuspension in TE pH 8.0, yields plasmid DNA of a quality suitable for restriction enzyme digestion. This method was used to obtain plasmid DNA for screening following transformation.

Large scale preparations of plasmid DNA were obtained from 250 ml cultures, according to the Alkaline lysis method of (Birnboim & Doly, 1979), and essentially as described in Section 1.38 of the Handbook of molecular cloning (Sambrook *et al.*, 1989), the exception being that since 250 ml cultures were used (as opposed to 500 ml) all volumes quoted in the the Handbook of molecular cloning were halved. Cultures were harvested in a Beckman JA-10 rotor at 4000 rpm, 4°C for 10 minutes. Following washing the cells were resuspended in lysis buffer and transferred to a sterile 50 ml falcon tube. Later centrifugation steps were carried out in the 11133 rotor of a Sigma 3K12 centrifuge, and the pellets finally resuspended in 8 ml of sterile analar water.

Following isolation by alkaline lysis the DNA was purified by equilibrium centrifugation in Ceasium chloride/Ethidium bromide continuous gradients, as described in Section 1.42 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). The high concentration of Ethidium bromide present denatures protein and binds to any nucleic acids present, however RNA plus linear or nicked plasmid DNA bind more ethidium than circular supercoiled plasmid DNA. This causes them to equilibrate at a different point on the column, and thus allows the separation of highly pure undamaged plasmid DNA. Prior to equilibrium centrifugation, but after the addition of all components (in a falcon tube), the samples were centrifuged at 5000 rpm in the 11133 rotor of a sigma 3K12 centrifuge at room temperature for 5 minutes. Equilibrium centrifugation was carried out using disposable 11.5 ml Ultracrimp

polyallomer tubes [Sorvall], in a T865.1 rotor [Sorvall] at 45000 rpm, 21°C for 24 hours in a Sorvall OTD65B ultracentrifuge. The sample (usually visible as a darker band in daylight) was removed by piercing the side of the tube with a needle and syringe and removing the band. Ethidium bromide was removed from the sample by extraction against an several volumes of water saturated butanol as described in Section 1.46 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). Finally the sample was dialysed against 5 changes of 5 litres of TE pH 8.0, with at least 1 hour between changes. DNA isolated in this manner was suitable for all enzymatic treatments including DNA sequencing and PCR.

Genomic DNA isolation and purification

Genomic DNA was isolated by lysis with lysozyme followed by protein digestion (using proteinase K), and phenol/chloroform extraction. Great care was taken at all stages to avoid contamination and fragmentation since this DNA was to be used as a source for the Genomic library, sterile plasticware was used extensively and any equipment not used for the first time was washed extensively with sterile deionised water.

DNA was isolated from pure cultures of both strains using 200 ml of a 'HL' 500 ml liquid culture, grown for 48 hours in the case of strain 151 and 24 hours for strain 149. High molecular weight genomic DNA was isolated according to the methods of (Birnboim & Doly, 1979) and (Klug & Drews, 1984), as described in depth by MacKenzie, (1990). Cells were harvested in sterile plastic Falcon tubes [Elkay] by centrifugation at 5000g, 4°C, in the 11133 rotor of a 3K12 bench top centrifuge [Sigma Ltd.] for 10 minutes.

Both strains of *Rv. gelatinosus* used in this study, but especially strain 151, secrete a 'muco-polysaccharide' into the medium, in which many cells appear to become trapped. This phenomenon appears to occur as the cultures are reaching the stationary phase of the growth curve, and can by avoided in the main by harvesting cells before they reach their full density. However a small amount is produced at all stages of culture and this resulted in the production of a 'two-tier' pellet following the centifugation stage described above. The lower tier contained a relatively hard pellet of cells (as judged by its dark brown colour) whilst the upper tier appeared translucent and to consist of the muco-polysaccharide, this upper tier was poured off and discarded before continuing with the isolation procedure.

The pellets were resuspended in 10 ml of Lysis buffer (%50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA pH 8.0), and re-pelleted using the same centrifugation conditions. They were then resuspended again in 10 ml Lysis buffer, which acts to support the cell walls during subsequent lysozyme treatment. Once resuspended 1 ml of lysozyme solution (100 mg ml^{-1} Lysozyme [Sigma] in Lysis buffer) was added. The cells were then incubated in a water bath at 28°C for 30 minutes. Following this incubation period 1 ml Proteinase K solution (5 mg ml^{-1} in Lysis buffer) and 50 µg DNase free RNase was added and mixed by gently inverting the falcon tube 6 times. These digest RNA and protein and thus aid the isolation of pure DNA. The detergent SDS was then added from a 10% stock solution to give a final concentration of 1%, at

this point the solution became extremely viscous as the cell contents were released. The SDS was mixed in by 6 gentle inversions of the falcon tube, and the tube incubated for 1 hour at 56°C (the optimum temperature for proteinase K activity).

Genomic DNA was purified from this mixture by extraction with phenol/chloroform (prepared using Analar grade phenol and chloroform and buffered to pH 8.0 as described in Appendix B4 of the Handbook of molecular cloning (Sambrook *et al.*, 1989) and dialysis against TE buffer (composition of TE buffer is given in Appendix B20 of the Handbook of molecular cloning (Sambrook *et al.*, 1989)). An equal volume of phenol/chloroform was added to the lysed cells and mixed by placing on a Multimix MMJ rocking and rolling table [Luckham] for 1 hour. The mixture was then centrifuged at 5000g, 20°C for 15 minutes in the 11133 rotor of a 3K12 bench top centrifuge [Sigma Ltd.]. The aqueous (upper) phase was removed using a truncated 1 ml micropipette tip (to avoid shearing the DNA), into a fresh sterile 50 ml falcon tube and a second extraction performed.

To remove any remaining traces of phenol and the other extraction chemicals the samples were dialysed against 5 changes of 5 litres TE buffer pH 8.0 (25 mM Tris-HCl, 10 mM EDTA, pH 8.0), with a minimum of 1 hour between changes at room temperature, or overnight at 4°C. Following dialysis the DNA was transferred to a sterile falcon tube and stored at 4°C until used.

Isolation and purification of total RNA

RNA was isolated from cells grown under a range of growth conditions using a modified version of the "acid guanidium thiocyanate-phenol-chloroform extraction" protocol described by Chomczynski & Sacchi, (1987). This utilises guanidium thiocyanyate which acts as a deproteinisation agent and inhibits ribonucleases which may degrade the samples.

RNA was isolated from cells in early log phase to ensure that no self-shading occurred (thus altering the irradiance conditions), and also to ensure that strain 151 produced a minimal amount of mucus, which could potentially have interfered with the RNA extraction. Three 1 ml aliquots were removed from the culture bottles into sterile 1.5 ml microfuge tubes and stored at -20°C, absorption spectra were recorded from these samples at a later point. The remaining culture solution (from a 100 ml bottle) was used for the preparation of RNA. All equipment was either sterile plasticware, or had been treated with di-ethylpyrocarbonate (DEPC) before use, to ensure the removal of ribonucleases. A protocol for the DEPC treatment of apparatus intended for use in the preparation of RNA is described by Blumberg, (1987).

Cells were harvested by centrifugation at 10000 rpm in the 12156 rotor of a Sigma K20 centrifuge, for 4 minutes. During this period a pestle and mortar were cooled by pouring liquid nitrogen into the bowl of the pestle. Following centrifugation the supernatant was discarded, and a sterile 3 ml plastic pipette used to add the cells one drop at a time to a pool of liquid nitrogen in the pestle. The cells were then ground to a powder and transferred to a 15 ml glass centrifuge tube [Corex] containing 5 ml of Solution D (a recipe for the production of Solution D is given in appendix 2.1). These were mixed together using the plastic pipette before adding 0.5 ml of 2M Sodium

acetate (pH 4.0), 5 ml of water saturated phenol and 1 ml of chloroform (water saturated phenol was prepared as described in Appendix 2.2). Centrifugation at 10500g , 4°C in the 12156 rotor of a Sigma 3K12 centrifuge for 10 minutes, was usually sufficient to pellet the protein at the phenol/aqueous phase interface. On occasion the protein failed to separate fully at this first step and a second 10 minute centrifugation was carried out. The aqueous phase was then transferred to a fresh corex tube and the phenol extraction procedure repeated twice more, before extraction in 5 ml of chloroform. Following this final extraction the aqueous phase was transferred to a fresh tube and an equal volume of Isopropanol added and mixed, the sample was then placed at -20°C for 30 minutes to precipitate the RNA. The RNA was then collected by centrifugation at 10500g, 4°C for 10 minutes, and the pellet washed with 80% Ethanol (20% DEPC treated water). Finally the pellet was air dried for ~5 minutes and resuspended in 100 µl of DEPC treated sterile analar water. RNA samples were stored frozen at -70°C until used.

Since *in vivo* degradation of RNA transcripts occurs continuously, it was important to ensure that RNA was isolated from the bacteria as quickly as possible after their removal from the growth chamber. Once the bacteria were added to the liquid nitrogen (as described above), the worst of this danger is passed. The time taken to reach this point was between 6 and 10 minutes for all RNA samples prepared.

Methods used in Genomic library construction

Some considerations

A good discussion of the construction of genomic libraries is given in Kaiser & Murray, (1985), and relevant information concerning bacteriophage λ can also be found in Section 2.3 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). Both these sources also provide technical advice and detail protocols useful in genomic library construction. A genomic library is defined as collection of clones which together encompass the complete genome of an organism. For an organism such as a bacterium, with a relatively small genome (*Rb. capsulatus* is approximately 3.8 Mb - see the Introduction for references), a lambda replacement vector is a useful vehicle for a genomic library. The ability to clone segments of up to 20 kb means that the complete genome of an organism such as *Rb. capsulatus* can be contained by less than 1000 clones, to screen such a number of clones in the search for a gene is an achievable target.

Although the use of commercial kits and prepared vectors has taken some of the work out of genomic library construction, it remains important to ensure that the library is truly representative of the source genome. Care must be taken in the initial stages to prevent contamination of the source material with alien DNA, and at later stages to rule out the possibility of recombination within and between clones. The latter problem is addressed by the use of vectors and bacteria which are impaired in the functioning of several cellular DNA manipulation systems, this will described in more detail below. The former problem is addressed in two ways, firstly by ensuring that the source bacterial culture is pure (see the section "Culture of photosynthetic bacteria" above), and secondly by using acid washed glass ware, or virgin plasticware during the course

of DNA isolation.

Isolation of high molecular weight genomic DNA

DNA was isolated from the two strains of *Rv. gelatinosus* as described in the section "Genomic DNA isolation and purification" above. At all stages the DNA was treated gently to avoid shearing, ideally genomic DNA should consist of fragments well in excess of 50 kb in size because it will be further reduced in size in the next step.

Digestion with Sau3AI

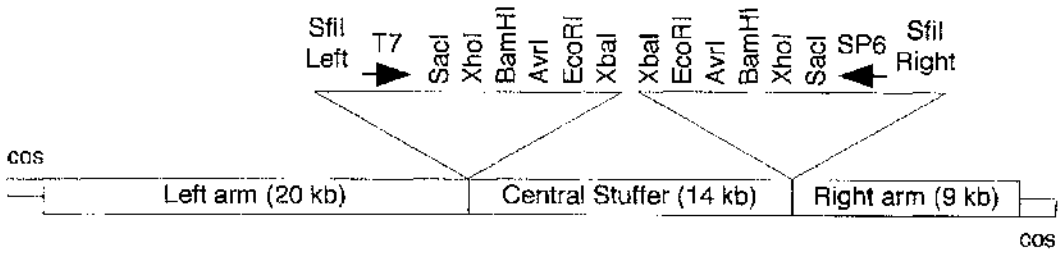
The wild-type bacteriophage λ genome is 49 kb long, deletion of genes not essential for replication allows the insertion of a foreign 20 kb fragment between the two λ 'arms', variations above or below this size lead to inefficient replication. Thus it is necessary to produce a sample of genomic DNA consisting of fragments at or near 20 kb in length, furthermore these fragments should have ends compatible with those of the λ vector arms, to facilitate their ligation.

The isolated DNA from strain 151 was digested with a range of enzymes, of which Sau3A1 proved suitable for the production of 20 kb fragments. This enzyme cuts very frequently and a partial digest must be carried out in order to yield fragments of the correct size, in this case the partial digest was achieved by performing a time course digestion. An initial pilot digest was carried out to determine the digest time which yielded the highest proportion of 20 kb fragments. Two premixtures were constructed, one containing enzyme (described in Appendix 5.1) which was incubated for 15 minutes on ice and then warmed to 37°C for 5 minutes, and a second containing genomic DNA (Appendix 5.2) which was incubated in water bath at 37°C for 20 minutes. Following the incubation period the two mixtures were mixed thoroughly using a micropipette, and a zero time point aliquot (39 μl) immediately removed into 0.5M EDTA pH 8.0 (4.2 μl) and incubated at 65°C for 10 minutes to stop the digestion. Further aliquots (42 μl) were removed into EDTA at time points over the next hour, specifically at 1, 2, 3, 4, 5, 7.5, 10, 12.5, 15, 20, 45 and 60 minutes after the start of digestion.

Analysis by agarose gel electrophoresis indicated that a digest time of 8 to 12 minutes yielded the best sized fraction (The gel can be seen in the section "Results Part 2: Construction of Genomic Libraries & Cloning of the puc Genes"). The next step was to scale up the digest for the production of sufficient quantities of 20 kb fragments to make a library. The volume of all ingredients was increased such that the digest could be carried out in a 1.5 ml microfuge tube, the components of the larger mixtures are given in Appendices 5.3 and 5.4. An initial attempt at the scaled up digest did not work well, probably due to a lack of mixing between the two solutions. Thus it was repeated, paying special attention to the mixing of the two solutions at the start of the digest. During the first repeat of the large scale digest aliquots were removed over the course of 1.5 hours, at 2, 4, 8, 16, 20, 25, 30, 35, 40, 60, and 90 minutes. Electrophoretic analysis of these samples revealed an optimum digestion time of 4 to 8 minutes, thus the digestion process was repeated 2 further times removing aliquots only between 4 and 10 minutes (for strain 149 2 to 4 minutes appeared best).

Figure 16

The λGEM-11 vector.



Note: The λGEM-11 vector is a derivative of the λEMBL3 vector (Frischauf *et al.*, 1983). The multiple cloning sites (expanded to show detail) contain recognition sites for several restriction enzymes, flanking this are SfiI sites. SfiI is a rare site cutting enzyme and potentially allows the insert DNA to be excised from the vector and mapped via an end labelling protocol. The bacteriophage T7 and SP6 promoters facilitate the production of RNA probes for use in genome 'walking'. The overhanging 12 bp cohesive ends of λ are labelled cos.

An aliquot of the samples from each digest was analysed by agarose gel electrophoresis and all the samples containing a substantial proportion of 20 kb fragments were pooled for use in the next step of the procedure.

Size fractionation of Sau3AI digested genomic DNA

The digested genomic DNA was size fractionated using sucrose density gradient centrifugation, this is described in Section 9.27 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). Sucrose solutions containing 10 and 40% sucrose (plus 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0 and 10 mM NaCl), were used to pour continuous gradients as described by MacKenzie, (1990). The DNA sample was heated to 68°C for 10 minutes then cooled on ice before loading onto the gradient. Following centrifugation at 22000 rpm in a Sorvall AH629 swing-out rotor for 22 hours, the solution was gently siphoned off through capillary tubing in ~0.5 ml aliquots, as described by MacKenzie, (1990).

The collected aliquots were analysed by agarose gel electrophoresis (0.4% gel), 30 µl of each aliquot was mixed with an equal volume of sterile analar water, load buffer added and the sample heated to 65°C for 10 minutes before loading. An aliquot or aliquots which were shown to contain fragments of the correct size were selected and purified by ethanol precipitation.

Purification of size fractionated genomic DNA

The selected aliquots were dialysed against 3 changes of 5 litres of TE pH 8.0, then transferred to microfuge tubes. A tenth volume of 3M sodium acetate pH 5.2 and 2 volumes of ethanol was added and mixed. The tubes were then placed at -20°C for 30 minutes, following which they were centrifuged at 21000g, 0°C for 1 hour in the 12154 rotor of a Sigma 3K20 centrifuge. The resulting pellets were washed with 70% ethanol then vacuum dried before resuspending in 10 µl of sterile analar water. The DNA sample was then ready for ligation to the vector.

Ligation to the vector arms

Before describing the method used to ligate the insert genomic DNA to the vector DNA I will first give a brief description of the vector used. The vector chosen for the construction of both libraries was λGEM-11 [Promega Corp.], this is a derivative of the widely used λEMBL3 vector (Frischauf *et al.*, 1983). Figure 16 (opposite) illustrates the basic structure of λGEM-11, two features set this vector apart from the EMBL3 vector. Firstly the presence of Sfi I restriction sites outside the polylinker facillitates the high resolution mapping of insert DNA using an end labelling protocol, and secondly the opposing bacteriophage T7 and SP6 promotor sites allow the production of RNA probes for use in genome walking. Although neither of these features was required for this project, it was thought that they may be of use in future projects utilising the genomic library created as part of this project. It is also of note that the vector is able to replicate inserts between 9 and 23 kb, however inserts that vary from the optimum insert size replicate at lower efficiency.

The vector was purchased in a pretreated form, it had already been digested with

the restriction enzymes XbaI, EcoRI and BamHI, and had also been treated with a phosphotase (purchased from Promega Corp). The digestion with restriction enzymes serves to prevent religation of the vector arms to each other, and to allow the ligation of genomic DNA to the vector, since BamHI digested ends are compatible with the Sau3AI digested ends of the prepared genomic DNA. The phosphotase treatment removes the 5′ phosphate groups from the DNA terminii, this also effectively prevents the vector from ligating to itself, resulting in very low backgrounds of unaltered vector when the library is plated out.

Prior to ligation the vector DNA was heated to 42°C for one hour to anneal the λ cohesive ends, this helps to ensure that any given genomic DNA fragment is ligated to both arms, and furthermore that the ligation product is a concatenate, the preferred substrate for λ packaging. The genomic DNA and λGEM–11 arms were added to the ligation mixture in an approximate 1:1 molar ratio. Ligation was carried out at 18°C for 48 hours, using T4 Ligase (BRL) in the Ligase buffer (Appendix 6.4) with ATP at a concentration of 1mM. Three control tubes were also set up, one using a test insert supplied with the vector, this should ligate and be packaged at high efficiency and thus checks that the vector is functioning correctly. In a second control the vector arms are ligated to themselves, this tests the effectiveness of the previously described phosphotase treatment. The third control contained no ligase, only vector arms in ligase buffer, when packaged this indicates the level of background contributed by undigested vector. The exact constituents of the ligation reactions are given in Appendix 6.

The packaging reaction

The ligated vector and genomic DNA was packaged into phage particles using a Gigapack II Gold *in vitro* packaging kit [Stratagene Ltd.]. This packaging kit consists of extracts produced from two lots of cells infected with different λ strains defective in coat protein production, each extract is produced from a different mutant phage. Coat proteins accumulate within the cells which are then lysed to yield the extracts. When mixed together the two extracts provide all the necessary components to package concatameric λ genomes (including those carrying a foreign insert).

Besides the (putative) genomic library 5 controls were also packaged, these included the three controls described in the ligation section above and two packaging controls. One packaging control was a positive control, this was wild-type λ DNA (λcI857sam7) supplied with the packaging kit, which acts to test the efficiency of the packaging kit. The second control contained only the packaging extracts to check that these contributed no background.

Packaging of the genomic library was carried out according to the manufacturers instructions by adding the a 4 μl of the ligation reaction to one extract immediately it thawed, then immediately adding the second extract and gently mixing using a micropipette. The control DNA was treated differently, an aliquot of the DNA to be packaged was first placed in a 0.5 ml microfuge tube and then one fifth of a single packaging mix (as supplied) added to each tube. The two extracts were added quickly in succession, then mixed. Following mixing all reactions were microfuged briefly (5 second pulse) to remove bubbles, and then incubated at room temperature for 2 hours.

At the end of the incubation period several volumes of SM buffer were added (not containing gelatin, see Appendix 7.1), that is 500 µl to the genomic library and 100 µl to each of the control reactions. Chloroform was also added to each reaction, 20 µl to the Library reaction and 4 µl to each of the control reactions, thus effectively stopping the reactions and sterilising them. The assembled phage were stored at 4°C until used.

Titering the library

Bacteria for use in the culture of phage were streaked out from a glycerol stock on to LB agar, and incubated overnight at 37°C, following which it was stored at 4°C for several days. When bacteria were required for phage culture 50 ml of LB medium was inoculated with a single colony from the plate, and incubated at 37°C with shaking for 4 to 6 hours, the $A_{600\,nm}$ was not allowed to get higher than 1 cm^{-1}. The cells were then harvested at 2000 rpm, 4°C for 10 minutes in the 11133 swing-out rotor of a Sigma 3K12 centrifuge. The pellet was resuspended in a solution of 10mM $MgSO_4$ to give an $A_{600\,nm}$ of 2 cm^{-1}. It is necessary to use a young bacterial culture when plating the library because phage will also attach to dead cells and cellular debris, any phage which does this represents a lost clone, and lowers the titer of the library.

The *E. coli* strain KW251 was used for plating the library and all controls except the test DNA supplied with the packaging kit, this was plated on VCS257 [Promega Corp.]. The genotype of strain KW251 is given in Appendix 4.1. When plating 200 µl of KW251 prepared as described above was placed in a sterile glass test tube and phage diluent (from the packaging reactions, diluted in SM buffer, Appendix 7.1) added and mixed, the phage were then left to adhere to the bacteria at 37°C fro 15 minutes.

BBL agar plates (Appendix 7) were poured on a flat surface, and once set were dried until the surface became slightly wrinkly. These were then prewarmed at 37°C before use. When the adhesion period had finished the phage/bacteria mixture was added to 3-4 ml of BBL top agarose, mixed by swirling, and poured onto the prewarmed agar plates. Once set these plates were incubated at 37°C overnight.

After approximately 12 hours of growth plaques become visible on the plate, these clear patches in the agarose which is clouded by *E. coli* (this is termed a 'lawn' of bacteria) represent areas where phage have lysed the bacterial cells. Each plaque should represent a single clone, the plaques were counted and the efficiency of the packaging calculated as plaque forming units per ml (pfu ml^{-1}), for both control and library packaging reactions. If the efficiencies were close to those expected it was possible to move on to the screening of the library.

Primary screening of the library

For screening of the library 90 mm square petri dishes were used since this allows for easier access when performing plaque lifts (see next paragraph). LB agar was used as the bottom layer and BBL top agarose for the top layer, the plates were poured as described in the previous section. A volume of packaged genomic library that contained ~1000 pfu was plated on each plate, this number of plaques can be grown on a 90 mm plate without the plaques merging. Merging of plaques is to be avoided since it engenders the possibility of recombination between clones. Approximately

5000 plaques were screened in the primary screening of both strains of *Rv. gelatinosus*. Filters were also lifted from plated control packaging reactions to act as controls in the subsequent hybridisation.

Following overnight growth at 37°C, and when the plaques had reached a reasonable size but before they began to merge, the plates were removed from the incubator and placed at 4°C for 1 hour. This serves to slow phage growth, and also causes the agarose top layer to become somewhat more solid, such that it is less likely to be disturbed during the next step. Plaque lifts were performed according the protocol described by Kaiser & Murray, (1985), using Hybond-N nylon membrane [Amersham Ltd.]. Filters were lifted in duplicate from each plate, leaving the first filter in place for 30 seconds and the second for two minutes. DNA was fixed to the nylon membrane by exposure to UV for 2 minutes, once they had been dried and wrapped in cling-film.

The filters lifted from plates of the strain 151 library were probed using the α-consensus oligonucleotide as described in the sections "Preparation of nucleic acid probes" and "Nucleic acid hybridisation" above. These sections also describe the probe carrying the strain 151 pucBA genes which was used to screen the strain 149 genomic library.

Positively hybridising clones appear as spots on the autoradiogramme, the duplicate filters taken from each plate allow differentiation between background spots and truly hybridising plaques, since the latter always appear at the same point on both filters. The autoradiogramme can then be placed beneath the plates and corresponding plaques picked from the agar. Plaques were picked off as agar plugs as described by Kaiser & Murray, (1985), and placed in 1 ml of SM buffer to which a drop of chloroform was added (in order to kill any bacteria present). These picked plaques were stored at 4°C until the used in the secondary screening process.

Secondary screening

Prior to secondary screening each plaque solution was streaked across a lawn of bacteria using a sterile toothpick, after culture a well isolated plaque was picked for use in secondary screening. This process ensures that the screened plaques represent a single clone. For the purposes of rescreening the isolated plaques several 90 mm square plates were poured, using LB agar as the bottom layer and LB top agarose as the top layer. Plating bacteria were mixed with the top layer before pouring but no phage were added at this stage. Once the top layer had set 1 μl of each phage suspension resulting from the primary screening was spotted onto the agarose surface, in a grid pattern noting the position of each clone. The plate was then incubated overnight, a filter lifted the following day and this probed using the same probe as for the primary screening (as described above). Positively hybridising clones from the secondary screening were then cultured on a larger scale, DNA isolated for analysis and subcloning of a fragment carrying the target genes.

Culture of λ clones

The culture of isolated λ clones for the isolation of DNA was carried out in three stages, firstly each clone was purified by streaking a sealed Pasteur pipette which had

been dipped in the phage diluent, across a bacterial lawn. From this a well isolated plaque was chosen and picked off into SM as described above, this diluted plaque served as a stock culture from which the clone could be cultured when needed. The second stage involved the production of plate lysates for each clone, these were used to inoculate larger liquid cultures. The latter two stages of the protocol are described in Section 2.60 of the Handbook of molecular cloning (Sambrook *et al.*, 1989) and are outlined below.

Plate lysates were produced by mixing 100 µl of phage suspension with 200 µl of plating bacteria, and plating these in LB top agarose over a bottom layer of LB agar. Following overnight culture confluent lysis was generally observed. If necessary the process was repeated using extra phage solution to ensure confluent lysis occurred. The phage were harvested from this plate by pipetting 5 ml SM buffer onto the surface of the plate, and leaving for at least 1 hour at room temperature to allow the phage to diffuse out of the agarose. This 5 ml of SM containing phage was then removed from the plate using a micropipette, placed in 15 ml sterile falcon tubes and centrifuged for 10 minutes, 4°C, 2000g to remove cellular debris and agarose. The phage containing supernatant was then used to inoculate large scale cultures as described below, and in "Infection at high multiplicity" Section 2.72 of the Handbook of molecular cloning (Sambrook *et al.*, 1989).

A 2 ml aliquot from an overnight 50 ml culture of KW251 was used to inoculate 200 ml of LB in a 1 litre flask, which was then incubated at 37°C, 200 rpm in a rotary incubator. The A_{600} of the growing culture was monitored and at the point it reached an A_{600} of 0.4 cm^{-1} the plate lysate was added. After 4 or 5 further hours of growth the culture was usually observed to have undergone complete lysis, at this point an 400 ml of chloroform was added to lyse any remaining bacterial cells, and an excess of DNase and RNase added to digest away bacterial debris. Incubation was continued for a further 10 minutes, then the culture was transferred to 250 ml plastic bottles [Nalgene] and centrifuged at 10000 rpm, 4°C for 15 minutes in a J-14 rotor [Beckman] to further remove bacterial debris. The supernatant was then transferred to a second 250 ml bottle containing PEG6000 [Sigma] and NaCl to give a final concentration of 10% PEG and 1M NaCl. These ingredients were dissolved by Rocking and Rolling, and the bottles left on ice overnight to precipitate the phage. The phage were pelleted at 10000 rpm, 4°C for 20 minutes in a J-14 rotor [Beckman], and the supernatant discarded, the bottles were then rotated 180° in the rotor and centrifuged for a further 5 minutes, after which any remaining supernatant was removed with a pipette. The pellets were then resuspended in 10 ml of SM.

Phage were further purified by ceasium chloride density gradient centrifugation as described in Section 2.79 of the Handbook of molecular cloning (Sambrook *et al.*, 1989), with ceasium chloride at a concentration of 7.5g ml^{-1}. This centrifugation step was carried out in disposable 11.5 ml Ultracrimp polyallomer tubes [Sorvall], in a T865.1 rotor [Sorvall] at 46000 rpm, 21°C for 12 to 14 hours in a Sorvall OTD65B ultracentrifuge. The sample (usually visible as a white against a dark background) was removed by piercing the side of the tube with a needle and syringe and removing the band (the tube must first be pierced at the top to allow air to enter without disturbing the

gradient). These concentrated phage were stored in microfuge tubes at 4°C until DNA was isolated.

Prior to the isolation of DNA from purified phage the samples were dialysed against a buffer containing 50mM Tris-HCl (pH 8.0), 10mM NaCl and 10mM $MgCl_2$ to remove ceasium chloride. Dialysis was carried out against 3 changes of 5 litres of this buffer, with at least 1 hour between changes at room temperature or overnight at 4°C. Following dialysis the phage solution was transferred to sterile microfuge tubes and DNA isolated as described below.

To the phage solution EDTA was added to a concentration of 20mM, Proteinase K to a concentration of 50 $\mu g\ ml^{-1}$ and SDS to 0.5%. The resulting mixture was then incubated to 37°C for 1 hour to lyse the phage. After lysis the solution was cooled to room temperature and extracted twice against an equal volume of phenol/chloroform solution, then once against chloroform alone. The aqueous phase was then dialysed against 2 changes of 5 litres TE pH 8.0 for 1 hour at room temperature, and finally transferred to a sterile microfuge tube.

Restriction mapping

The purified genomic clones were mapped using the enzyme XhoI for both strains. The DNA (~250ng) was digested using XhoI [Gibco-BRL] according to the manufacturers instructions, and analysed on a 0.8% agarose gel. When a single copy of the probed sequence resides within the genome the fragments produced usually form a set of overlapping clones. Analysis of the pattern of fragments, looking for common fragments between clones, allows a restriction map of the area up to approximately 20 kb either side of the probe target sequence to be drawn up.

The resolution of maps produced in this way is dependent on the number of clones examined, this may lead to some uncertainties in the map. Further problems are also encountered when several sites occur within a few hundred base pairs, often it is impossible to distinguish their physical arrangement within the clone. However these low resolution maps are sufficient for the isolation of the target genes.

Subcloning into a sequencing vector

A λ clone carrying the target sequence in a central position was selected, and digested with the chosen restriction enzyme. For both strains of *Rv. gelatinosus* from which libraries were made the enzyme used was XhoI, since this produced fragments around 2 kb long, a size that is suited to further manipulation and subcloning into a plasmid vector. The digested DNA was 'shotgun' cloned into the XhoI site of the pBluescript SK- vector [Stratagene], by ligating the digested λ DNA to XhoI digested vector DNA as described in Appendix 9.1. A description of the pBluescript vector is given in Figure 17 (following pages). Prior to the ligation reaction the digested DNA was purified by phenol/chloroform extraction and ethanol precipitation, the pellets being resuspended in sterile analar water.

Figure 17

Description of the pBluescript vector, taken from the Promega Corp. catalog of 1993. The SK- version was utilised in this project.

pBluescript® SK (+/-) Phagemid

The pBluescript® SK (+/-) phagemid is a 2958 basepair phagemid derived from pUC19. The SK designation indicates the polylinker is oriented such that *lacZ* transcription proceeds from Sac I to Kpn I.

f1 (+) origin: f1 filamentous phage origin of replication allowing recovery of the sense strand of the *lacZ* gene when a host strain containing the pBluescript phagemid is co-infected with helper phage.

f1 (-) origin: f1 filamentous phage origin of replication allowing recovery of the antisense strand of the *lacZ* gene when a host strain containing the pBluescript phagemid is co-infected with helper phage.

ColE1 origin: Plasmid origin of replication used in the absence of helper phage.

Lac I: Certain sequences of the Lac I gene have been deleted. The Lac I gene is non-functional.

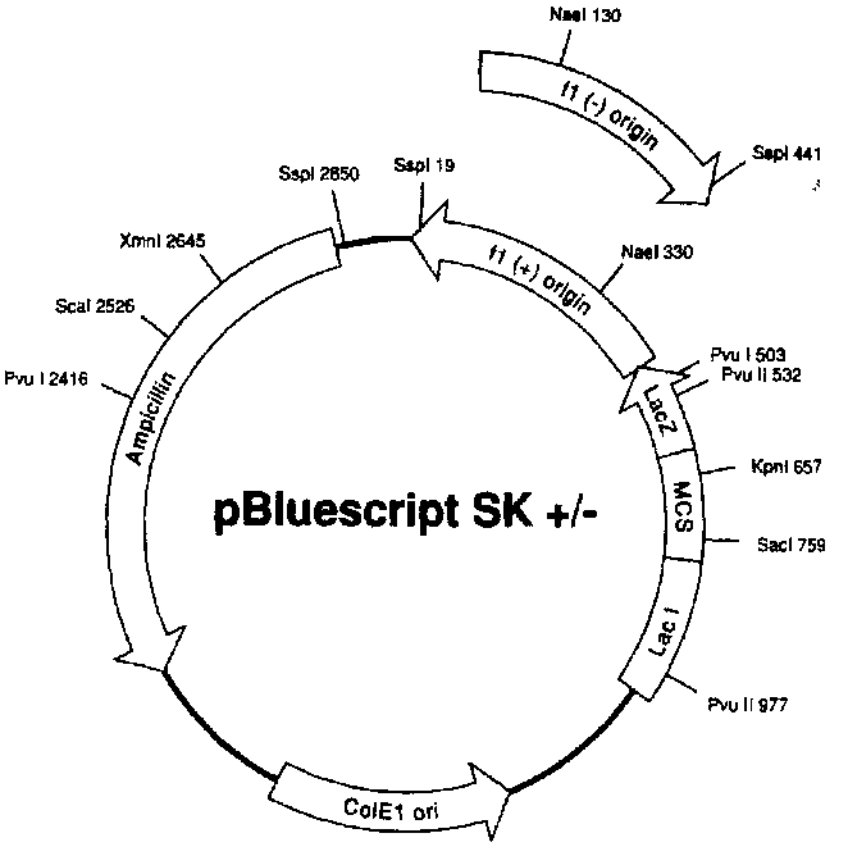
LacZ: This portion of the *lacZ* gene provides α-complementation for blue/white color selection of recombinant phagemids. An inducible lac promoter upstream from the *lacZ* gene permits fusion protein expression with the β-galactosidase gene product.

MCS: Multiple cloning site flanked by T3 and T7 promoters; please see sequence below.

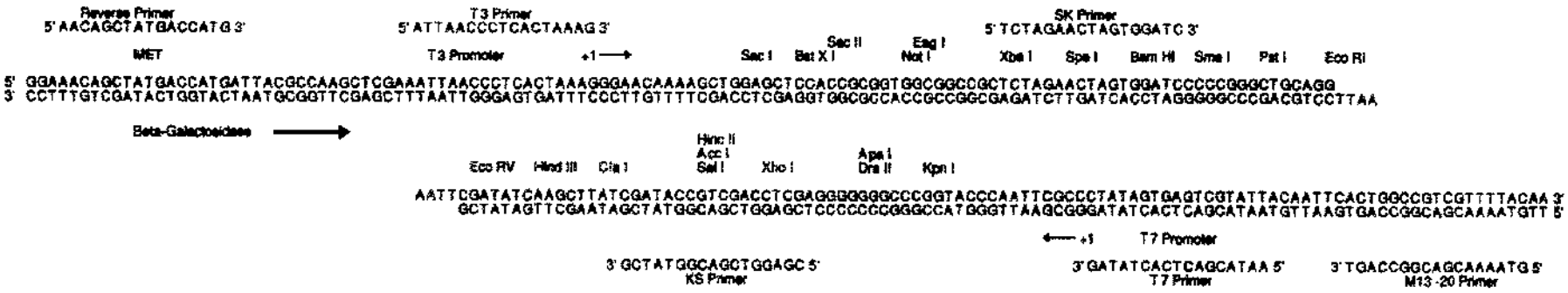
Ampicillin: Ampicillin resistance gene for antibiotic selection of the phagemid vector.

Please note: The upper strand is designated the (+) strand and the lower strand is designated the (-) strand.

GenBank® # 52325 (SK+), 52324 (SK-)



The pBluescript polylinker



Note: The fragments cloned in the course of this study were all cloned into the XhoI restriction site located at the T7 side of the polylinker. The T7 and T3 primer binding sites were used during sequencing of the cloned inserts.

The next step in subcloning was to transform the plasmids into the *E. coli* strain DH5α (Genotype is given in Appendix 4.2). Competent DH5α cells were prepared using two different methods, for strain 151 incubation in ice cold 0.1M CaCl was used to make the cells competent and transformation carried out by adding 2 μl of the ligation reaction to 200 μl of cells, as described in section 1.83 of the Handbook of molecular cloning (Sambrook *et al.*, 1989). For the subcloning of strain 149 puc genes the protocol of Chung *et al.*, (1989), was used. This latter protocol involves the use of DMSO to make the cells competent and is described in Appendix 10.

Transformed cells were plated onto LB plates containing 50 μg ml^{-1} of the antibiotic Ampicillin [Sigma], 80 μg ml^{-1} of fresh X-gal [Boehringer-Mannheim] and 20mM IPTG [Boehringer-Mannheim]. This allows the selection for bacteria containing plasmid, since pBluescript carries an ampicillin resistance gene, and also blue/white selection for the identification of plasmids carrying an insert.

Blue/white selection (via α-complimentation) is possible because DH5α has a deletion in the lacZ gene (lacZΔM15) which the pBluescript vector can complement, since it carries a segment of this gene. However the multiple cloning site (MCS) of pBluescript is located within the N-terminal end of this segment, thus when an insert is cloned into pBluescript the lacZ gene product (β-galactosidase) can no longer be produced. Thus addition to the media of IPTG, an inducer of lacZ expression, and X-gal, a substrate for β-galactosidase which is converted to a blue substance, causes colonies carrying unaltered pBluescript to turn blue, whilst those carrying a cloned insert remain white.

White colonies obtained from the plating of transformed DH5α were picked form the plates and cultured in LB (with 50 μg ml^{-1} added) as described in the section "Culture of E. coli" above, and small scale preparations of plasmid DNA isolated as described in "Plasmid isolation and purification" above. These samples were analysed by agarose gel electrophoresis and hybridisation with the probes used to screen the genomic library, and positive clones selected for large scale DNA preparation, analysis and sequencing.

Generation of nested sets of deletions

A further feature of the pBluescript vector is the positioning of sites within the MCS to facilitate the generation of unidirectional deletions of the inserted DNA. Sets of clones with progressively longer deletions were used in DNA sequencing, this is described in the section "DNA sequencing" below. The deletions were created using the enzyme Exonuclease III [Gibco-BRL], this has a useful property in that it digests away the 5′ strand of DNA molecules with blunt or 5′ overhanging terminii but not those with 3′ overhangs. The pBluescript MCS is arranged such that enzymes creating a 3′ overhang cut towards the outside of the MCS, whilst those creating blunt or 5′ overhangs cut more centrally (see Figure 17). Thus one can easily cut the plasmid in such a way that the vector is protected from Exonuclease III digestion, whilst the cloned insert is not. The original plasmid DNA was digested with Exonuclease III in a time course reaction, with aliquots being removed every thirty seconds into a stop solution. The resulting single stranded extensions were digested away with nuclease S1 [Gibco-

BRL], and a fill-in reaction then performed using Klenow fragment (of polymerase 1)[Gibco-BRL], this ensures that both terminii of the plasmid are blunt. The plasmids in each aliquot were then recircularised by blunt end ligation and transformed into DH5α using the protocol of Chung *et al.*, (1989) described in Appendix 10.

Three colonies were isolated from each time point, small scale preparations made and examined on an agarose gel. From these a set of clones differing in size by approximately 200 bp were selected, and these cultured using the large scale protocol described in "Plasmid isolation and purification" above. The purified DNA was then used for DNA sequencing as described below.

A full description of the protocol used can be found in Gardiner, (1992), whilst general advice and alternative protocols are given in Section 13.34 of the Handbook of molecular cloning (Sambrook *et al.*, 1989).

DNA sequencing

The strategy employed to sequence the cloned DNA was to make nested deletions where possible, sequencing from one of the universal primer sites in pBluescript. This allowed the 200 to 300 bp sequenced from each deletion clone to be assembled into a contig of one strand, custom primers were then synthesised and used to sequence the second strand. The design and synthesis of primers is described in the section "Oligonucleotide design and preparation" above. Where it was not possible to create deletions sequencing was achieved by primer walking, on both strands of the DNA.

DNA for sequencing was purified by equilibrium centrifugation in ceasium chloride gradients as described above, this provides highly pure supercoiled plasmid DNA suitable for sequencing. Plasmid DNA was prepared for sequencing using two different methods, in the main the DNA was denatured by alkali and used directly in the sequencing reactions. Alkali denaturation utilising NaOH was carried out according to the method described in the manual of the USB Sequenase Kit used for sequencing (see below), modified from the methods of Chen & Seeburg, (1985) and Hattori & Sakaki, (1986), and also described in Section 13.71 of the Handbook of molecular cloning (Sambrook *et al.*, 1989).

Some templates were first made single stranded using T7 gene 6 exonuclease [USB corp.], this exonuclease digests double stranded DNA in the 5′-3′ direction from terminii that blunt, or that have 5′ or 3′ single stranded overhangs. The products of the digestion of linear double stranded DNA with T7 gene 6 exonuclease are single stranded half molecules, this single stranded product may be used directly in sequencing reactions.

Templates to be made single stranded were treated with the enzyme PvuII [Gibco-BRL], pBluescript has two PvuII sites lying either side of the MCS, digestion of pBluescript alone with PvuII yields a fragment 445 bp in size (see Figure 17). If the plasmid is carrying cloned DNA inserted into the MCS then the result of the PvuII digest is to cut out the cloned insert plus the small amount of vector DNA (assuming that there are no PvuII sites within the insert). T7 gene 6 exonuclease treatment then yields two single stranded half molecules which stop approximately midway through

the insert DNA, allowing the use of universal primer sites in pBluescript and of correctly designed custom primers.

Since both PvuII and T7 gene 6 exonuclease are able to function in the medium salt Sequenase buffer [USB] used for DNA sequencing (see below), it was possible to carry out these digestion steps on the DNA and then perform the sequencing reaction in the same microfuge tube. The protocol used is described in Appendix II.

Sequencing was carried out using Sequenase 2.0 DNA sequencing Kits [USB Corp.] and TAQuence 2.0 DNA sequencing Kits [USB Corp.], according to the manufacturers instructions. Both these kits utilise the Sanger dideoxy sequencing protocol (Sanger *et al.*, 1977), in which a synthetic oligonucleotide bound to a single stranded DNA template is used to prime synthesis of a new strand by a DNA polymerase. The polymerase is capable of incorporating radiolabelled nucleotides, allowing the new strand to be labelled, and can also incorporate dideoxynucleotides which terminate the growing chain. The two steps, labelling and termination are usually performed in two separate sequential reactions. Addition of the radiolabelled and dideoxynucleotides in the correct proportion to standard deoxynucleotides, will yield an array of labelled transcripts of different sizes. The radiolabelled fragments can then be size separated on a polyacrylamide gel and the DNA sequence read. The Sequenase 2.0 kit utilises a modified form of T7 DNA polymerase, described by Tabor and Richardson, (1989).

The Sequenase kit was used as a standard sequencing system, it was found however that pausing by the polymerase (at secondary structure in the template DNA) and gel compressions (secondary structure in the synthesised transcript), were most easily overcome by the use of Taq DNA polymerase and the 7-deaza-dGTP nucleotide analogue as supplied in the TAQuence kit. The latter analogue forms weaker secondary structures, thus allowing the elimination of gel compressions, whilst the use of TAQ polymerase allows sequencing reactions to be carried out at temperatures up to 70°C, effectively destabilising secondary structure within the template and preventing pausing.

Further modifications to the original Sanger sequencing protocol suggested by USB Corp. were also used. These included the addition of pyrophosphotase to the sequencing enzyme mix, to improve occasional weak band intensity problems, and the use of Glycerol enzyme dilution buffer which in turn forced the use of glycerol tolerant sequencing gels. The dilution of Sequenase in 50% glycerol has two advantages, firstly it allows the dilution of enzyme to a working concentration well in advance, simplifying the sequencing procedure. Secondly it is a way of adding extra glycerol to the sequencing reaction, this improves the stability of the enzyme and allows higher incubation temperatures and longer incubations during the labelling and termination steps.

One problem associated with the extra glycerol added in the dilution buffer and also by the use of T7 gene 6 exonuclease produced single stranded DNA directly in the sequencing reaction, is that the glycerol interacts with boric acid in the TBE buffer commonly used for DNA sequencing gels. This interaction causes distortion of the DNA sequence at the top of the gel, but can easily be overcome by the use of TTE

buffer, which replaces boric acid with taurine [Sigma] but is otherwise the same. A recipe for TTE buffer can be found in Appendix 12.1.

Sequencing gels were made at a concentration of 5 to 8% Acrylamide depending on the sequencing range (distance from the primer) desired, and also contained 50% urea as a denaturing agent. The stock acrylamide solution contained 40% acrylamide at a ratio of 19:1 of acrylamide to bisacrylamide [BDH Ltd., Electran grade]. This was used to make a 100ml gel mix from fresh, 30 µl of TEMED and 1ml 10% ammonium persulphate added to initiate polymerisation. The gels were poured between two glass plates (34 by 45 cm) using strips of 3MM paper as spacers and allowed to set for a minimum of two hours before use. Sharkstooth combs were used [Gibco-BRL], and the samples heated to 75-80°C for 10 minutes before loading onto gels which had been prewarmed to ~55°C.

The gels were run in vertical kits designed and built in house. Either a Bromma A2197 [LKB] or a Biochrom 2103 [LKB] power supply was used to run the sequencing gels at the wattage required to maintain a gel temperature of 55°C. Heat was dissipated from the gel and spread more evenly over it by the use of a 0.5 cm thick aluminium plate attached to one of the glass plates. The length of time that sequencing gels were run was dependent on the distance one was trying to sequence from the primer, but varied between 3 and 14 hours. Following the run the gels were fixed by immersion in 5 litres of 10% acetic acid/10% methanol in water, for 1 hour. They were then transferred to a sheet of 3MM paper and vacuum dried using a Model 483 slab dryer [BioRad], drying was carried out under vacuum at 80°C for 30 minutes or 60°C for 1.5 hours. Once dry the gels were autoradiographed as described in the section "Autoradiography" above.

Sequence analysis

The University of Wisconsin Genetics Computer Group [GCG Inc.] software package was used extensively in the analysis of DNA and protein sequence. The names of individual programs within this package are given in the results sections when used, as are the names of any other programs and packages used.

PCR

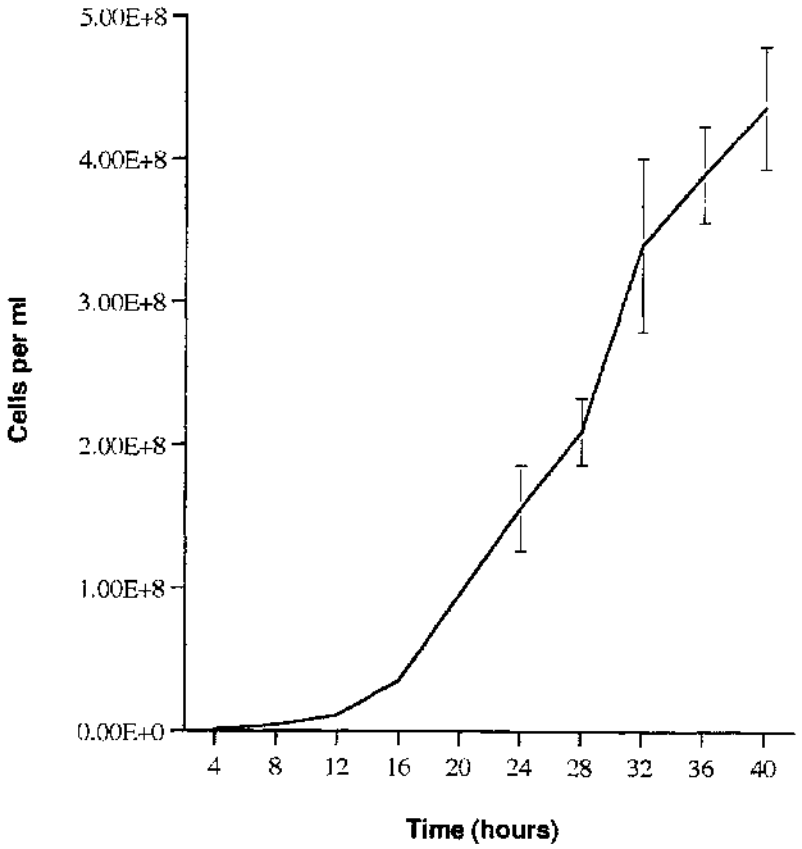
The polymerase chain reaction (PCR) is now used widely in a range of applications, it was used in this project as a means of producing DNA fragments to use as probes in hybridisation studies. It was also used to produce a mutated fragment for use in the heterologous expression studies detailed in the results section, the mutations however were outside the reading frame of any genes, and merely introduced new restriction enzyme sites needed for subcloning the genes into a new vector.

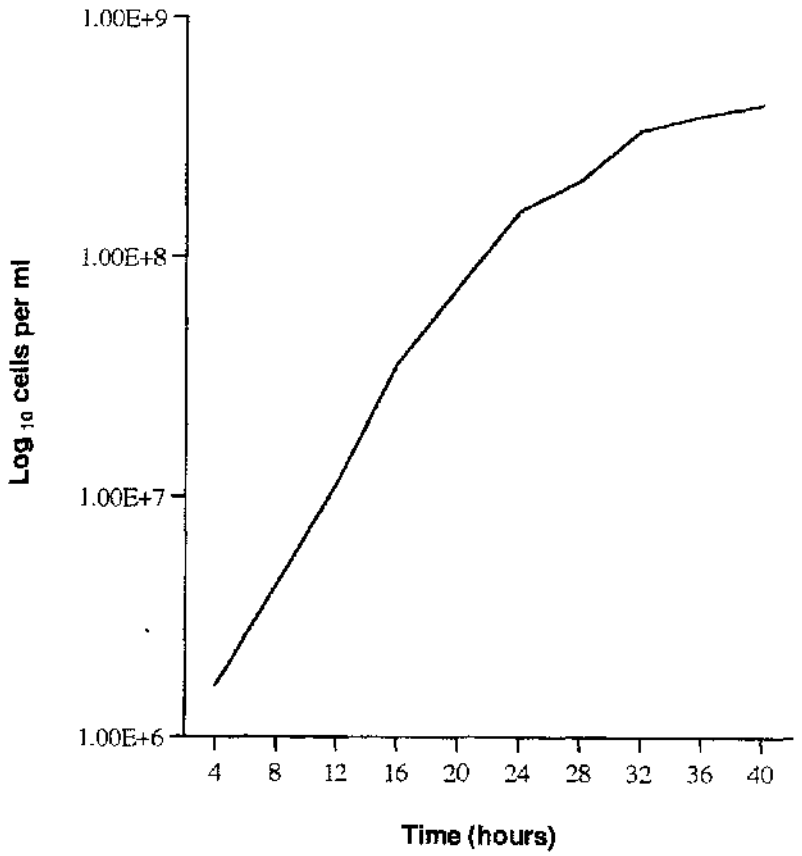
Taq DNA polymerase [Promega Corp.] was used with the buffer supplied, also added to the reaction were dNTPs [Pharmacia] to a final 200µM each, $MgCl_2$ to 1.5mM, and primers to 0.5µM. The template DNA was added at a concentration that gave approximately 10^4 copies of the target sequence. PCR was carried out using a Crocodile II Incubation System [Appligene], the cycling parameters were adjusted to suit the primers being used, but generally involved an initial 4 minute denaturing melt

at 95°C, then cyclical 1 minute anneal/extend/denature steps at ~60°C/72°C/95°C. It was the annealing temperature that was changed to suit the primer. The number of cycles used was usually about 30, exact parameters are given for each reaction in the results sections.

Figure 18

Two graphs showing the number of cells against time during the growth of strain 149. The graph on the left has a linear y axis, whilst the graph on the right has a logarithmic y axis. The vertical bars represent 95% confidence limits. It should be noted that no 20 hour sample was collected, and thus has not been counted or plotted.





Results Part 1: Comparative Physiology of the *Rv. gelatinosus* Strains 151 & 149

Introduction

As stated in the thesis introduction, it had been noted that the two strains of *Rv. gelatinosus* used in this study appeared to synthesise different amounts of LH2 complex when cultured under identical conditions. The work detailed in this section was carried out in order to confirm this observation, and also to define the culture conditions that would be useful for the gene expression studies described in the "Results part 5" section. Initially cell counts were carried out on samples from an experiment monitoring the growth of strain 149 cultures (the samples were collected by Dr. Andrew Gall as detailed in Gall, (1994), and the cells counted by myself). These cell counts ensured that absorption at 650 nm was a reliable diagnostic for cell density, a fact that was important to check because absorption at 650 nm is attributed to the CM, and production of ICM could thus potentially interfere with measurement of cell density. Examination of the cell count and growth curve data also indicated a useful culture harvesting time for the experiments which followed. These included the culture of strains 151 and 149 under three different light regimes followed by measurement of their whole cell absorption spectra, to examine their response to irradiance level, and the gene expression analysis detailed in "Results part 5".

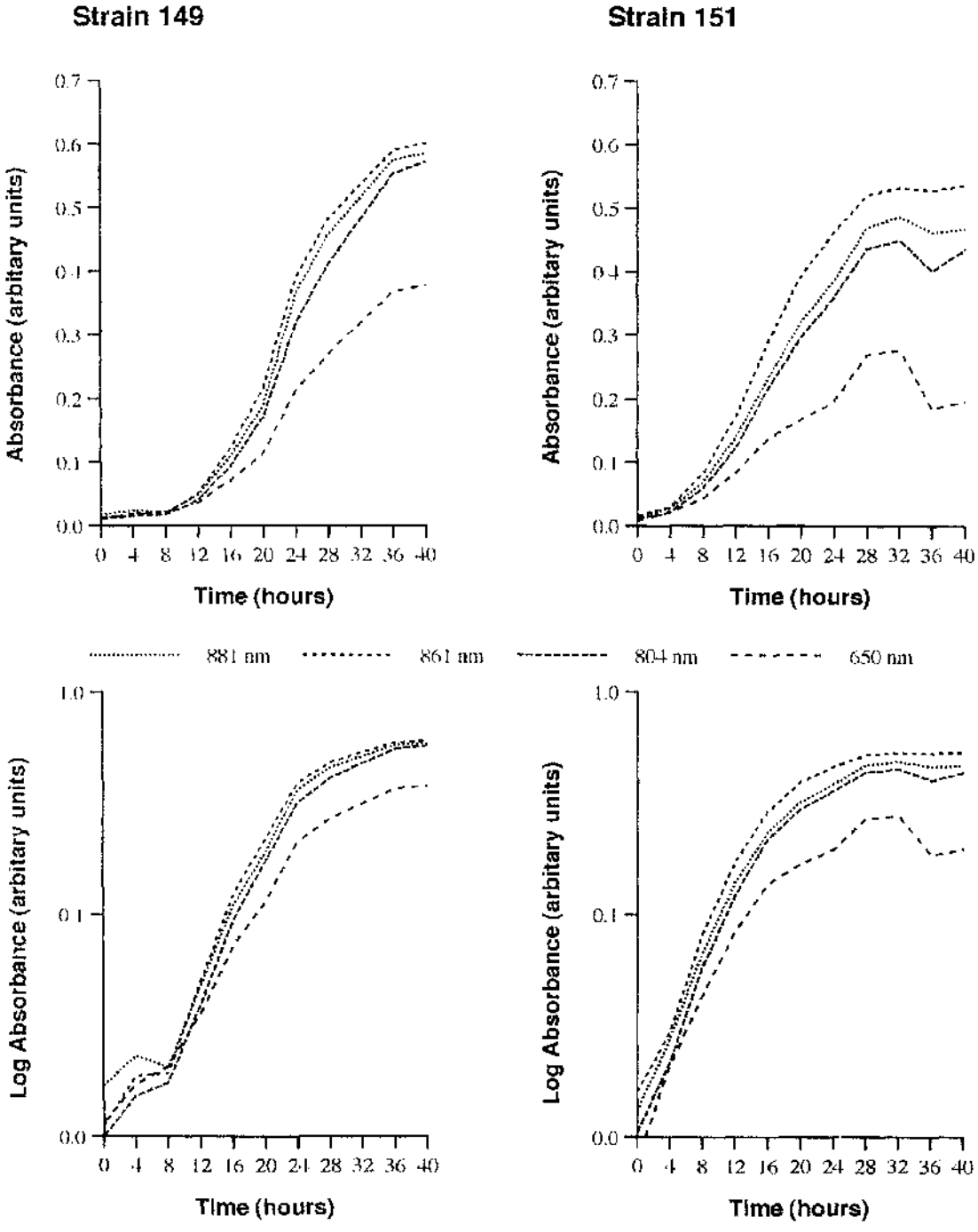
Results

Growth of *Rv. gelatinosus*

Samples from strain 149 taken every 4 hours during the course of growth were supplied by Dr Gall, and the cells counted by myself as described in the Materials and Methods section "Cell Counts". The resulting graph of cell number against time is illustrated in Figure 18 (opposite), a second graph with the same data plotted on a logarithmic y axis is also given. The growth curve is sigmoidal in much the same way as a graph showing the way A_{650nm} (representative of cell density) changes during growth as can be seen in Figure 19 (overleaf). The absorption peaks at 804, 861 and 881 nm representing LH2 and LH1 complexes are also plotted in Figure 19. The peak at 861 nm is predominantly attributable to absorption by the dimeric Bchl in LH2 whilst that at 881 nm is attributed to the dimeric Bchl molecules of LH1, thus these two peaks are clearly representative of LH2 and LH1 (see the introduction section "Light-harvesting complexes" for a more detailed discussion). Interpretation of the 804 nm peak is somewhat more complicated, since both complexes contribute to absorption at this wavelength, however the size of this peak relative to those at 881 and 861 nm indicates which LHC is dominating the NIR absorption spectrum.

Figure 19

Whole cell absorption at selected wavelengths during growth of strains 149 and 151 of *Rv. gelatinosus*. Absorbance values at a series of wavelengths were measured during growth of cultures and plotted together on the graphs below. The wavelengths chosen include the NIR absorbance peaks of LH1 and LH2, as well as the absorbance at 650 nm which provides a measure of cell density. For comparison with the graph of strain 149 cell density in Figure 18 it should be noted that the A_{650nm} of the 20 hour sample (0.1129 cm^{-1}) is approximately equivalent to $1x10^8$ cells ml^{-1}. The upper graphs have a linear y axis whilst the lower graphs have logarithmic y axes. The graphs on the right displaying data collected from strain 151 clearly illustrate the problems associated with recording absorbance spectra from this strain. Beyond a culture time of 32 hours the absorbance values appear to drop, this is most likely to result from aggregation of cells in the mucus produced by the bacteria. The main text provides a more expansive discussion of this problem, but it is clear that data from beyond the 32 hour point should be disregarded, and those before it treated with caution.



The similarity between the A 650 nm and the cell number indicates that ICM has no or a negligible affect on absorbance at 650 nm, and that A_{650nm} is a reliable measure of *Rv. gelatinosus* cell density.

A potential problem with the culture of strain 151 is that mucus production by this strain appears to cause cell aggregation, and makes the measurement of accurate absorption spectra difficult. This has an effect visible in Figure 19 (opposite), in that samples taken as the culture reached stationary phase appear to vary unpredictably in absorption. To avoid this problem it was decided to harvest cells of both strains in early log phase, this has a second advantage in that it should ensure there is little self shading of cells during culture which could alter the perceived irradiance level.

Culture of *Rv. gelatinosus* at different irradiance levels

In order to analyse the response of *Rv. gelatinosus* to irradiance levels the two strains were cultured under a range of different lighting conditions, as described in the Materials and Methods section "Anaerobic liquid culture". Initially 5 different irradiance levels were used, however repeated attempts to culture the bacteria at the two lowest irradiance levels failed, and the experiment was continued using only the highest 3 levels. These were measured to be 100, 40 and 18 W m^{-2}, and will be referred to as high-light (HL), medium-light (ML) and low-light (LL) to ease discussion.

An initial culture of each strain was started under a HL irradiance level before splitting culture into several bottles using a 5% inoculum and growing at the three different irradiance levels. The HL and ML cultures were sub-cultured 3 times using an inoculum of 5%, this was to ensure full adaption to the growth conditions. Absorption spectra were recorded from the fourth culture as described below. The initial subculture from the nominal HL starting culture into the LL environment failed to grow, thus a larger inoculum (20%) from the third ML culture of each strain was used to initiate the low light cultures. As with the HL and ML cultures the LL grown bacteria were sub-cultured three times before the spectra was recorded, however a 10% inoculum was used at each transfer. In all cases sub-culturing was carried out at a point before self shading of the bacterial cells would become significant and thus alter the irradiance level. This was judged by eye, but appears to have been successful since the A_{650nm} of the harvested cultures varied between 0.012 cm^{-1} (strain 149 LL) and 0.046 cm^{-1} (strain 149 HL), thus no culture was grown over an A_{650nm} of 0.05 cm^{-1}. The period taken for the cultures to reach this density was approximately 24 hours, the HL cultures growing slightly faster, and the LL cultures slower.

Whole cell absorption spectra were recorded on a Shimadzu UV-2101PC scanning spectrophotometer as described in the Materials and methods section "Absorption spectroscopy". Spectra from two separate aliquots of each culture were recorded and the mean calculated, this mean spectra was then used in any further data analysis and manipulation.

The results of this experiment are presented in a series of graphs on the following pages. The spectra presented were smoothed using a 7 point moving average, this removed noise resulting from the low cell densities being measured (all samples had an A_{650nm} of < 0.05 cm^{-1} at harvest). The spectra were then normalised to an absorbance at 650 nm of 1 cm^{-1} to allow direct comparison of the LHC levels, only the region from 750 to 900 nm is displayed.

Figures 20 and 21 (overleaf) illustrate the response of each strain to culture at different irradiance levels. The difference between the two strains is clearly visible in Figure 20, which compares the spectra of the two strains at each irradiance level. The most obvious difference is that the strain 151 LH2 peak is much larger than that of strain 149 at each irradiance level, suggesting that more LH2 is present in the strain 151 samples. This is further supported by the differences in the shape of the NIR absorbance peaks, strain 149 has relatively a flat peak in the 830-900 nm region since the LH1 (881 nm) and LH2 (861 nm) peaks merge, in contrast the strain 151 LH2 absorption dominates that of its own LH1 producing a much sharper looking NIR peak.

Figure 20

Normalised absorption spectra of *Rv. gelatinosus* strains 151 and 149 cultured at 3 different irradiance levels. Each panel represents a different light level (given above the spectra), and shows smoothed normalised absorption spectra from each strain (see legend). The main text has details of the how the original spectra were recorded and the data manipulated. The y axis represents absorbance in arbitary units and has been omitted for clarity. The spectra are all normalised to an absorbance at 650 nm of 1 per cm, and the y axes for each graph have been scaled to the same size, thus all spectra illustrated are directly comparable.

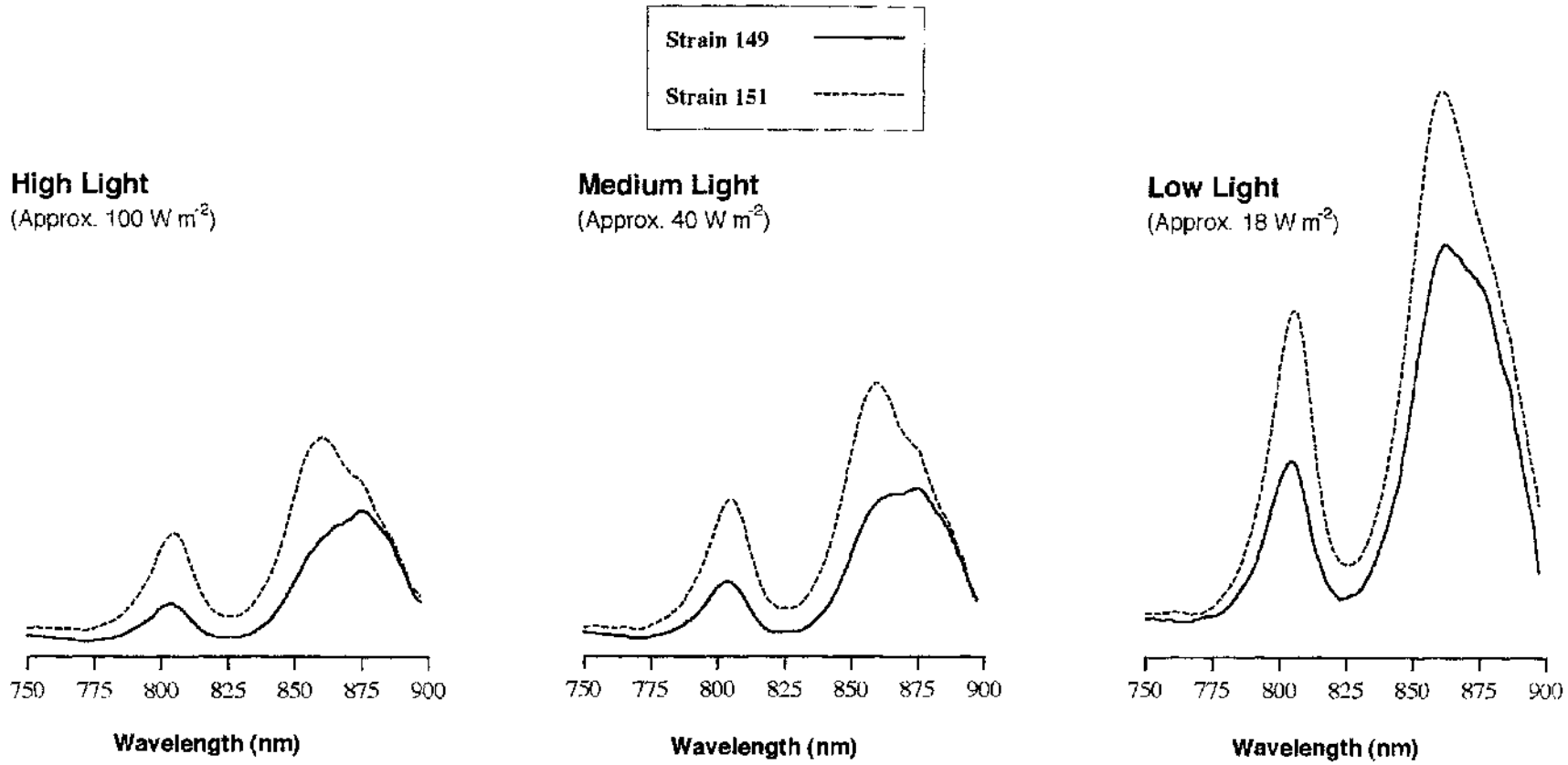
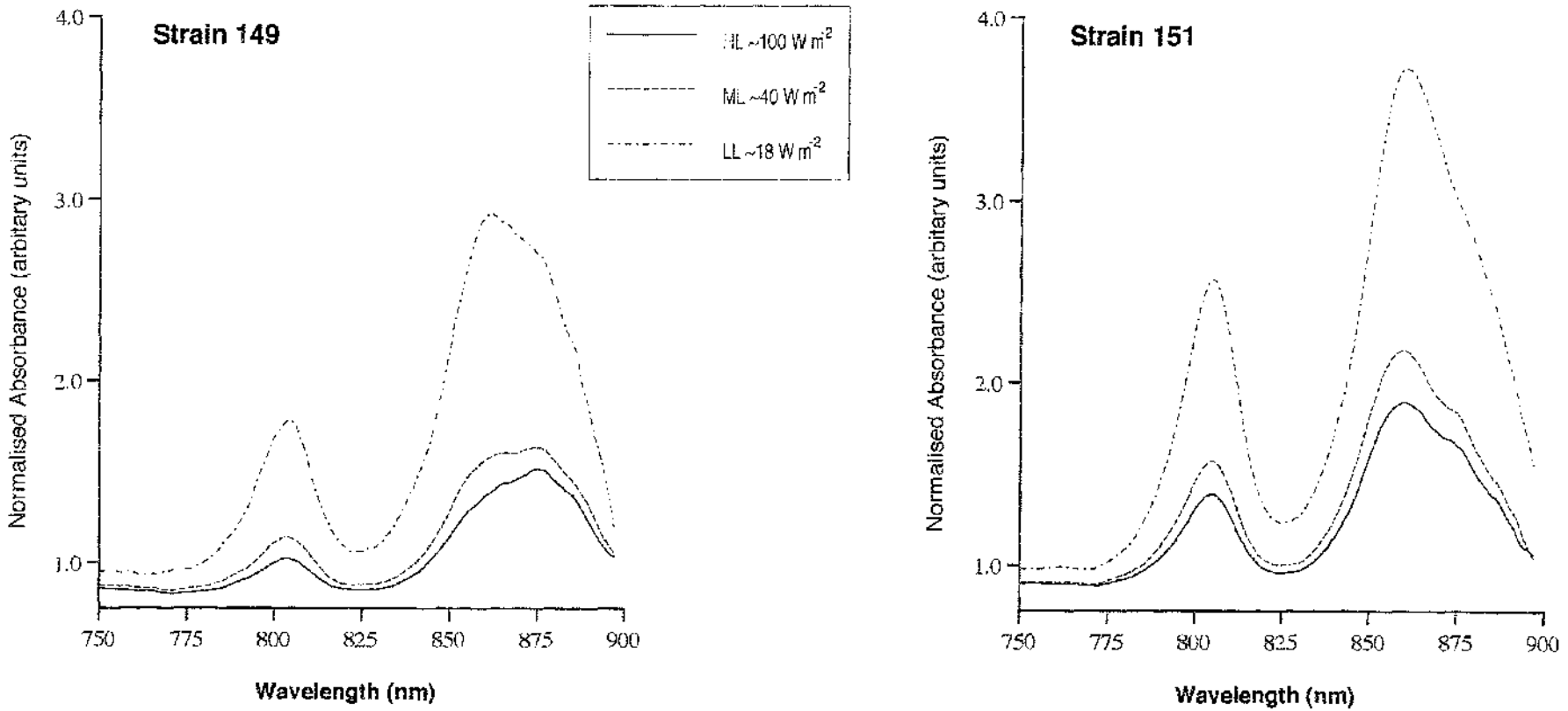


Figure 21

Normalised absorption spectra of *Rv. gelatinosus* strains 151 and 149 cultured at three different irradiance levels. The spectra on the left were recorded from strain 149 cultured at high light (HL), medium light (ML), and low light (LL). The spectra on the right were recorded from cultures of strain 151 grown under the same conditions. All spectra were normalised to an absorbance at 650 nm of 1 per cm, details of how the spectra were recorded and manipulated can be found in the main text.



This also indicates that strain 151 has more LH2 relative to the amount of LH1 than strain 149, the level of LH1 apparently remaining relatively static in the face of varying irradiance.

The larger 804 nm peak of strain 151 also indicates that more LH2 is present than in strain 149. LH1 contributes to the absorption at 804 nm but at a much lower level than LH2 (see Figure 5 in the Introduction for an example of an LH1 absorption spectrum). Thus when LH1 dominates NIR absorption the 804 nm peak is proportionately smaller in comparison to the 861 and 881 nm peaks, this can be seen to be the case for strain 149 in Figure 21, only under LL does the 804 nm peak show a significant increase in comparison to that at 861 nm. In contrast the strain 151 804 nm peak is 50% the size of the 861 nm peak under HL, and grows even larger as the irradiance level is lowered.

Another interesting comparison concerns the ratio of HL:ML:LL A_{861nm} in the two strains. In strain 149 a ratio of 1 : 1·1 : 2·1 is found, the comparable ratio for strain 151 calculated relative to the strain 149 HL A_{861nm} reading, is 1·4 : 1·6 : 2.7. Similarly comparing HL:HL, ML:ML and LL:LL A_{861nm} between the two strains one finds ratios of 1 : 1·4, 1 : 1·4 and 1 : 1·3, suggesting that strain 151 synthesises approximately one third more LH2 than strain 149 at any given irradiance level (assuming the amount of LH2 present is proportional to the LH2 absorbance peak).

Despite these obvious differences it is clear from the spectra shown in Figure 21 that *Rv. gelatinosus* responds to irradiance level in a manner similar to other photosynthetic bacteria, HL suppressing the formation of LH2. This is evident in the size of the absorbance peaks at 804 nm and 861 nm representing LH2, which are larger under lower irradiance levels. Both strains apparently respond in a similar manner as can be seen if one looks at the variation in A_{861nm} between the different lighting conditions. The ratio of A_{861nm} values for HL:ML:LL in strain 149 is 1 : 1·1 : 2·1, an extremely similar variation is seen in strain 151 which has a ratio of 1 : 1·1 : 2·0 at the same wavelength. Thus the amount of variation which can be induced by changing the irradiance levels appears very similar in the two strains.

Conclusions

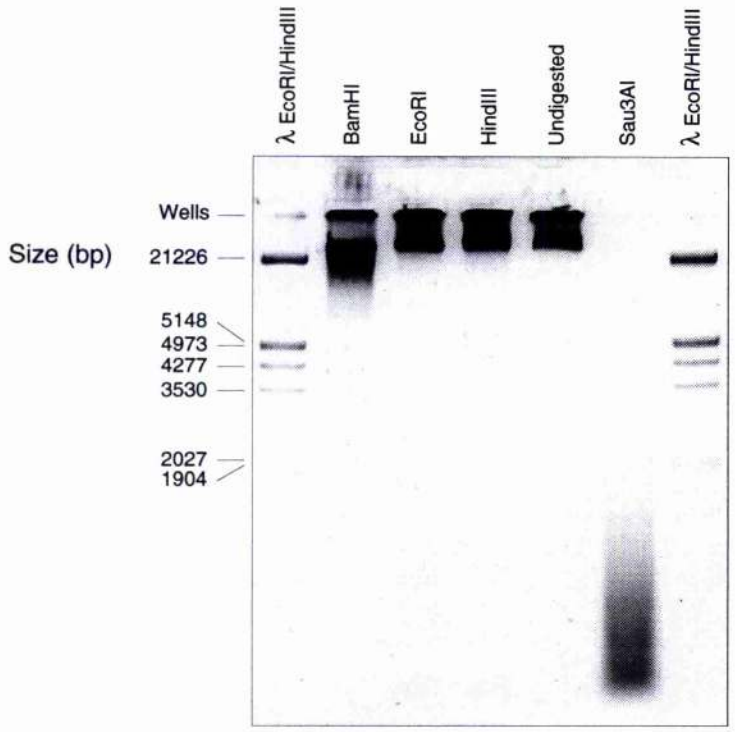
The culture of the two *Rv. gelatinosus* strains at three different irradiance levels proved an effective method for highlighting the differences between them. From the results obtained it appears that strain 151 synthesises approximately one third more LH2 than strain 149 when cultured under the same conditions. This confirms the initial observation that strain 151 synthesises more LH2 than strain 149. However it is interesting to note that the degree of variation in LH2 synthesised appears similar in both strains, the LL irradiance level inducing an approximate two fold increase over the amount of LH2 present at HL, and the ML an approximate 10% increase in comparison to HL.

The data relating to the cell counts and growth curves suggest that the growth of *Rv. gelatinosus* follows a sigmoidal pattern that can be followed by measurement of cell culture absorption at 650 nm. The use of the three defined irradiance levels combined with cell harvest at early log phase proved sufficient for the observation of changes in

LH2 levels, and thus provide a useful protocol for the analysis of LH2 synthesis.

Figure 22

An agarose gel illustrating how genomic DNA from *Rv. gelatinosus* strain 151 is resistant to cleavage by several restriction enzymes.



Note: Samples (20 μg) of strain 151 genomic DNA were digested with the restriction enzymes BamHI, EcoRI, HindIII and Sau3AI and run on a 0.3% agarose gel. The labels above the gel indicate which enzyme was used. A sample of undigested genomic DNA was also loaded (fifth from the left) for comparison. To the left and right sides of the gel are DNA markers, these were made by digesting phage λ DNA with EcoRI and HindIII enzymes, the size of the resulting fragments is given at the side of the gel (250 ng of markers were loaded).

Results Part 2: Construction of Genomic Libraries & Cloning of the puc Genes

Introduction

This section details the construction of genomic libraries for the two strains of *Rv. gelatinosus*, plus the subsequent screening, selection and subcloning procedures. The library of strain 151 was created first, certain parameters optimised during this process were then applied to the construction of the strain 149 library.

The process of library construction, screening and subcloning is a lengthy one involving several separate but successive stages. These stages are described in some detail in the Materials and Methods, the results of each step are given below together with any conclusions which were essential for the subsequent steps, more general conclusions are discussed at the end of the section.

Results

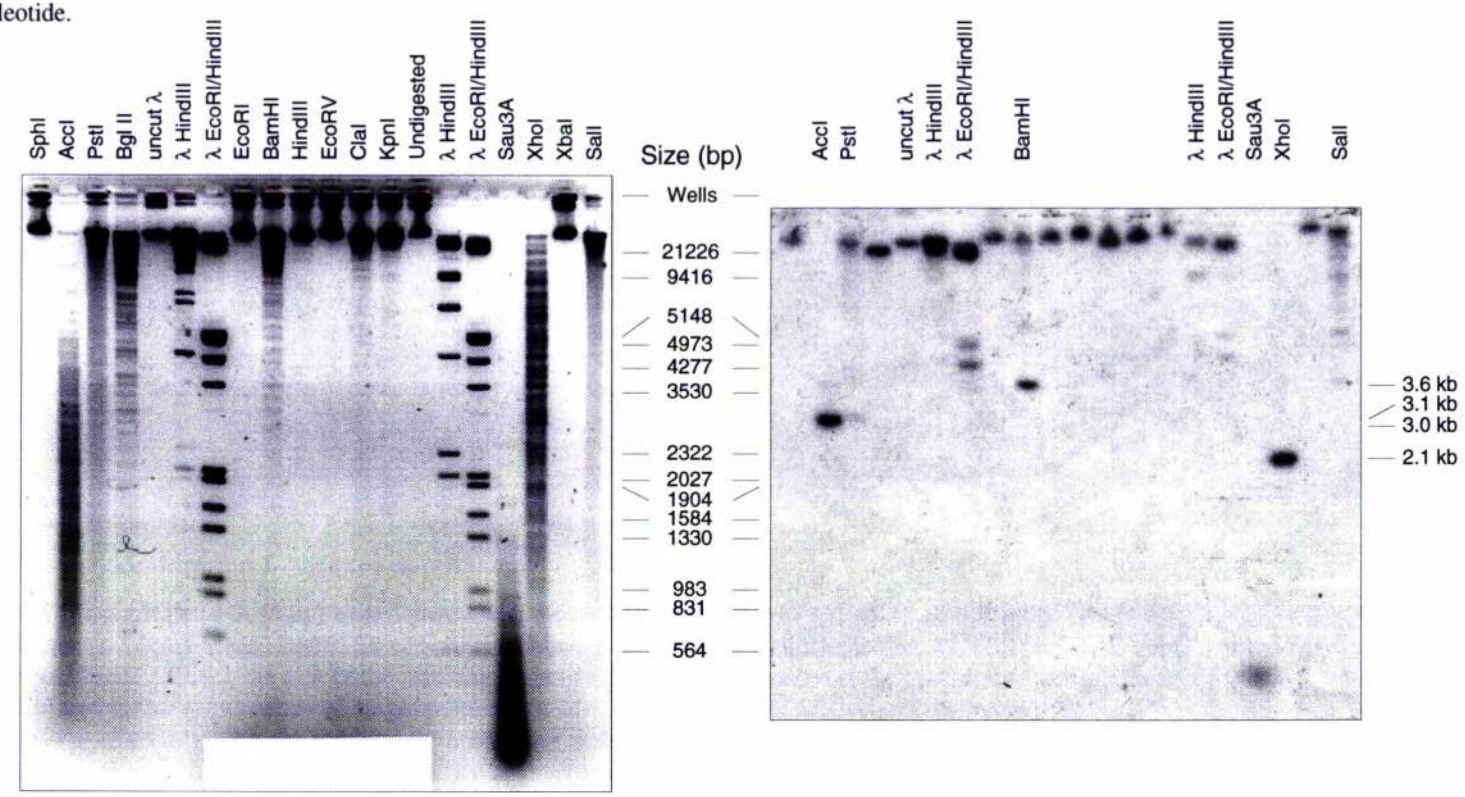
Test restriction digests of genomic DNA

Genomic DNA from *Rv. gelatinosus* strain 151 isolated as described in the Materials & Methods section was digested with the restriction enzymes BamHI, EcoRI, HindIII and Sau3AI. This served two functions, firstly it confirmed the purity and size of the DNA and secondly it showed that those enzymes to be used in the library construction were able to digest the DNA. Figure 22 (opposite) shows the resulting digests after they had been run on a 0.3% agarose gel, interestingly the DNA was not cut by EcoRI or HindIII, since the DNA in these lanes appears identical to the undigested genomic DNA. BamHI caused a minor change in size and thus must have digested the DNA to a limited extent, but did not cut extensively. Much of the DNA remains uncut by these three enzymes and is trapped in the wells at the top of the gel, this is not the case with DNA digested using Sau3AI, which cut the DNA into small fragments (visible as a smear at the bottom of the gel).

To investigate further the inability of commonly used restriction enzymes to digest the *Rv. gelatinosus* strain 151 genomic DNA, samples were digested with a wider range of enzymes and run on a 0.8% agarose gel. The results are illustrated in Figure 23 (overleaf), several enzymes cut the DNA as can be seen by the smearing and bands present in lanes labelled AccI, PstI, BglII, XhoI and SalI. However many enzymes either cut to a limited extent or didn't cut at all. Since many restriction enzymes are sensitive to contaminants in the DNA (such as metal ions) it was decided to further purify an aliquot of genomic DNA by ethanol precipitation. Digestion with the same enzymes then produced fairly similar results, except for the enzymes PstI, EcoRI, EcoRV and SalI, which had somewhat improved digestion.

Figure 23

Agarose gel analysis of the digestion of *Rv. gelatinosus* strain 151 genomic DNA with a range of restriction enzymes, and autoradiograph of a subsequent Southern blot probed with the α consensus oligonucleotide.



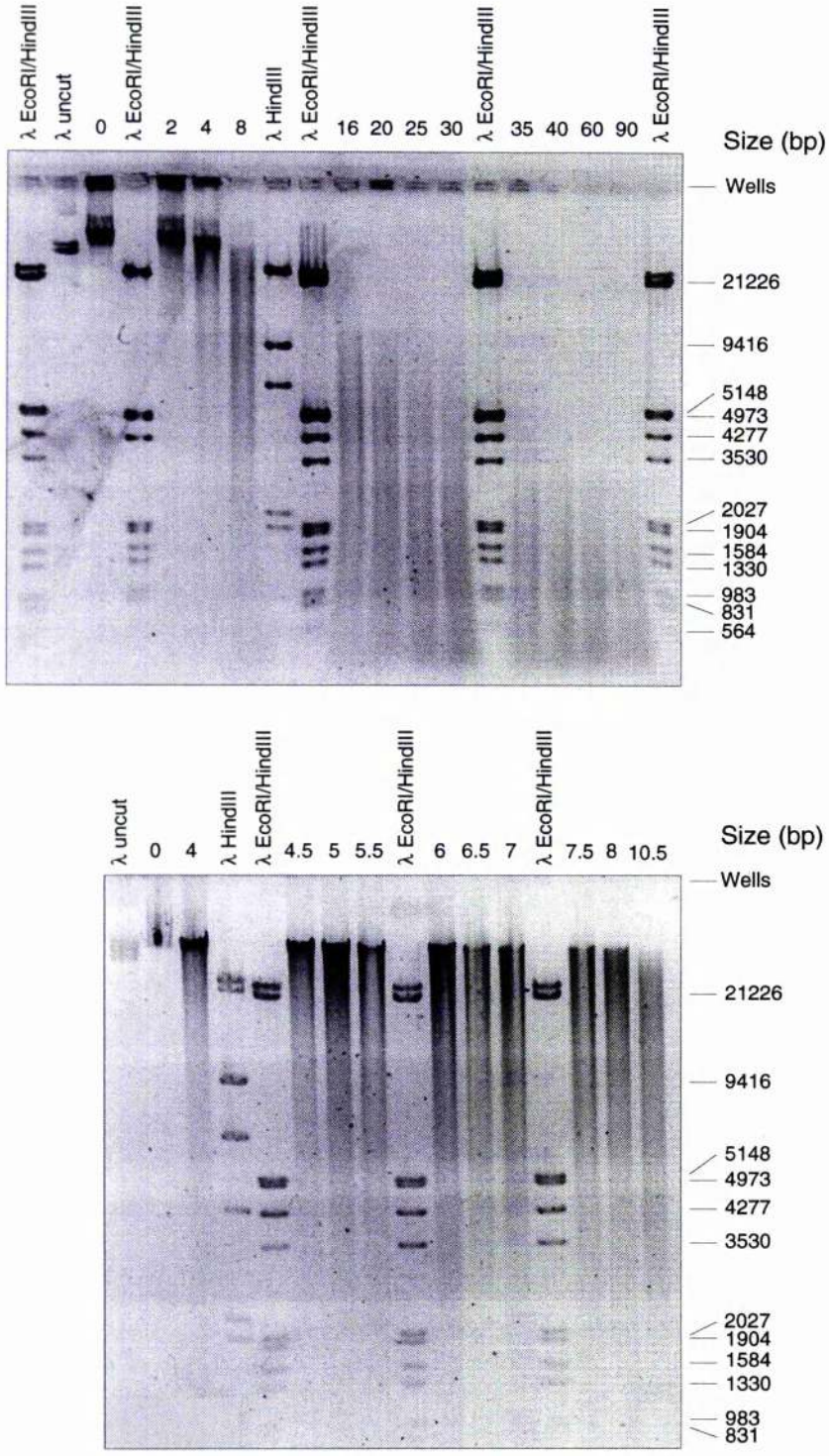
Note: Samples of strain 151 genomic DNA (5 µg) were digested with the restriction enzymes named above each lane and run on a 0.8% agarose gel (pictured on the left). A sample of undigested strain 151 genomic DNA was also loaded for comparison (labelled 'Undigested'). DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes (and HindIII alone), the size of the resulting fragments is given in the centre (1 µg of markers were loaded in the lefthand marker lane, 500 ng in the righthand marker lane). The white area at the bottom centre of the gel represents a label in the photograph which has been removed for clarity. The DNA was transferred to nylon membrane and probed with the α consensus oligonucleotide, the resulting autoradiograph is shown on the right. Only those enzymes which produced fragments of useful sizes are labelled, the sizes of the pucA carrying fragments generated are given on the right of the autoradiograph.

This screening of restriction enzymes was also a convenient point to pinpoint those which would be useful in the later subcloning of the puc genes. Thus a Southern blot was performed and the blot probed with the α consensus oligonucleotide. The resulting autoradiograph is also shown in Figure 23 (opposite). All samples hybridised with the oligonucleotide probe as can be seen by the bands at the top of the autoradiograph, several enzymes also yielded hybridising fragments of a size useful for subcloning the puc genes. These have been labelled above the autoradiograph and are BamHI (3.6 kb fragment), SalI (3.6 kb fragment), PstI (3.1 kb), AccI (3 kb) and XhoI (2.1 kb). The phage λ DNA used as markers can also be seen to have hybridised to the probe, but this appears to be non-specific binding, a problem often associated with oligonucleotide probes.

Thus although the strain 151 genomic DNA showed a somewhat abnormal susceptibility to digestion by some common restriction enzymes, enzymes were identified which could be used for the construction of the genomic library (Sau3AI) and for subcloning the puc genes (XhoI).

Figure 24

Agarose gel analysis of the digestion of *Rv. gelatinosus* strain 151 genomic DNA by Sau3AI restriction enzyme during a time course reaction.



Note: Samples of strain 151 genomic DNA (30 µg) were digested with Sau3AI in a time course reaction and aliquots run on a 0.4% agarose gel. The upper gel shows an initial digest, with aliquots removed over a long time period, the lower gel shows a later digest with aliquots taken over a shorter period (the time the aliquot was removed is given above each lane. A sample of undigested λ DNA was also loaded for comparison (labelled 'λ Uncut'). DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes (and HindIII alone), the size of the resulting fragments is given to the right of the gels (1 µg of markers were loaded).

Size fractionation of genomic DNA

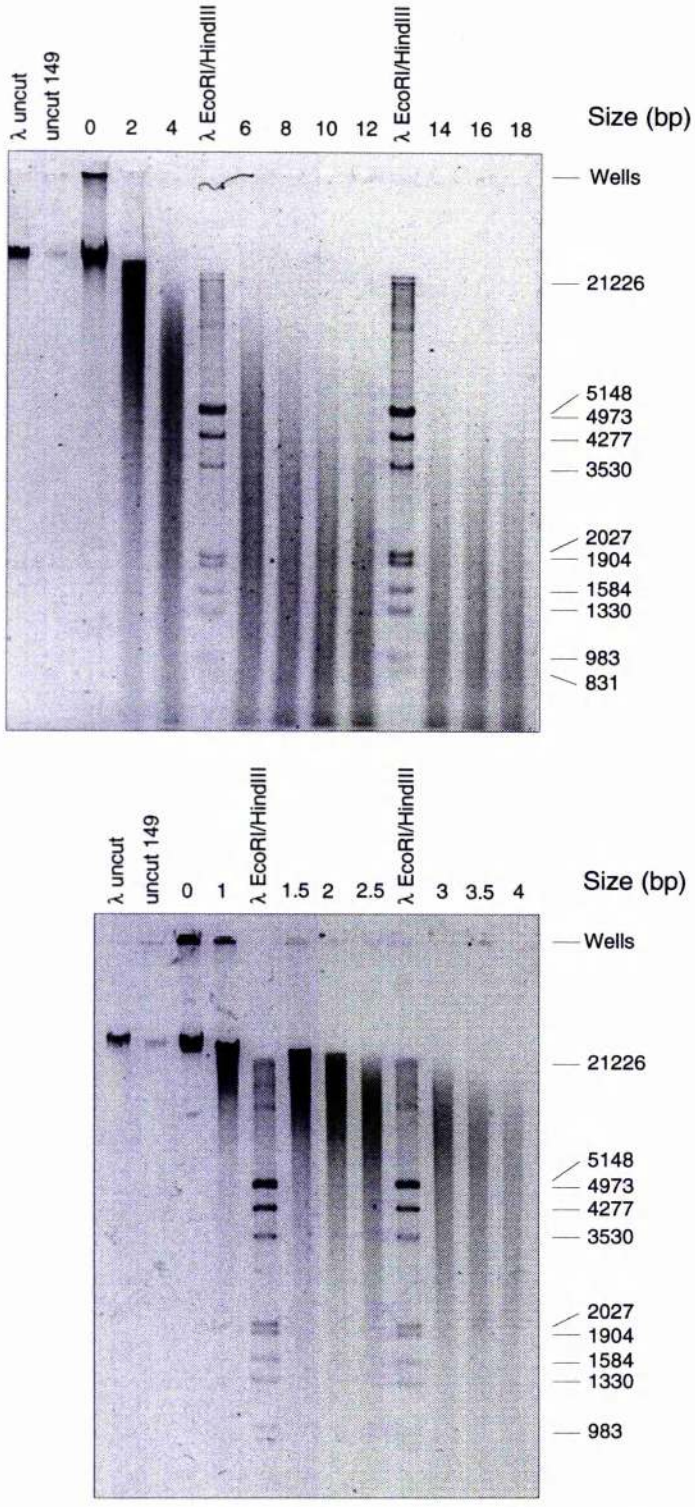
Sau3AI digestion

Having established that the restriction enzyme Sau3AI digested the genomic DNA of strain 151 satisfactorily, it was used to generate fragments of a size suitable for cloning in a λ replacement vector. This was achieved by carrying out time course reactions as described in the Materials and methods section "Digestion with Sau3AI". In an initial reaction aliquots were removed over a wide time period, these were analysed by agarose gel electrophoresis to identify a time range which yielded fragments of the correct size. Figure 24 (opposite) shows the results of these initial digest on strain 151 and Figure 25 (overleaf) those from strain 149. If one looks at the upper gel pictured on Figure 24 it is possible to see that the genomic DNA present in the zero time point sample is larger in size than λ DNA since it lies higher up the gel than the uncut λ DNA in the lane to the left. As the digest progresses this genomic DNA band gets progressively smaller and becomes a smear from the 8 minute time point onwards. By the 35 minute time point the DNA is almost completely digested and is visible as a smear at the bottom of the gel. Looking more closely at the 2 to 16 minute time points it appears that the 4 and 8 minute samples contain the largest proportion of fragments around 20 kb in size. Figure 25 (overleaf) shows that the digestion of genomic DNA by Sau3AI occurred faster when strain 149 DNA was used, the DNA being almost completely digested in around 12 minutes. The optimal digestion time for the generation of 20 kb fragments was also shorter at 2 to 4 minutes, as can be seen if one looks at those time points in Figure 25. This difference in the time taken to digest the genomic DNA is almost certainly due to the lower concentration of genomic DNA used in the strain 149 digestion (see the Materials and methods section "Digestion with Sau3AI" for more details).

Using these optimal times of 4 to 8 minutes for strain 151, and 2 to 4 minutes for strain 149, further digests were carried out but aliquots removed from the reactions at and between these time points. The results of one such digest for each strain are displayed in the gels pictured in the lower halves of Figure 24 and Figure 25. The lower part of Figure 24 pictures a gel analysing the digest of strain 151 DNA, this clearly did not progress as well as the initial digest since there is a proportion of fragments larger than the 21 kb λ DNA marker in all samples through to the 8 minute time point. There is however also a smear running down each lane to well below the 21 kb marker, indicating that all samples contain some 20 kb fragments. This incomplete digestion was probably a result of incomplete mixing of the reaction mix due to the high viscosity of genomic DNA. The digests were repeated three times, the mixing problem was encountered each time and the results obtained were similar to those displayed in Figure 24. However since many samples contained a proportion of 20 kb fragments, it was decided to pool the samples containing the highest proportion of 20 kb fragments for the subsequent size separation step.

Figure 25

Agarose gel analysis of the digestion of *Rv. gelatinosus* strain 149 genomic DNA by Sau3AI restriction enzyme during a time course reaction.

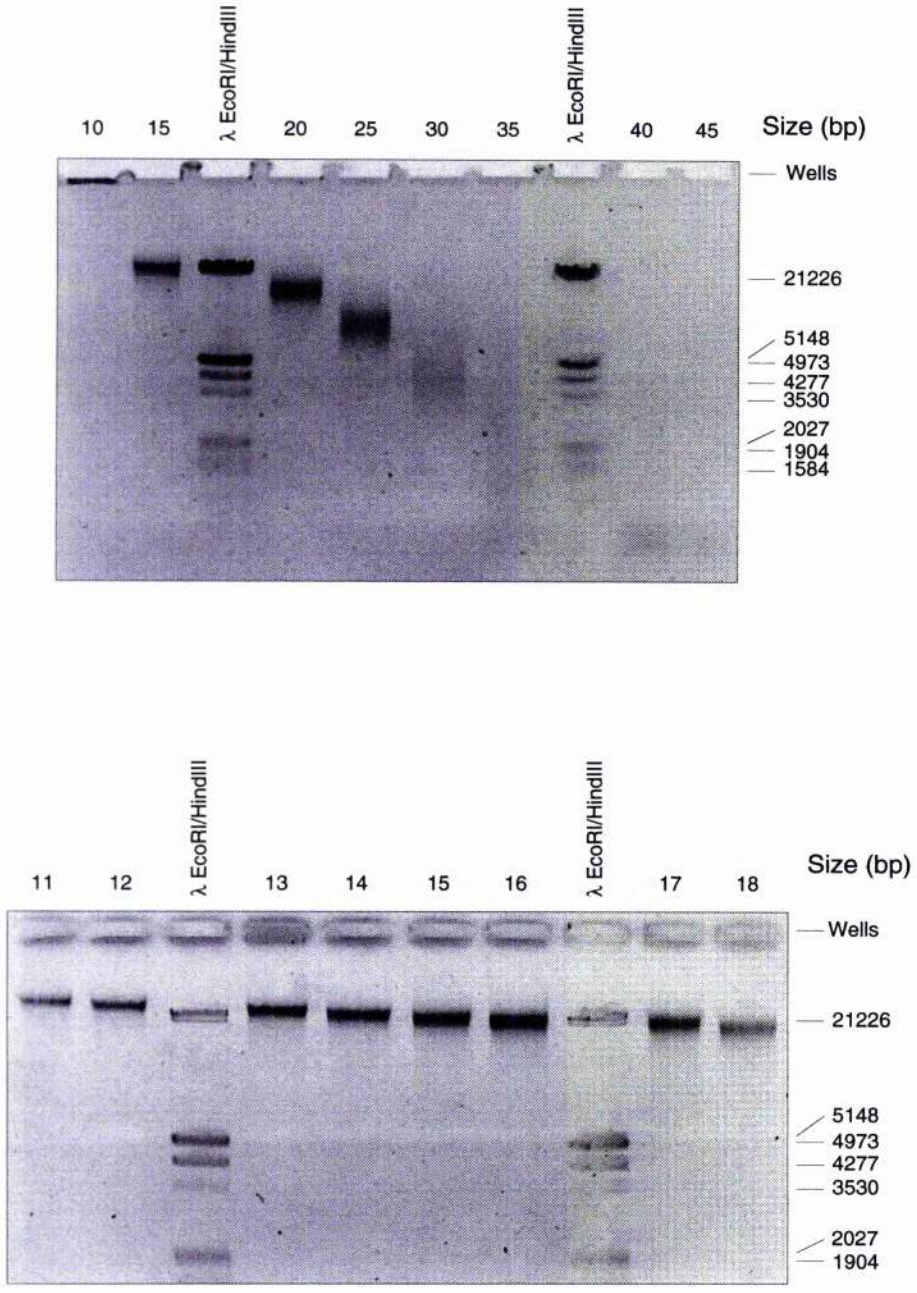


Note: Samples of strain 149 genomic DNA (260 µg) were digested with Sau3AI in time course reactions and aliquots run on a 0.4% agarose gel. The upper gel shows an initial digest, with aliquots removed over a long time period, the lower gel shows a later digest with aliquots taken over a shorter period (the time in minutes is given above each lane). A sample of undigested λ DNA was also loaded for comparison (labelled 'λ Uncut'). DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes (and HindIII alone), the size of the resulting fragments is given to the right of the gels (1 µg of markers were loaded).

The digestion of strain 149 DNA over the 1 to 4 minute range is seen clearly in the gel forming the lower part of Figure 25 (opposite). This appears to have worked more effectively than the digestion of strain 151 DNA since the bulk of each sample gets progressively smaller as the digest continues. It seems likely that the mixing problem was not encountered with strain 149 because less DNA was added to the reaction (see the Materials and methods section "Digestion with Sau3AI" for more details). Samples from 1 to 3 minutes each have a smear beneath the 21 kb λ DNA marker indicating that they all contain 20 kb fragments. As with strain 151 three such digests were carried out on strain 149 genomic DNA, and the samples containing the highest proportion of 20 kb fragments pooled for use during the subsequent size separation step.

Figure 26

Agarose gel analysis of the digestion of *Rv. gelatinosus* strain 151 genomic DNA by Sau3AI restriction enzyme during a time course reaction.



Note: Fractions from a sucrose gradient run on a 0.4% agarose gel. (fraction number is given above each lane. The upper gel shows every 5th fraction between 10 and 45, whilst the lower gel focuses in on fractions 11 to 18. DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes, the size of the resulting fragments is given to the right of the gels (500 ng of markers were loaded).

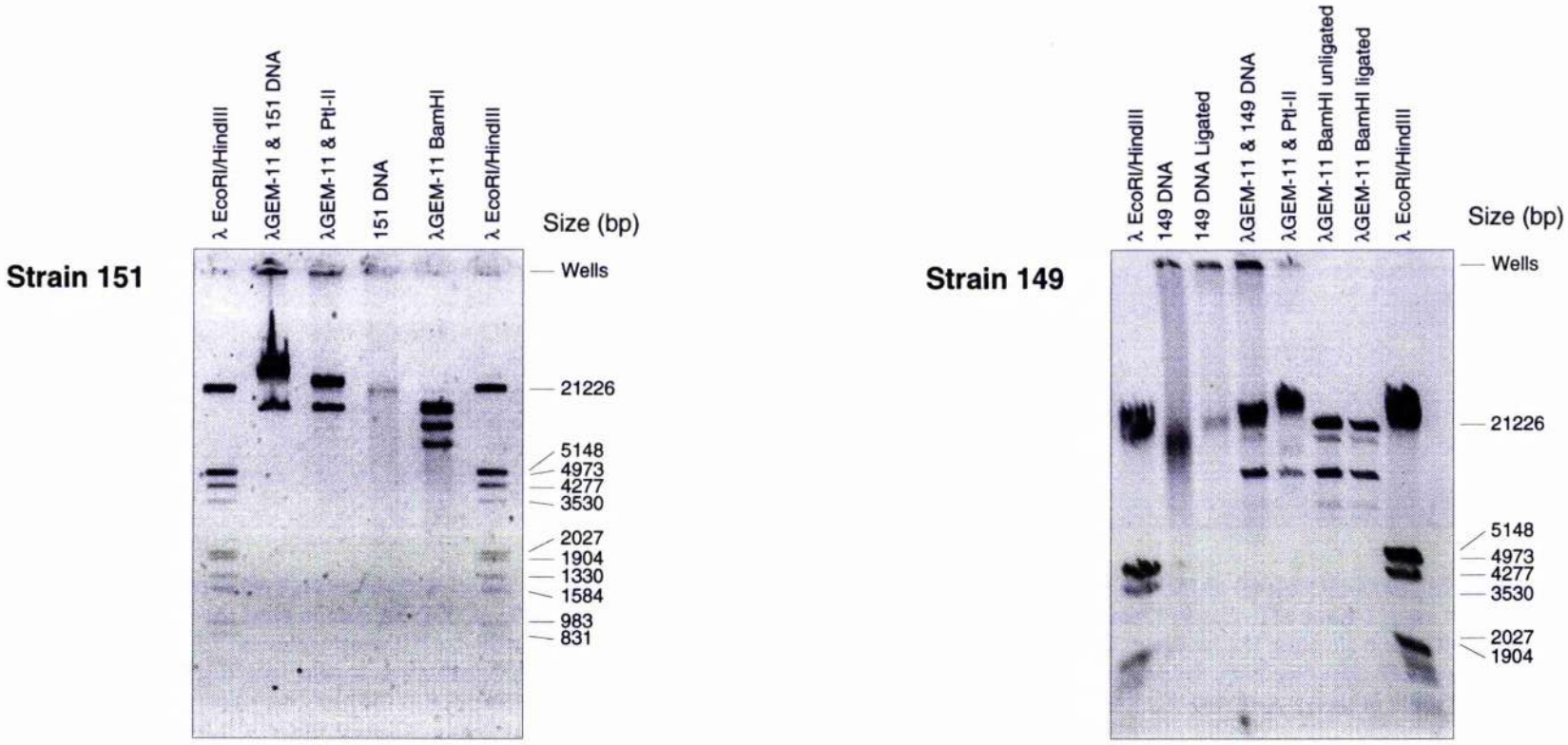
Sucrose gradient fractionation

As described in the previous section Sau3AI digestion yields a range of fragment sizes, however fragments of around 20 kb are cloned most efficiently in λ. It is thus desirable to size separate the digested fragments and isolate those of 20 kb size, this was done by running the pooled samples from the previous step on a sucrose gradient as described in the Materials and methods section "Size fractionation of Sau3AI digested genomic DNA". Fractions of approximately 0.5 ml were collected from the gradients and a small aliquots analysed on agarose gels to determine those containing fragments of the correct size. Initially a selection of fractions was checked representing the whole gradient, the upper gel pictured in Figure 26 (opposite) shows every fifth aliquot from the sucrose gradient carried out using the pooled strain 151 samples. Fraction 10 which came form near the bottom of the gradient appeared to contain very large fragments of DNA since the only band is just beneath the well, as one moves up the gradient the bands of DNA move progressively further down the gel and become more diffuse. Since fraction 15 appears to contain fragments of a size similar to the 21 kb λ DNA marker, and fraction 20 appears much smaller (~15 kb), fractions from 11 to 18 were analysed on a second gel. This gel is pictured in the lower part of Figure 26 (opposite), once again the fragments get progressively smaller as one moves up the gradient. The DNA band in fraction 17 lies just below the 21 kb marker and is thus likely to contain fragments of DNA around 20 kb in size, this fraction was selected for the construction of the strain 151 library.

Very similar results were obtained during the size fractionation of the Sau3AI digested strain 149 DNA, the resulting agarose gels appear almost identical to those for strain 151 and are thus not presented. The selected samples for both strains were then used in a ligation reaction the results of which are presented in the next section.

Figure 27

Agarose gel analysis of the ligation reactions performed during construction of the *Rv. gelatinosus* strain 151 and strain 149 genomic libraries.



Note: The gel on the left shows the ligation reactions carried out during the construction of the strain 151 genomic library and that on the right the strain 149 library. DNA markers were loaded in the left and right edge lanes of each gel (100 ng of markers were loaded on the strain 151 gel, 200 ng on the strain 149 gel). The strain 151 gel to the left (0.4% agarose) has from left to right: prepared 151 genomic insert DNA ligated to the vector arms, λGEM-11 ligated to the test insert PtI-II, prepared 151 genomic insert DNA alone, and the λGEM-11 vector arms alone (note that this latter sample was buffered differently and appears to have run further down the gel than equivalent sized markers). The strain 149 gel to the right (0.3% agarose) has a slightly different layout, from left to right: prepared strain 149 insert DNA alone, 149 insert DNA ligated to itself, 149 DNA ligated to the λGEM-11 vector arms, λGEM-11 ligated to the test insert PtI-II, λGEM-11 arms ligated to themselves and λGEM-11 arms unligated.

Ligation

The sized fractionated genomic DNA was prepared for ligation as described in "Purification of size fractionated genomic DNA" in the Materials and methods, and ligation carried out as described in the following section "Ligation to the vector arms". Following ligation the reactions and unligated ingredients were examined on an agarose gel to assess the success of the ligation reaction.

On the left of Figure 27 (opposite) are the results of ligation reactions carried out during the construction of the strain 151 library. The lane second form the left contains an aliquot of the vector (λGEM-11) ligated to strain 151 genomic DNA, this ligation was used in the construction of the library. Two bands are visible, the lower band is the central stuffer fragment which had been cut from the vector and should not participate in the ligation reaction, the larger upper band which streaks out towards the well is the ligated vector arms and genomic insert. The streaking which occurs in this band and the fact that some of the sample remains in the well suggest that this ligation has worked well, forming the long concatameric molecules which are packaged most efficiently by phage λ. It is also notable that the ligated DNA fragments are much larger (at around 50 kb) than either the genomic DNA alone (4th lane from the left) or the vector arms (5th lane from the left), and further that those bands have disappeared from the sample, suggesting that almost all of the insert and vector DNA added has been ligated together. The control reaction of the vector arms ligated to a test insert (PtI-II) was loaded in the third lane from the left and has also worked well, the vector arm and insert bands having disappeared and been replaced by a single band at around 44 kb in size. The vector arms as supplied were run in the lane 5th from the left, four bands can be seen representing (from the bottom up): the 9 kb vector arm, the 14 kb central stuffer fragment, the 20 kb vector arm, and at 29 kb the two vector arms connected by means of their 12 bp cohesive ends[10].

The ligation reactions carried out at the time the strain 149 library was constructed are shown in the right hand gel on Figure 27 (opposite). The prepared strain 149 insert DNA ligated to the vector arms was run in the lane 4th from the left, as with strain 151 the product of this ligation is larger than the insert or vector DNA alone and is concatameric. The product differs from that of the strain 151 ligation however, in that it is smaller than the product of the control ligation of vector to PtI-II, visible in the lane to the right. This suggests that the prepared strain 149 insert DNA is smaller than the intended 20 kb, but should be of sufficient size since the packaging mix used can package λ DNA molecules between 23 and 9 kb. The 6th and 7th lanes from the left both contain vector DNA, however ligation was attempted on the sample in the 7th lane, since the resulting bands are almost identical to those of the unligated vector in the 6th lane ligation could not have taken place. This confirms that the phosphotase treatment of the λ arms carried out by the manufacturers was effective, since it should prevent ligation of the λ arms to each other.

[10] This sample had no ligase buffer added, as a result the DNA has acted differently to the other samples, and the bands thus appear to be smaller than the quoted size.

Figure 28

Expected and observed plating efficiencies of the *Rb. gelatinosus* strain 151 and strain 149 genomic library packaging reactions and controls.

Strain 151

Reaction	Expected efficiency (pfu μg^{-1})	Observed efficiency (pfu μg^{-1})
Genomic Library	1×10^6	8.7×10^5
Control insert	1×10^7	1.3×10^7
Packaging control	2×10^9	2×10^9
Ligated vector arms	negligible	zero
Unligated vector arms	negligible	zero
Packaging mix only	zero	zero

Strain 149

Reaction	Expected efficiency (pfu μg^{-1})	Observed efficiency (pfu μg^{-1})
Genomic Library	1×10^6	1×10^6
Control insert	1×10^7	1.3×10^7
Packaging control	2×10^9	2×10^9
Ligated vector arms	negligible	zero
Unligated vector arms	negligible	zero
Packaging mix only	zero	zero

Note: The efficiencies are given as plaque forming units (pfu) per μg of λ arms packaged. The expected packaging efficiencies are as suggested by the manufacturer of the packaging mix (Stratagene Ltd.). The packaging control reaction utilised wild-type λ and *E. coli* strain VCS257 as supplied by Stratagene Ltd., other reactions were plated in *E. coli* KW251.

Titering the packaged libraries

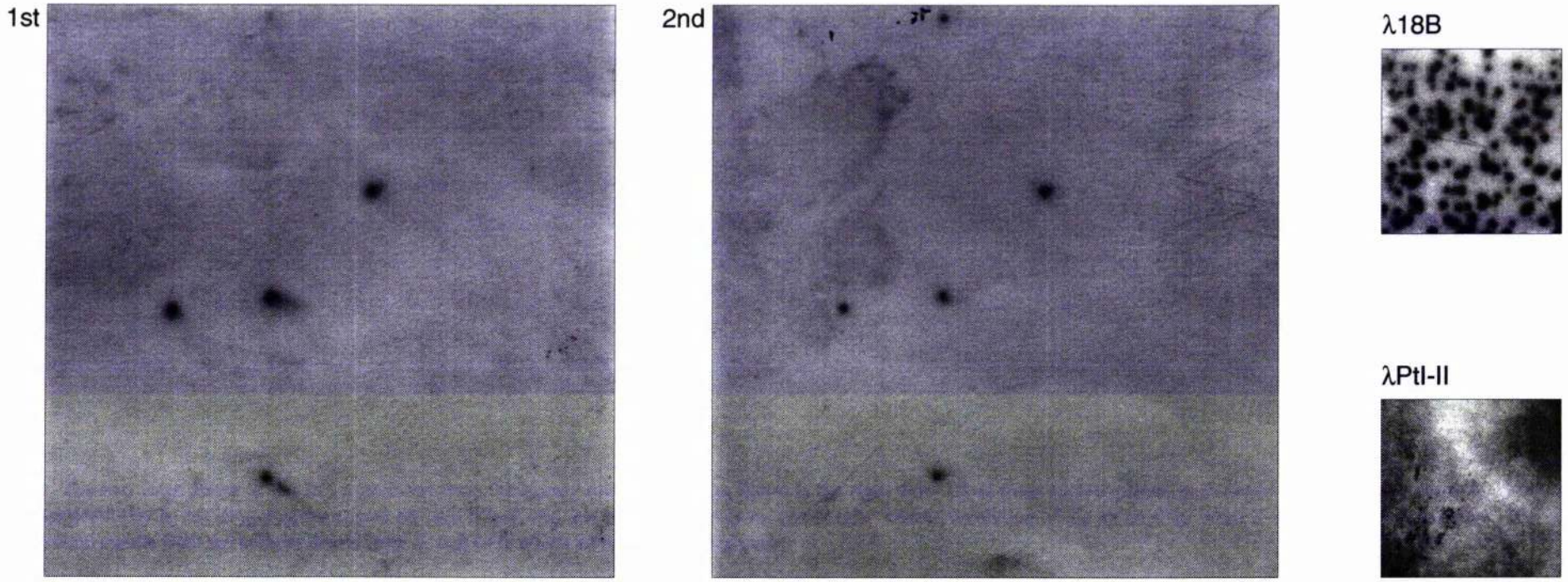
The ligated libraries and controls were packaged as described in the Materials and methods section "The packaging reaction", as were several packaging controls. The packaged phage were then plated as described in the Materials and methods section "Titering the library", and the efficiency of the packaging reactions calculated. The expected efficiency and the observed efficiency of these reactions is given in Figure 28 (opposite). The efficiency is measured in terms of the number of plaque forming units (pfu) generated per μg of λ arms added to the packaging reactions.

According to the manufacturers instructions the packaging mix used should achieve efficiencies of 1 x 10^6 pfu μg^{-1} when the genomic library is packaged, this level of efficiency was achieved for strain 149, but the efficiency was lower when the strain 151 library was packaged. This shortfall could not have been due to problems with the packaging mix since the packaging control plated at full efficiency, nor could there have been a problem with the plating bacteria because the packaged control insert ligation (utilising PtI-II) plated more efficiently than expected. Thus there is no clear reason why this shortfall in efficiency occurred. The controls labelled as ligated and unligated vector arms were carried out to check the level of background contributed by uncut vector, which may have been present in the digested vector DNA. The plating of packaging mix only allows one to ensure that there are no viable phage in the packaging mix. All of these latter controls worked as expected for both libraries and suggest that the level background phage present in the genomic library is zero, or so close to zero that it can be ignored for practical purposes.

Although full efficiency was not achieved in the *Rv. gelatinosus* strain 151 genomic library more than enough clones (around 600,000) were generated for effective screening and cloning (around 1000 clones probably cover the whole genome). It was therefore decided to continue and screen the library as planned. The strain 149 library packaged at full efficiency thus there was no problem in proceeding to the screening stages with the library of this strain.

Figure 29

Autoradiographs of filters lifted from the plated *Rv. gelatinosus* strain 151 genomic library, and probed with the α consensus oligonucleotide during primary screening of the library.



Note: The two large filters to the left were lifted from the library plates whilst the filters to the right were lifted from control plates, both negative and positive controls. Note that genuine positively hybridising plaques appear on both filters, and that plaques often have 'comet tails' caused by sliding of the nylon filter when it is removed from the plate. The high background results from the cellular debris from *E. coli* cells which adhere to the membrane.

Primary screening

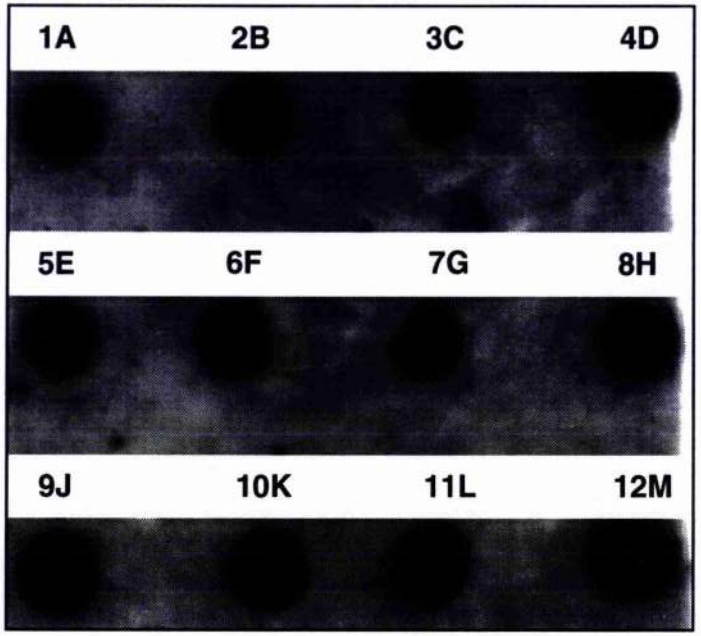
Once the efficiency of packaging had been determined as described in the previous section, a large number of pfu (~5000) were plated for the purposes of screening as described in the Materials and methods section "Primary screening of the library". Several filters were lifted in duplicate, however only one pair is shown, these are pictured in Figure 29 together with positive and negative control filters.

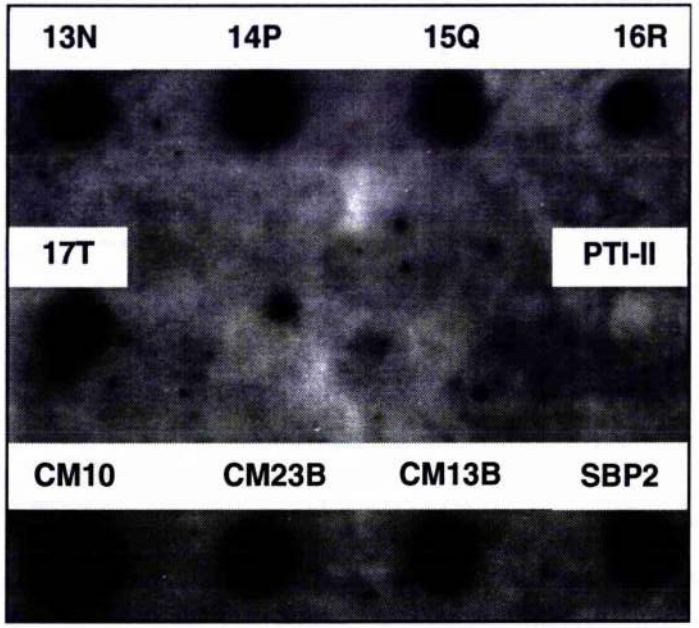
The two large autoradiographs displayed in Figure 29 (opposite) are of filters lifted from plates of clones from the strain 151 library, which were probed with the α consensus oligonucleotide. Positively hybridising plaques are distinguishable from background spots on these filters because they occur in the same position on both filters in a pair, in this case 4 plaques can be seen. They also often have a characteristic 'comet tail' because some spreading occurs when the filter is removed from the plate. This is particularly noticeable on the first filter lifted (pictured on the left of Figure 29). Shown to the right of the two large autoradiographs are fragments of two smaller autoradiographs, taken from filters lifted off control plates. The positive control plate contained the clone λ18B at relatively high density, this clone contains some puc genes from *Rps. acidophila* and was known to hybridise with the probe. The negative control filter was created from the control insert (Ptl-11) carrying phage, plated whilst titering the library (see previous section). Both these controls hybridised in the expected manner, indicating that the hybridising plaques were likely to be true positives.

The strain 149 library was screened using the pucBA genes cloned from strain 151 but otherwise the procedure and results were similar. Approximately 40 positive plaques were picked from the library plates of each strain rescreened.

Figure 30

Autoradiographs of filters lifted from the plated *Rv. gelatinosus* strain 151 genomic library, and probed with the α consensus oligonucleotide during primary screening of the library.





Note: 1 ml of phage solution was spotted onto a lawn of KW251. Following growth filters were lifted and probed with α consensus oligonucleotide. The clone marked PTI-II acts as a negative control, whilst the 4 clones at the bottom of the right hand autoradiograph act as positive controls since they are known to contain sequence homologous to the probe.

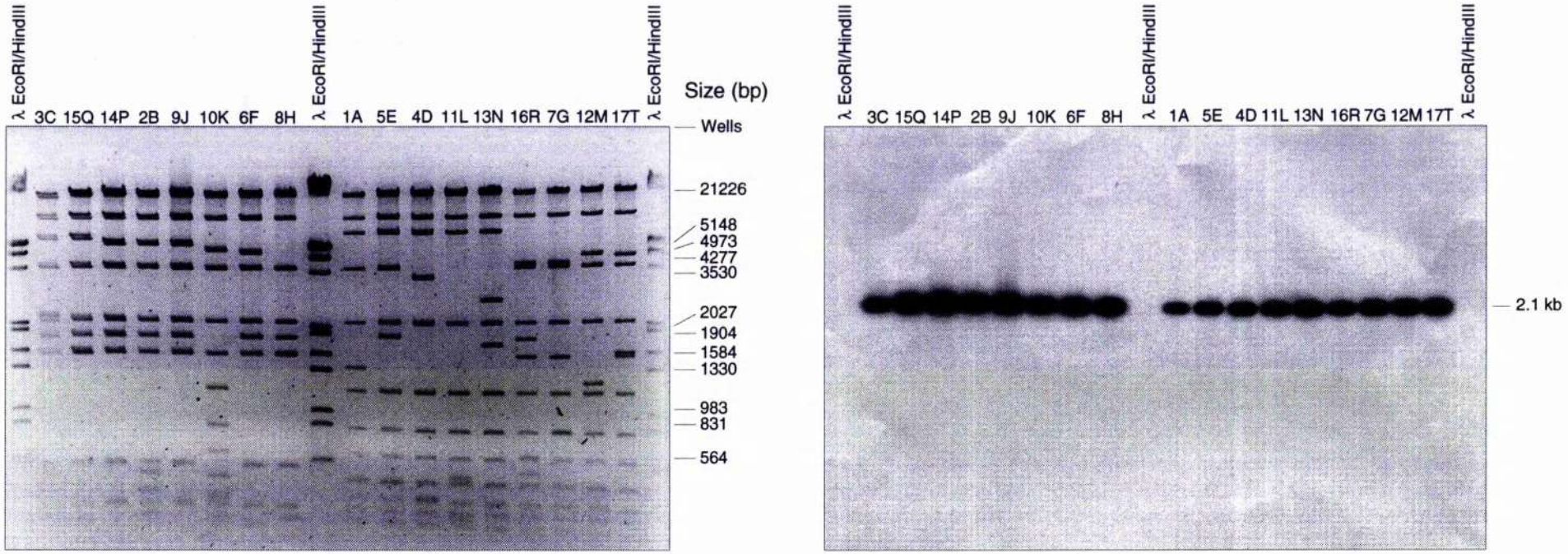
Secondary screening

The positively hybridising clones identified in the primary screening were purified, then rescreened by spotting them in a grid onto a lawn of plating bacteria as described in "Secondary screening" in the Materials and methods. Filters were then lifted and hybridised to check that the purified clones were still true positives. Following this several of the picked clones appeared not to be positives, a result of contamination with other phage during the initial picking process due to the high density of plating (around 1000 plaques per 100 cm^2 plate). However sufficient positive clones were identified, seventeen were selected for strain 151 and spotted onto a lawn once more, the results of this third screening are displayed in Figure 30 (opposite). If one compares the hybridisation signal obtained for each strain 151 clone with that of the 4 positive controls (λCM10, λCM23B, λCM13B, and λSBP2 at the bottom right), which all carry copies of the puc genes from *Rps. acidophila* it is clear that all 17 clones are strong positives. A negative control represented by the λPtI-II test clone can be seen not to hybridise with the probe, further supporting the conclusion that all the clones spotted are true positives.

During the secondary screening process of the strain 149 genomic library many of the primary positives were lost. However 8 positively hybridising clones were identified. In an attempt to avoid confusion between clones of the two strains those from strain 151 were named using letters and numbers, placing the number foremost (e.g. λ17T), whilst those from strain 149 were named with a letter formost (e.g. λF6). Following identification of the positive clones their DNA was isolated and restriction mapped, the results of this mapping are described in the following section.

Figure 31

Agarose gel analysis of the digestion of *Rv. gelatinosus* strain 151 genomic λ clones with XhoI, and autoradiograph of a subsequent Southern blot probed with the α consensus oligonucleotide.



Note: DNA purified from seventeen isolated λ clones was digested with XhoI and analysed on a 0.8% agarose gel (pictured above left). Approximately 1.2 µg of each clones DNA (named above each lane) was loaded. DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes, the size of the resulting fragments is given in the centre (200 ng of markers were loaded). The DNA was transferred to nylon membrane and this probed with the α consensus oligonucleotide, the resulting autoradiograph is shown on the right. The estimated size of the single hybridising band is given to the right of the autoradiograph.

Mapping the λ clones

The positively hybridising clones identified during the screening process were cultured and their DNA isolated and purified to facilitate restriction mapping of the clones. Previous work presented at the beginning of this Results section had suggested that digestion or *Rv. gelatinosus* strain 151 DNA with the enzyme XhoI generated a single fragment carrying the target sequence. Since this 2.1 kb fragment was also of a size suitable for subcloning and sequencing it was decided to use this enzyme to map the isolated λ clones.

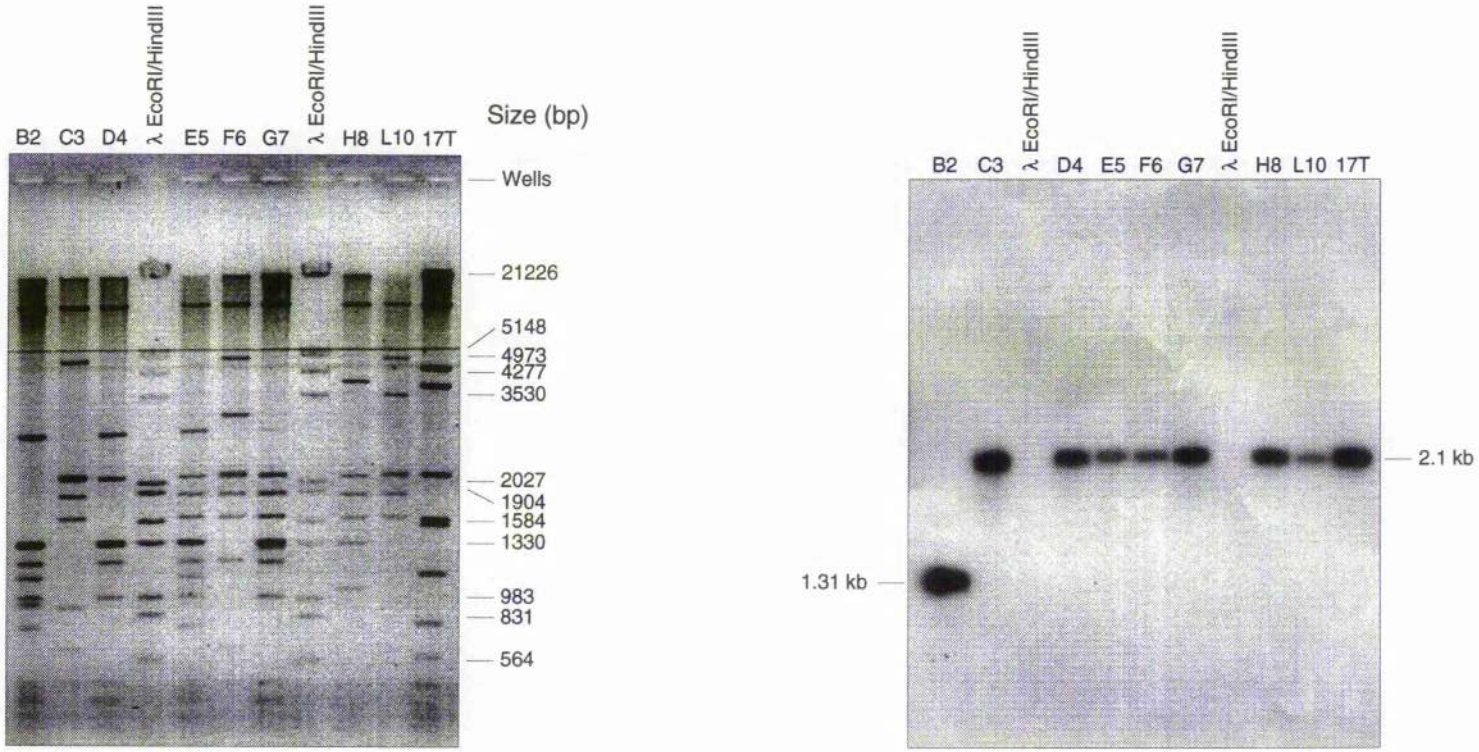
The seventeen strain 151 clones were digested with XhoI and analysed by agarose gel electrophoresis. The resulting gel was analysed and the clones arranged into groups with a similar restriction pattern. A second gel was then run with the samples loaded in an order based on their groups, this gel is pictured on the left of Figure 31 (opposite). The prearrangement of the clones makes the comparison of their restriction patterns somewhat easier. If one looks towards the top of the gel the two topmost fragments appear to be present in all 17 clones, these are the 20 kb and 9 kb vector arms. Only one other fragment is present in all clones, this can be seen lying just above the 2027 bp λ marker, and should contain the sequence to which the probe used to screen the library binds. To confirm this the gel was blotted onto nylon membrane and the filter probed with the α consensus oligonucleotide, the resulting autoradiograph is presented to the right of the gel on Figure 31. A single band of 2.1 kb in size can be seen to hybridised to the probe, confirming that this fragment carries the probe target sequence.

A similar result was obtained when the 8 strain 149 clones were digested with XhoI, the results of which can be seen on the left of Figure 32 (overleaf). The strain 151 clone λ17T was loaded on this gel at the right hand side for comparison. It is clear from the comparison of λ17T with the strain 149 clones that there is little in the way of common restriction fragments between the two strains. The only common fragment (apart form the vector arms) is a 2.1 kb fragment which is present in 7 of the strain 149 clones. Southern blotting of a similar gel followed by hybridisation with a fragment carrying the strain 151 pucBA genes yielded the autoradiograph pictured on the right of Figure 32. This autoradiograph shows that a single hybridising fragment is present in each clone, for 7 of the clones this matches the 2.1 kb fragment that the clones have in common. Clone B2 contains no 2.1 kb fragment, but has a hybridising fragment of 1.3 kb, indicating that the 2.1 kb fragment is truncated in this clone.

The gels pictured in Figure 32 were used to construct an XhoI restriction map covering the area of the *Rv. gelatinosus* strain 151 and 149 genomes represented by the isolated λ clones. The maps produced are described below.

Figure 32

Agarose gel analysis of the digestion of *Rv. gelatinosus* strain 149 λ clones with XhoI, and autoradiograph of a subsequent Southern blot probed with the 151 pucBA genes.



Note: The purified DNA from 8 strain 149 λ clones was digested with XhoI and analysed on a 0.8% agarose gel (pictured above left). Approximately 1 µg of each clone (named above the gel lanes) was loaded. For comparison the strain 151 clone λ17T was also digested, and loaded at the right hand edge of the gel. DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes, the size of the resulting fragments is given in the centre (500 ng of markers were loaded). The black line across the width of the gel approximately one third of the way down is an artifact caused by the imaging system used to photograph the gel and should be ignored. A similar gel on which 250 ng of DNA from the same clones were run was transferred to nylon membrane and probed with the strain 151 pucBA genes, the resulting autoradiograph is shown to the right. The sizes of hybridising bands are given at either side of the gel.

Figure 33

Restriction maps of the strain 151 & 149 genomic λ clones derived from XhoI restriction digests of the clones. Strain 151 λ clones are at the top, strain 149 at the bottom, each clone is labelled with its name at the right hand end. The lines marked 151 map and 149 map represent the respective genomes, the fragment sizes are given next to these maps. The vertical lines represent XhoI sites in the DNA. Where two or more fragment sizes are given in brackets the exact arrangement is not clear, additional XhoI sites are of course present in these regions but are not marked. The positively hybridising 2.1 kb XhoI fragment is shown in bold, an XhoI site is also present at each end of this fragment but is not marked.

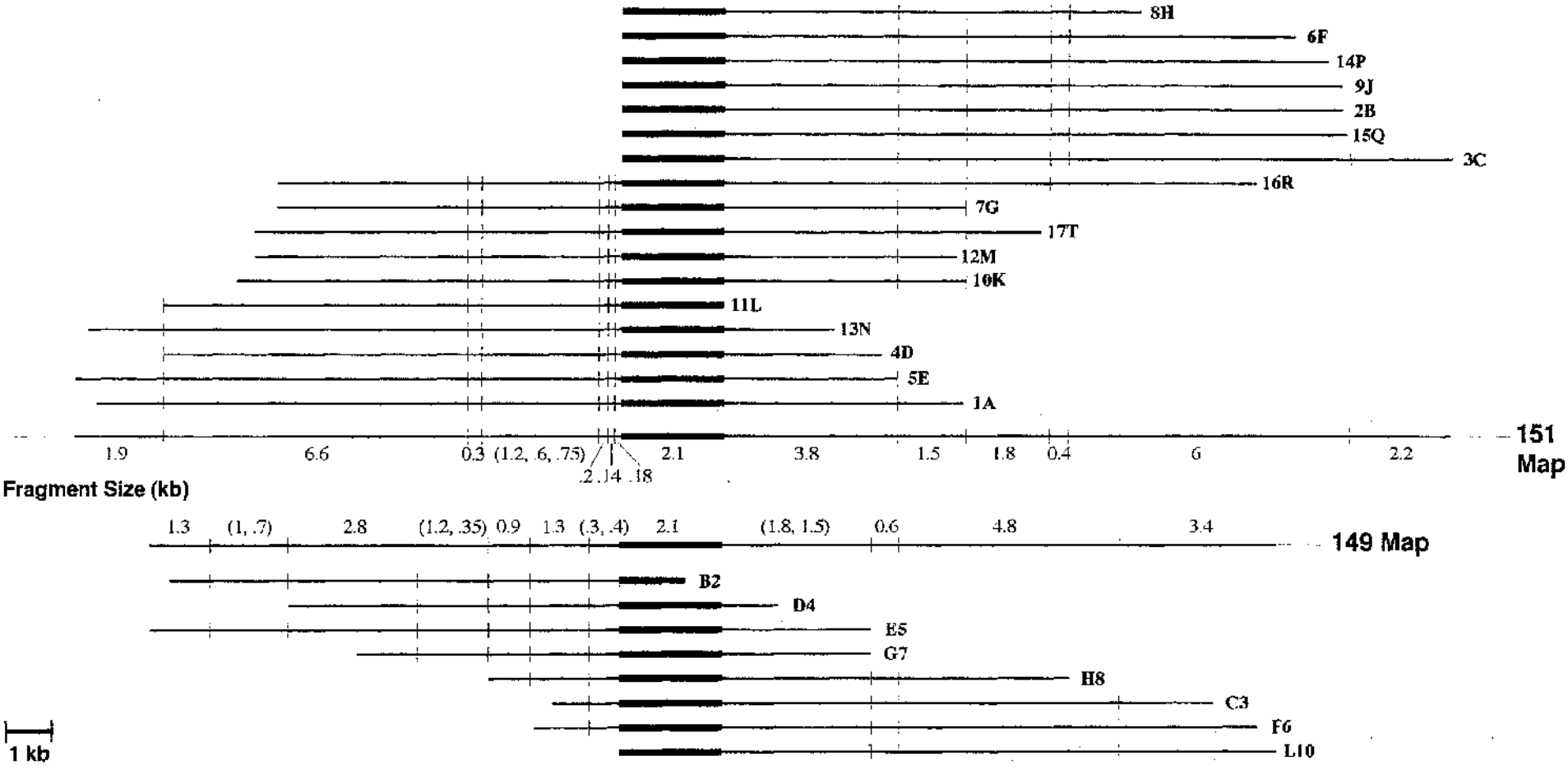


Figure 33 (opposite) displays the maps constructed from the strain 151 and 149 λ clone XhoI restriction digests. In the centre of the Figure are the two maps representing the genomes of strain 151 and 149, above and below these the extent of each clone is pictured. It should be noted that the correct orientation (left to right, or right to left) of the two maps with respect to each other is unknown. The strain 151 map extends over approximately 30 kb of the genome, whilst that for strain 149 covers approximately 27 kb. The slightly shorter 149 map may stem from the smaller size of the strain 149 λ clones, these ranged in size from 9.5 to 15.6 kb, and only two out of 8 clones were over 15 kb long. In contrast the strain 151 clones ranged from 10 to 17 kb, and 14 of the 17 clones were 15 kb or over.

More interestingly the pattern of XhoI sites appears to differ between the two strains, the only conserved sites in this area of the two genomes appear to be those at each end of the positively hybridising 2.1 kb fragment.

On the basis of the restriction maps pictured opposite λ clones were selected for the subcloning of the 2.1 kb fragments. For strain 151 the clone λ17T, with its centrally positioned 2.1 kb fragment was chosen. Whilst for subcloning the 2.1 kb fragment of strain 149 clone λF6 was selected. Clone λG7 has a more central 2.1 kb fragment than λF6, but is smaller (11.3 kb compared with 15.6 kb) and may not have replicated as effectively.

Conclusions

The results described above detail the successful construction of genomic libraries for the two strains 151 and 149 of *Rv. gelatinosus*. However it remains unclear why a lower plating efficiency was achieved for the strain 151 library. There are a number of reasons why this lower efficiency may have occurred, these include simple errors, such as in the measurement of DNA concentration, or more worryingly that a proportion of the strain 151 genome is 'unclonable'. The latter possibility would result in a part of the strain 151 genome not being represented within the library, the shortfall seems unlikely to be due to this however because the shortfall was so large (40%). A more likely possibility is that the prepared genomic insert was less than perfect. Shearing of the insert DNA can result in only one end being compatible with the vector arms, and contaminants could inhibit the packaging reaction. The size of the insert DNA is also important, the packaging kit used can package λ with inserts in the range of 9 to 23 kb, but those inserts varying from the 20 kb optimum are packaged with less efficiency. Since the strain 149 λ clones and thus the prepared genomic insert DNA isolated from 149 were smaller than those isolated from strain 151, the difference in efficiency between the two libraries may result from the different sized insert DNA. Thus there is a possibility that a proportion of the strain 151 genomic insert DNA was too large to be packaged.

Although the plating efficiency of the strain 151 library was lower than expected (for reasons which remain unclear), this did not prevent the isolation of the required clones. Ample clones were generated in each library, screening of which yielded a set of overlapping λ clones for each strain. These appear representative of the original genomes and (as was proven at a later point) contain the *Rv. gelatinosus* puc genes.

Thus in terms of the original aim to clone the puc genes the library construction was very successful.

During the course of library construction several interesting observations were made regarding *Rv. gelatinosus* and the differences between strain 151 and 149. Firstly the initial screening of strain 151 genomic DNA with various restriction enzymes combined with hybridisation analysis, suggest that the *Rv. gelatinosus* strain 151 puc genes are located within a single locus on the genome. This is further supported for strain 151 and iterated for strain 149, by the isolation of only a single set of overlapping λ clones from each strain. This suggests that *Rv gelatinosus* has only a single copy of its puc genes, like the *Rhodobacter* species, and unlike species such as *Rps. acidophila* which has multiple copies of its puc genes.

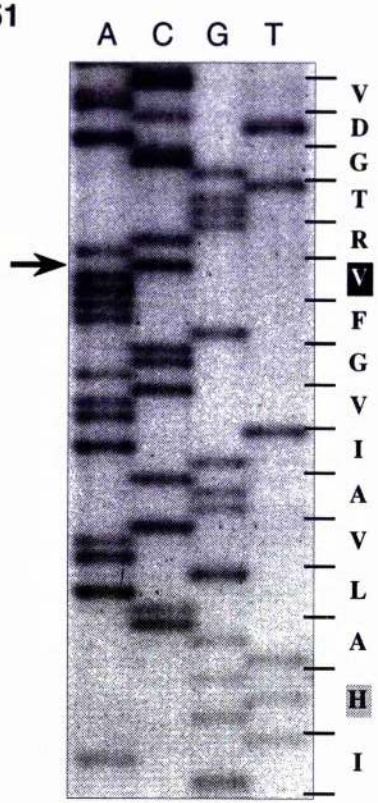
The restriction maps derived from the XhoI digested λ clones (Figure 33) also illustrate an interesting point. The two maps show almost no correspondence between the two strains, for this to be the case there must have been a substantial amount of change in the DNA of the two strains. This level of difference at the DNA level was unexpected but underpins the differences observed at the physiological level.

The final point arising from the results presented in this section concerns the digestion (or lack of digestion) of *Rv. gelatinosus* strain 151 genomic DNA by commonly available restriction enzymes. A number of explanations can be offered for this lack of digestion, aside form the previously mentioned contamination problem, which was addressed by further purifying the DNA. The first is that the DNA could be methylated at certain points, this is not generally a problem with bacterial DNA however and is certainly not the case in cloned DNA (cultured in *E. coli*). Thus the inability of enzymes like EcoRI to digest even cloned *Rv. gelatinosus* DNA allows this explanation to be excluded. The second possibility is that the restriction enzyme sites simply don't exist, or occur at a much lower frequency than is normally found. The latter explanation is extremely feasible since most restriction enzymes on the market have AT rich recognition sites, whilst the genome of *Rv. gelatinosus* is 72% GC (see Willems *et al.*, 1991). Such a high ratio of GC base pairs makes it statistically less likely that AT enzyme cut sites will occur. If one takes as an example the enzyme EcoRI which has the six base AT rich recognition site GAATTC, this would cut approximately once every 4096 bp in a base neutral DNA sequence, it can be calculated that a 70% GC rich template would be cut once every 16000 bp. Thus the simplest explanation of the reduced digestion observed for many enzymes may simply be that the GC rich DNA of *Rv. gelatinosus* is less likely to contain the AT rich recognition sites.

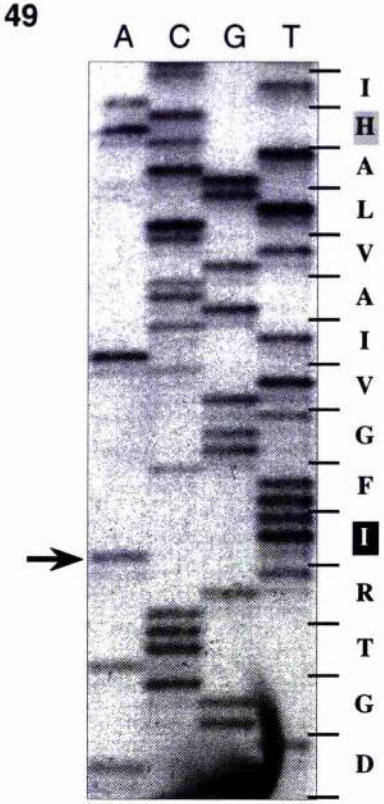
Figure 34

Results of sequencing reactions carried out to confirm the cloning of the puc genes and their source strain.

Strain 151



Strain 149



Note: Segments of the first sequencing gels from p17T2.1X (strain 151) and pF6X2 (strain 149), which show the sequence of the pucB gene. The single letter amino acid translation is given for each codon presented (the horizontal lines delineate each three base codon). The amino acid which differs between the two strains is shown as white text on black, whilst the conserved histidine residue present in most LH2 β-polypeptides is highlighted in grey. The arrows point the the nucleic acid residues which differ and cause the amino acid change. Note that the p17T2.1X (strain 151) sequence was obtained using the α consensus oligonucleotide, and thus represents the non-coding strand sequence. The sequence runs from the C-terminal end of the protein at the bottom of the gel, towards the N-terminal end at the top. For strain 149 (pF6X2) the sequence was obtained using the UR1C primer, which binds to the non-coding strand upstream of pucB, and reads the coding strand. The strain 149 β-polypeptide sequence presented thus runs in the opposite direction to that of strain 151. The black mark at the lower end of the strain 149 gel is an artifact generated during development of the autoradiograph.

Results Part 3:
Subcloning & Sequencing the Cloned Fragments

Introduction

This section details the subcloning of the puc genes from the isolated λ clones, with the subsequent DNA sequencing and Southern analysis carried out to confirm the source of the cloned fragments. Southern analysis using the pucC gene from *Rb. sphaeroides* was also carried out, and is described together with the subcloning and sequencing of a fragment homologous to the *Rb. sphaeroides* pucC gene. The construction of the nested deletions to facilitate sequencing and some DNA sequencing is also described.

Results

Subcloning the puc genes & confirming their source

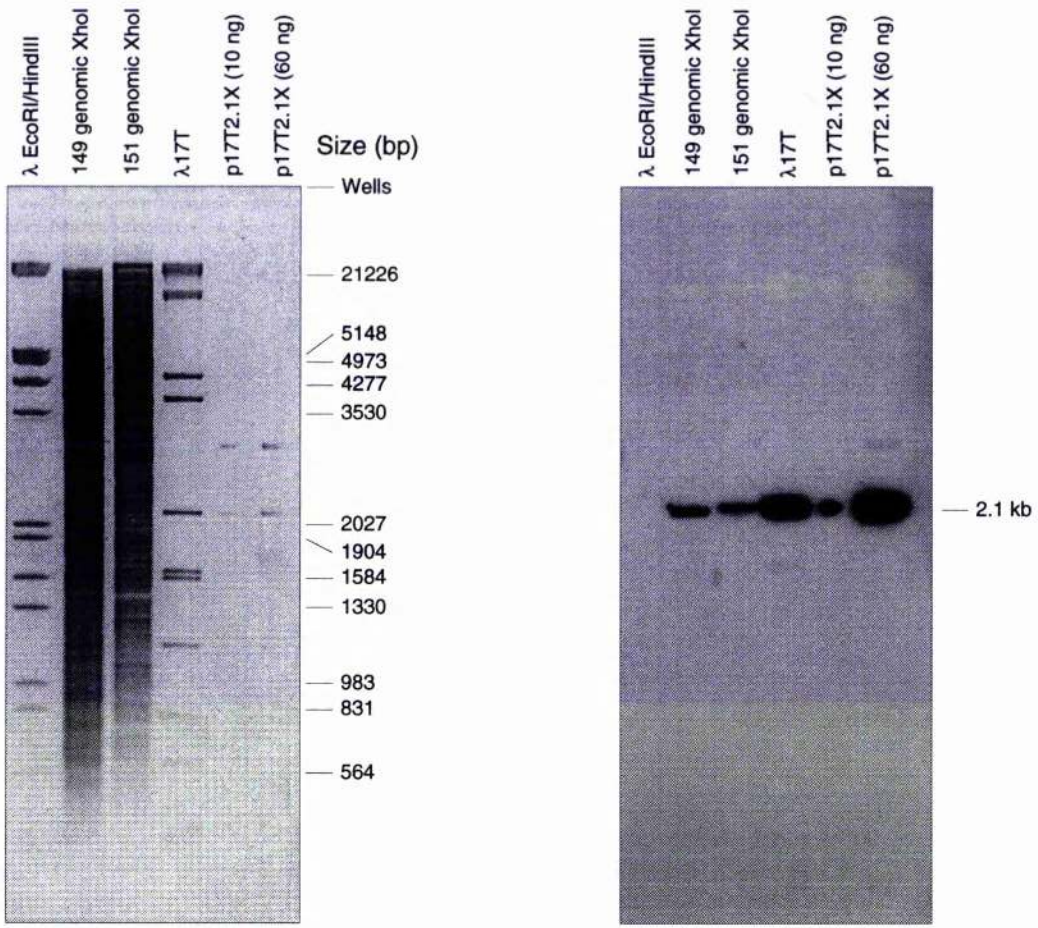
Having decided which λ clones were to be used for subcloning of the 2.1 kb fragments, the selected clones were digested with XhoI and the fragments shotgun cloned into the pBluescript plasmid as described in the Material and methods section "Subcloning into a sequencing vector". Shotgun cloning usually results in several subclones being isolated, carrying most of the λ clone's various fragments. Thus isolated clones from strain 151 were screened with the α consensus oligonucleotide and those from strain 149 with the strain 151 pucBA genes. Positively hybridising 2.1 kb subclones were identified and prepared for sequencing. The clone carrying the 2.1 kb fragment from strain 151 was named p17T2.1X, to indicate the λ clone from which it was subcloned and the size of the fragment. Similarly the clone carrying clone carrying the 2.1 kb fragment from strain 149 was named pF6X2.

For p17T2.1X (strain 151) an initial sequencing reaction was carried out using the α consensus oligonucleotide, which is homologous to the anti-sense strand of the pucA gene at its 5′ end. Whilst for pF6X2 (strain 149) a primer designed for sequencing the strain 151 puc genes known as UR1C was utilised. UR1C binds to the non-coding strand upstream of the pucB gene. A sequencing reaction with either primer would therefore read the sequence of the pucB gene (assuming that the *Rv. gelatinosus* pucBA genes are organised in a similar manner to those of other photosynthetic bacteria). Furthermore the β-polypeptide from both strains 151 and 149 had been protein sequenced and shown to differ by a single amino acid (see Zuber & Brunisholz, 1991). Thus a sequencing reaction carried out on both cloned 2.1 kb fragments could both confirm the presence of the puc genes and confirm the strain from which they had been cloned.

Figure 34 (opposite) shows part of the sequence obtained from these initial sequencing reactions on the cloned genes from strains 151 and 149, together with the predicted protein translation.

Figure 35

Agarose gel and autoradiograph confirming the source of the strain 151 cloned 2.1 kb fragment.



Note: All samples (excepting λ markers) were digested with the enzyme XhoI, and run on a 0.8% agarose gel (pictured on the left). DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes, the size of the resulting fragments is given on the right of the gel (500 ng of markers were loaded). 20 µg of digested genomic DNA was loaded for each strain, 420 ng of λ17T and two different amounts of the plasmid p17T2.1X (given above the lane). The gel was Southern blotted and the filter probed using the pucBA genes from strain 151, the resulting autoradiograph is shown on the right.

Since the two primers read opposite strands of the DNA the sequences are inverted relative to each other, as is the predicted sequence of the β-polypeptide. The conserved histidine residue of the β-polypeptide can be seen highlighted in grey, whilst the amino acid which differs between the strains is highlighted in black. There is a clear single DNA base-pair difference that gives rise to the amino acid change, the altered residue is highlighted with an arrow in Figure 34 (previous page). This residue forms the start of the codon GTT (on the coding strand, note that the non-coding strand is pictured in Figure 34), which codes for the amino acid valine. In strain 149 the codon has changed to ATT which encodes isoleucine.

Further confirming the source of the cloned fragments

Following sequencing of the strain 151 pucBA genes further confirmation of the source of the 2.1 kb cloned fragment was obtained by way of a hybridisation experiment utilising the strain 151 pucBA genes. Genomic DNA, the λ clone λ17T and plasmid p17T2.1X, were digested with XhoI and run on an agarose gel. This was then blotted and probed with a DNA fragment carrying the pucBA genes from strain 151, this probe was produced as described in the materials and methods section "PCR".

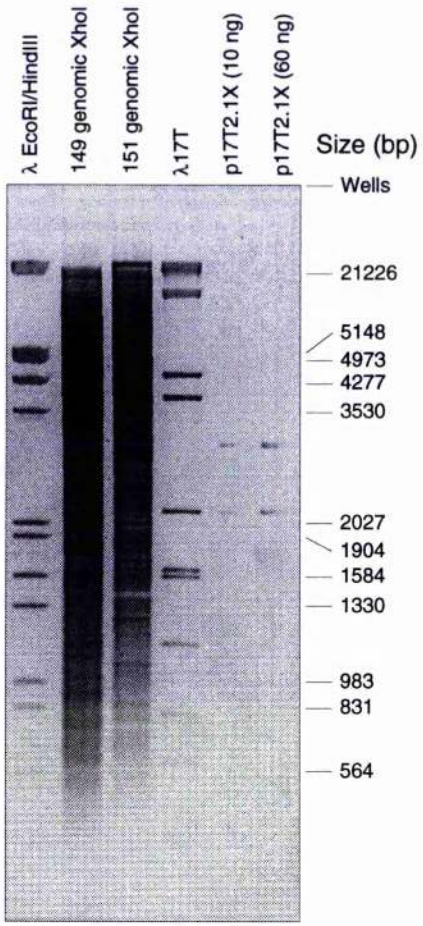
The resulting gel and autoradiograph are presented in Figure 35 (opposite). For purposes of comparison strain 149 genomic DNA was also digested with XhoI (second lane from the left), and run alongside the strain 151 genomic DNA (third lane from the left). There is an apparent difference between the two strains, with the digested strain 151 genomic DNA lying higher up the gel than that of 149. Equimolar amounts of the genomic DNA and p17T2.1X (in the 10 ng lane, fifth from the left) were loaded, whilst twice molar amount of λ17T was loaded. This was to ensure that when the blot was probed the hybridisation level would be fairly even across the different samples. The autoradiograph shown to the right side of Figure 35 shows a clear 2.1 kb hybridising band in each of the strain 151 samples, indicating that these fragments are comparable. In contrast the strain 149 genomic DNA contains a slightly smaller hybridising fragment, suggesting that the 2.1 kb fragment previously identified in strain 149 is not identical to that cloned from strain 151 (an observation that is confirmed by the sequencing studies described below). This difference in size was not observed when the strain 149 λ clones were compared with λ17T (see Figure 32). However the gel pictured in Figure 35 was twice the length (20 cm) of that in Figure 32 and allowed the resolution of the small difference in size between these fragments (the cloned fragments differ by 43 bp).

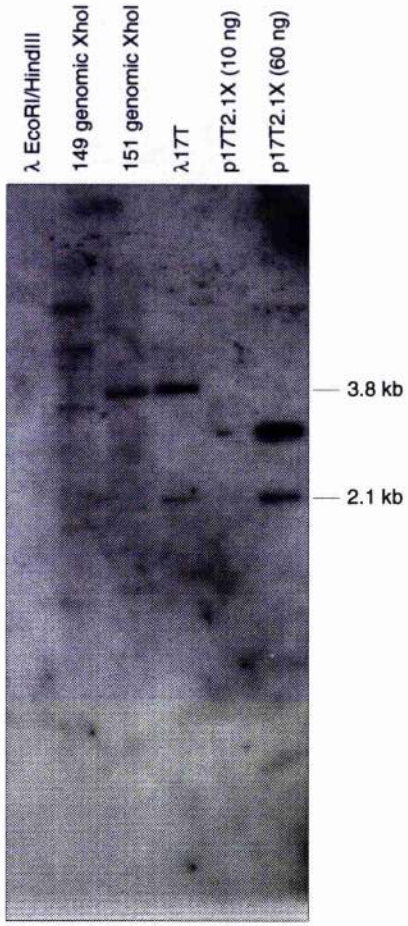
Identification of the pucC gene

Shortly after the hybridisation experiment described above was carried out probe carrying much of the pucC gene from *Rb. sphaeroides* became available. Since pucC was known to be essential for the synthesis of LH2 and expression of the pucBA genes of *Rb. sphaeroides* and *Rb. capsulatus*, it was interesting to ask whether *Rv. gelatinosus* contained the pucC gene, if so whether it lies close to the pucBA genes, and whether there was a difference between the two strains of *Rubrivivax* as regards pucC.

Figure 36

Agarose gel and autoradiograph illustrating *Rv. gelatinosus* genomic DNA hybridised with a probe carrying the *Rb. sphaeroides* pucC gene.





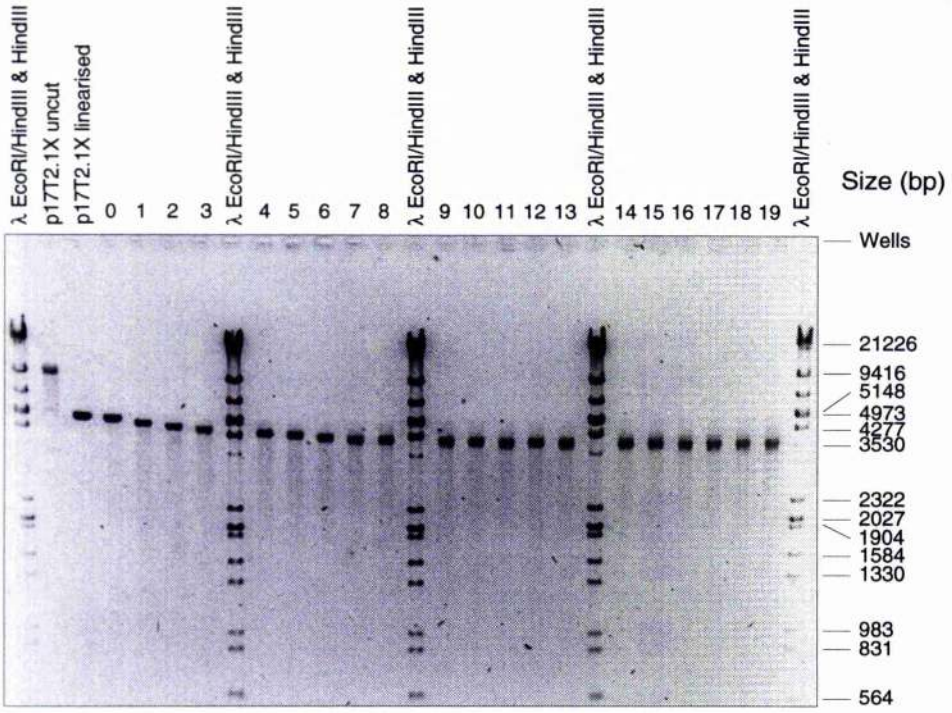
Note: All samples (excepting λ markers) were digested with the enzyme XhoI, and run on a 0.8% agarose gel (pictured on the left). DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes, the size of the resulting fragments is given on the right of the gel (500 ng of markers were loaded). 20 µg of digested genomic DNA was loaded for each strain, 420 ng of λ17T and two different amounts of the plasmid p17T2.1X (given above the lane). The gel was Southern blotted and the filter probed using the pucC gene from *Rb. sphaeroides*, the resulting autoradiograph is shown on the right.

To answer these questions the blot probed with the pucBA genes as described above was washed and reprobed with the *Rb. sphaeroides* pucC gene. The resulting autoradiograph is pictured in Figure 36 (opposite), together with a copy of the gel. The autoradiograph shows several bands which hybridised with the *Rb. sphaeroides* pucC fragment. Interestingly the probe hybridised with the 2.1 kb fragment present in p17T2.1X and λ17T, suggesting that at least a part of the pucC gene lies on this fragment. A more strongly hybridising fragment of approximately 3.8 kb in size is visible in the λ17T clone, an equivalently sized fragment is also seen within the strain 151 genomic DNA. This suggests that the pucC gene is present in strain 151 but has been cut into two parts by XhoI, an assertion supported by the map drawn up from the strain 151 λ clones (Figure 33) which indicates that the 3.8 kb fragment lies adjacent to the 2.1 kb fragment. In contrast with strain 151 there are no strongly hybridising bands present in the genomic DNA of strain 149, rather several fainter bands are seen. This may represent a difference between the two strains, the blot must be interpreted with caution however because the pBluescript vector DNA present in p17T2.1X also hybridises to the probe. It seems likely that this represents some chance homology between the probe and vector DNA, or possibly contamination of the probe with plasmid vector DNA.

The detection of part of the pucC gene in the strain 151 2.1 kb fragment DNA sequence (described below) suggests that the observed pattern of bands hybridising to the pucC probe is essentially correct. To investigate this potential difference further it was decided to sequence the rest of the pucC gene on the 3.8 kb fragment. A pBluescript clone carrying the 3.8 kb fragment from λ17T was isolated during the shotgun cloning of p17T2.1X (and named p17T3.8X), thus this clone was cultured and prepared for DNA sequencing.

Figure 37

Agarose gel illustrating the digestion of plasmid p17T2.1X by Exonuclease III, a step in the construction of nested deletions.



Note: Analysis of the digestion of p17T2.1X by Exonuclease III. Aliquots of the samples taken at each time point (marked above the lanes in minutes) were run on a 1% agarose gel. DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes and HindIII alone then mixing the two digests. The size of the resulting fragments is given on the right of the gel (200 ng of markers were loaded).

Construction of nested deletions

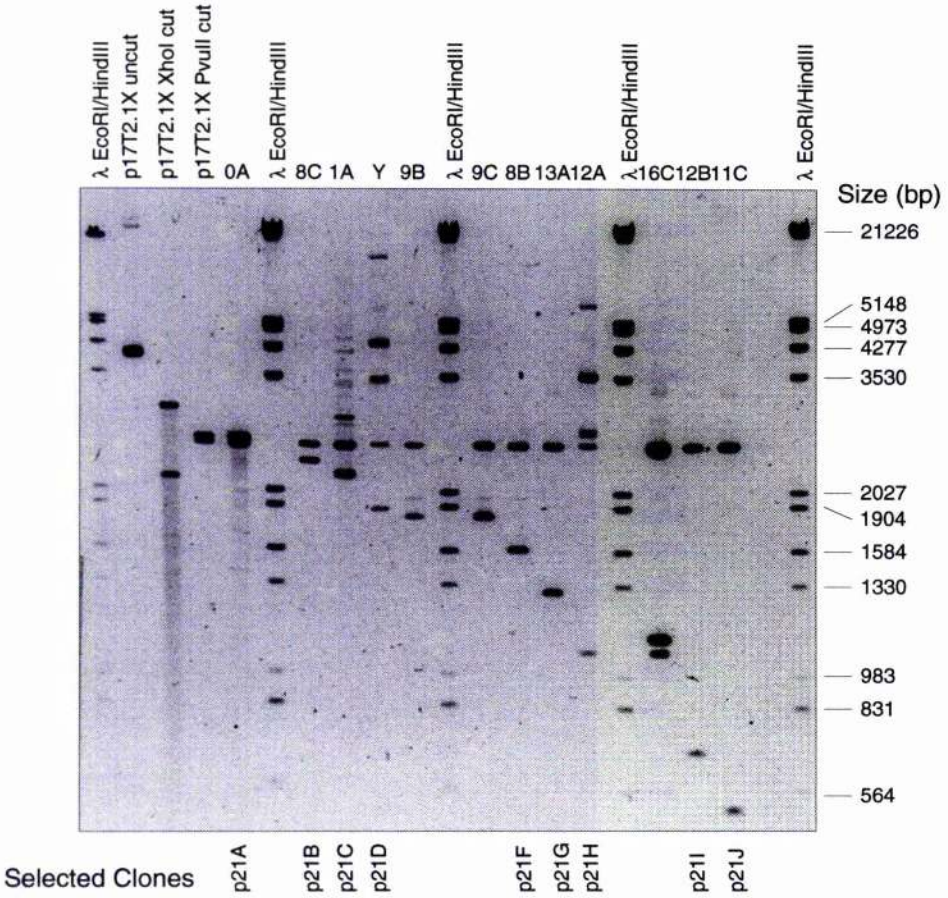
To facilitate the sequencing of the cloned fragments sets of nested deletions were constructed as described in the Materials and methods section "Generation of nested sets of deletions". Deletions were constructed from both p17T2.1X and pF6X2, but not p17T3.8X. It was not possible to construct deletions from p17T3.8X because no suitable restriction enzymes could be found which cut in the vector multiple cloning site (MCS), but not in the cloned insert DNA.

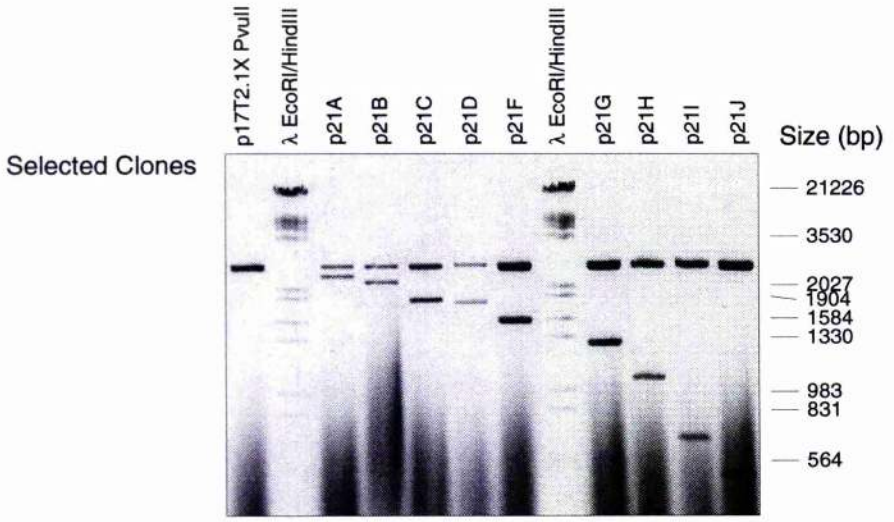
Figure 37 (opposite), illustrates the digestion of p17T2.1X by Exonuclease III. The full sized linearised plasmid can be seen in the lane third from the left. The plasmid decreases in size as the Exonuclease III progresses further into the insert DNA, as can clearly be seen if one looks at the size of the aliquots removed from the reaction over the 19 minute period. Each aliquot was religated following deletion and transformed into *E. coli* DH5α. Several transformants were then isolated from each plated aliquot and mini-preparations of DNA prepared for screening. The top gel pictured in Figure 38 (overleaf) shows several of the deletions isolated from p17T2.1X digested with the enzyme PvuII which cuts in the vector DNA either side of the MCS, allowing the size of the insert to be visualised. Thus one can see that the vector DNA remains the same size in each digested deletion at approximately 2.9 kb, whilst the insert DNA gets progressively smaller as one moves across the gel (the deletions were ordered this way). Nine of the clones pictured on the upper gel were selected with approximately 200 bp difference between each clone. These were then run on a final gel to check their size, this is pictured at the bottom of Figure 38.

Figure 39 (overleaf) pictures a similar set of deletions isolated from the strain 149 clone pF6X2. Once again one can see that the vector DNA remains the same size whilst a progressively smaller range of insert sizes can be seen, the deletions selected for sequencing are marked at the bottom of the gel.

Figure 38

Nested deletions constructed from p17T2.1X to facilitate sequencing of the cloned DNA.

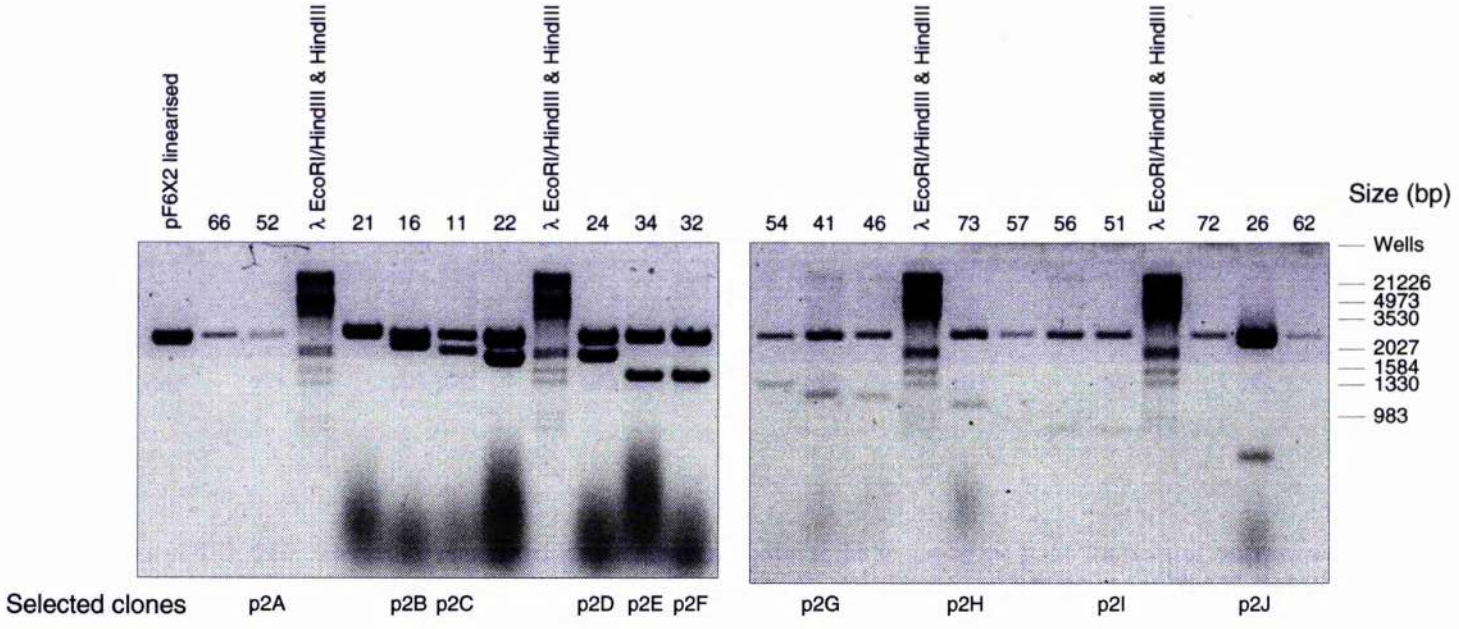




Note: All samples (excepting λ markers, and those samples marked otherwise) were digested with the enzyme PvuII and run on a 1.2% agarose gel. DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes, the size of the resulting fragments is given on the right of the gel (200 ng of markers were loaded). The upper gel shows a range of isolated deletions, a subset were selected for sequencing and renamed as shown at the bottom of the gel. These selected deletions were also run on the lower gel.

Figure 39

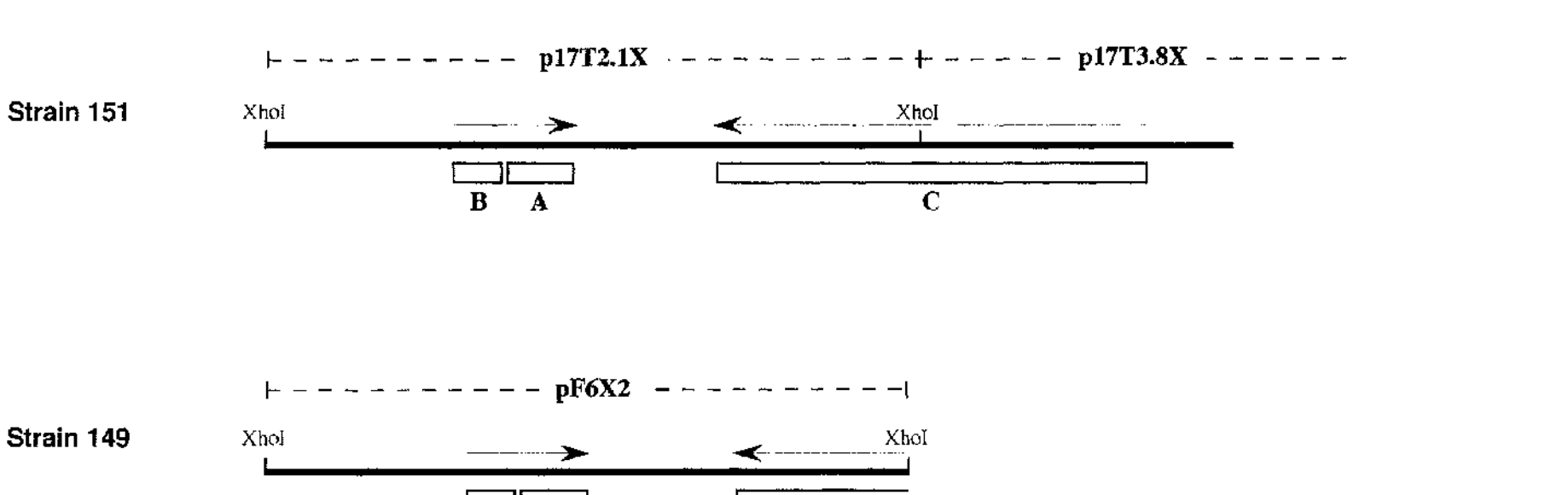
Nested deletions constructed from pF6X2 to facilitate DNA sequencing of the cloned DNA.



Note: Deletions isolated from pF6X2, those selected for sequencing were renamed as shown at the bottom of the gels. The isolated deletions were digested with PvuII and run on a 1.2% agarose gel. DNA markers were loaded, these were made by digesting phage λ DNA with both EcoRI and HindIII enzymes and HindIII alone then mixing the two digests. The size of the resulting fragments is given on the right of the gel (200 ng of markers were loaded). The smears at the lower end of the gel represent RNA present in the sample, no attempt was made to remove this from the 'mini-prep' samples.

Figure 40

A schematic view of the sequenced DNA fragments from *Rv. gelatinosus* strains 151 and 149.



Note: The diagrams above illustrate the extent of the DNA sequence information obtained from the two strains. Two fragments were sequenced from strain 151, p17T2.1X (2154 bp) was sequenced in its entirety, whilst only 1018 bp of the p17T3.8X insert were sequenced (its full size is about 3.8 kb). These two fragments together encompass the pucBA genes and the pucC gene. Only a single fragment was sequenced from strain 149, the clone pF6X2 carries the 2093 bp insert which was fully sequenced. This contains the strain 149 pucBA genes but only a portion of the strain 149 pucC gene. The positions of the predicted open reading frames for the pucBA and pucC genes are labelled and shown by open boxes, arrows indicate the direction in which transcription of the genes should occur. The XhoI sites delineating the cloned fragments are marked by vertical lines and labelled.

Sequencing

The deletions created from the two plasmids p17T2.1X and pF6X2 allowed one strand of the cloned DNA to be sequenced almost entirely from the T3 primer site of pBluescript. At three points (once in p17T2.1X and twice in pF6X2) the distance between the deletions was greater than estimated and a primer was synthesised to bridge these gaps. The opposing strand of DNA was then sequenced using synthesised oligonucleotide primers.

Since it was not possible to create a set of nested deletions from p17T3.8X a portion of this clone was sequenced by 'primer walking'. Initial reactions were carried out using the T3 and T7 primers sites of pBluescript to determine the end of the insert which was homologous to the pucC gene of the *Rhodobacter* species, this appeared to be the T3 end. Sequencing was then continued from this primer (and on the opposing strand back towards the T3 end) by synthesising a series of primers until the sequence obtained for both strands extended well beyond the apparent start of the pucC gene.

The sequence information recorded is given on the following pages and is summarised in Figure 40 (opposite). Figure 40 illustrates the extent of sequence information for the three cloned fragments, and shows the relative positions of the puc genes. The sequenced sizes of the approximately 2.1 kb fragments from the two strains were found to be 2131 bp for strain 151 and 2091 bp for strain 149. The smaller size of the strain 149 fragment is due in part to a mutation which moves the XhoI site within pucC ~100 bp closer to pucBA. The reason that the size difference is only 40 bp is that there have been many sequence changes, including deletions or insertions, between the pucA and pucC genes and upstream of pucB (these changes will be discussed in more detail in the next results section). Only 1018 bp of the p17T3.8X clone were sequenced since this was enough to encompass the rest of the pucC gene. Surprisingly the orientation of the *Rv. gelatinosus* pucC gene is opposite that of the pucBA genes, an arrangement not observed in any other species in which the gene has been sequenced.

The full DNA sequence from the two strain 151 clones p17T2.1X and p17T3.8X, and the strain 149 clone pF6X2 is presented on the following pages. In Figure 41 the clones p17T2.1X and p17T3.8X are presented as a single fragment since they represent a contiguous fragment of the strain 151 genome, the XhoI sites delineating the two clones are however marked. Also marked are the start and stop codons of the pucBA and pucC genes. Figure 42 displays the sequence of the strain 149 clone pF6X2, which also has the XhoI sites and puc genes marked.

Figure 41

The DNA sequence obtained from the *Rv. gelatinosus* strain 151 clones p17T2.1X and p17T3.8X. The positions of the XhoI sites which delineate the clones are highlighted. Also marked are the start and stop codons of the pucBA and C genes.

```
                ▼ XhoI
                .         .         .         .         .
   1  CTCGAGCAGCAGGTGCCGGCGCTGCTGGGCTATCAGTGGAAGGTGCGCGA   50
      GAGCTCGTCGTCCACGGCCGCGACGACCCGATAGTCACCTTCCACGCGCT
           ▲
                .         .         .         .         .
  51  CGAGCGCGCGTTCAACTTCGCGCGCCTGTTCTACCGGGCGCTGTTCGAAC  100
      GCTCGCGCGCAAGTTGAAGCGCGCGGACAAGATGGCCCGCGACAAGCTTG
                .         .         .         .         .
 101  GCGGCGCGCCGTCGTACCGCTACCTGGAATACGCGTTCATGCGGGCCCGG  150
      CGCCGCGCGGCAGCATGGCGATGGACCTTATGCGCAAGTACGCCCGGGCC
                .         .         .         .         .
 151  CGCCTGGCCTACGACGAGGCGGTGAACGAGGCCCAGGCTCGGCTGCTGGC  200
      GCGGACCGGATGCTGCTCCGCCACTTGCTCCGGGTCCGAGCCGACGACCG
                .         .         .         .         .
 201  CGACGGCGGTGACTGCGCCGCCGAGATCGAGGACGCGATGCGCGACCACT  250
      GCTGCCGCCACTGACGCGGCGGCTCTAGCTCCTGCGCTACGCGCTGGTGA
                .         .         .         .         .
 251  CCTGGATCTCGCCGGTGCTCGTCATGCAGATGGACTGAGCACCCCCTTGT  300
      GGACCTAGAGCGGCCACGAGCAGTACGTCTACCTGACTCGTGGGGGAACA
                .         .         .         .         .
 301  TCAGGGGGTCACTCGACATCGATTCGAAGCCCGACGCCGTGGGAAGCGGC  350
      AGTCCCCCAGTGAGCTGTAGCTAAGCTTCGGGCTGCGGCACCCTTCGCCG
                .         .         .         .         .
 351  CGCCTGCCGGGCGGCGGCGGGCGGCGAAAGACGCTGCGCACCCGGCGTCA  400
      GCGGACGGCCCGCCGCCGCCCGCCGCTTTCTGCGACGCGTGGGCCGCAGT
                .         .         .         .         .
 401  GGACGCCTGGCGCGGCCGGCAGACCTGAGCACTTTGTCATAGATCAGTCT  450
      CCTGCGGACCGCGCCGGCCGTCTGGACTCGTGAAACAGTATCTAGTCAGA
                .         .         .         .         .
 451  CGCGACACATTCAGCGTTGCGGGGGTGTCACGGCATGTTGACGCCCCCAC  500
      GCGCTGTGTAAGTCGCAACGCCCCCACAGTGCCGTACAACTGCGGGGGTG
                .         .         .         .         .
 501  CCTGTCAGCGTAGCGTTACACATGTCGGTAGGGAGAAATCCGAGGGCTAC  550
      GGACAGTCGCATCGCAATGTGTACAGCCATCCCTCTTTAGGCTCCCGATG
```

Figure 41 (continued)

```
551  AGAAGCCAACCCACGTGCGGGTCGAGCACGGAAACTGGTCCATGTCTCTG  600
     TCTTCGGTTGGGTGCACGCCCAGCTCGTGCCTTTGACCAGGTACAGAGAC

                pucB
     r.b.s      Start
601  AAGGAGAAACGCAATGGCAGATGATGCAAACAAGGTTTGGCCGTCCGGCC  650
     TTCCTCTTTGCGTTACCGTCTACTACGTTTGTTCCAAACCGGCAGGCCGG

651  TGACCACGGCTGAGGCCGAAGAGCTGCAAAAGGGTCTGGTTGATGGCACC  700
     ACTGGTGCCGACTCCGGCTTCTCGACGTTTTCCCAGACCAACTACCGTGG

701  CGTGTTTTCGGTGTTATCGCCGTTCTGGCACACATCCTGGCGTATGCCTA  750
     GCACAAAAGCCACAATAGCGGCAAGACCGTGTGTAGGACCGCATACGGAT

                                          pucA
                stop          r.b.s       Start
751  TACGCCGTGGCTCCACTAAGTCACTAGCAGGAACTAAAAAATGAACCAAG  800
     ATGCGGCACCGAGGTGATTCAGTGATCGTCCTTGATTTTTTACTTGGTTC

801  GCAAAGTCTGGCGCGTCGTTAAGCCGACCGTTGGTGTTCCCGTTTACCTG  850
     CGTTTCAGACCGCGCAGCAATTCGGCTGGCAACCACAAGGGCAAATGGAC

851  GGCGCCGTGGCCGTCACGGCCCTGATCCTGCACGGCGGCCTGCTGGCCAA  900
     CCGCGGCACCGGCAGTGCCGGGACTAGGACGTGCCGCCGGACGACCGGTT

901  GACCGACTGGTTCGGTGCCTACTGGAACGGTGGCAAGAAGGCTGCTGCGG  950
     CTGGCTGACCAAGCCACGGATGACCTTGCCACCGTTCTTCCGACGACGCC

951  CTGCCGCCGCCGTCGCCCCGGCCCCGGTCGCGGCCCCGCAGGCTCCGGCG  1000
     GACGGCGGCGGCAGCGGGGCCGGGGCCAGCGCCGGGGCGTCCGAGGCCGC

     stop
1001 CAGTAAATAGCGTCTGTCAGCGTATGTCCGGCATGGGGCTGGGGTTCGCG  1050
     GTCATTTATCGCAGACAGTCGCATACAGGCCGTACCCCGACCCCAAGCGC

1051 GACCCCGGCCCCATGTAGTTTCAAGGGCGGTGCAGTGCCTCTGGCTTCTG  1100
     CTGGGGCCGGGGTACATCAAAGTTCCCGCCACGTCACGGAGACCGAAGAC

1101 TGAAAAAGCGCCGCCACAAGACCCGGACACGTCCGAACCGCTGACGCGGC  1150
     ACTTTTTCGCGGCGGTGTTCTGGGCCTGTGCAGGCTTGGCGACTGCGCCG
```

Figure 41 (continued)

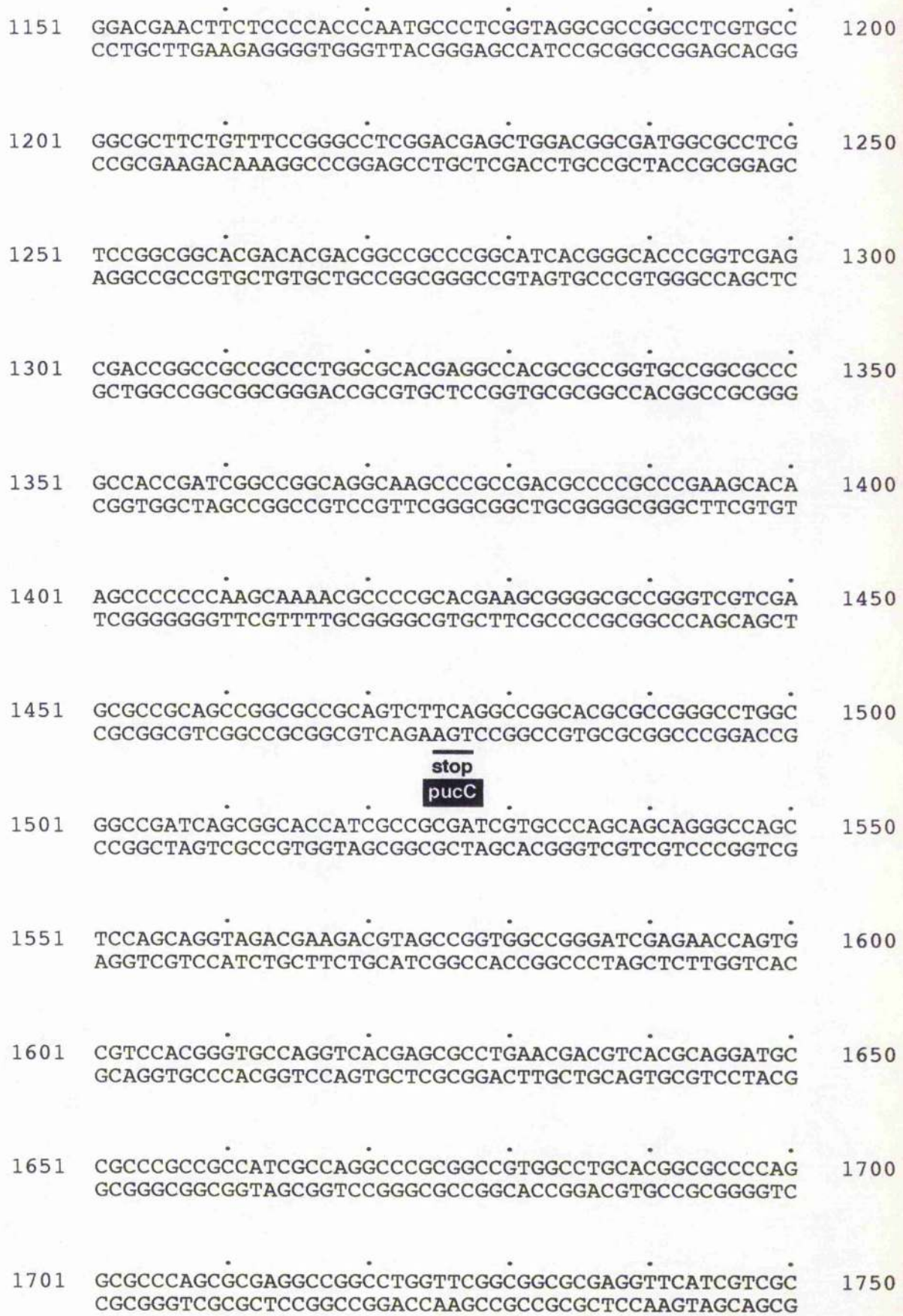


Figure 41 (continued)

```
1751  CGTCAGCGTGCCGTGCCCGAACAGGCCGCCGCCGAAGCCGATCAGCGCCA  1800
      GCAGTCGCACGGCACGGGCTTGTCCGGCGGCGGCTTCGGCTAGTCGCGGT

1801  CGCCGGTGGCGAAGACGAAGGGCGAGTTCATCGGTGCGGCGATGATGACG  1850
      GCGGCCACCGCTTCTGCTTCCCGCTCAAGTAGCCACGCCGCTACTACTGC

1851  AAGATGAAGGCCGGGATGCCGACCAGCGCGCCCCAGCAGGCCATGCGGAA  1900
      TTCTACTTCCGGCCCTACGGCTGGTCGCGCGGGGTCGTCCGGTACGCCTT

1901  CGGGTCCATGCCGCGGCTCAGCACCTTGGACGCCAGGCCGAAACCGAACA  1950
      GCCCAGGTACGGCGCCGAGTCGTGGAACCTGCGGTCCGGCTTTGGCTTGT

1951  GGCCCCCCAGCGCCAGCGTCGCCGTCAGCTTGGTCGTGTCGCCGACCGTC  2000
      CCGGGGGGTCGCGGTCGCAGCGGCAGTCGAACCAGCACAGCGGCTGGCAG

2001  AGCGCCAGCACCTCGCCGCCATAGGGTTCGAGCAGCACGTCCTCCATCGT  2050
      TCGCGGTCGTGGAGCGGCGGTATCCCAAGCTCGTCGTGCAGGAGGTAGCA

2051  GAAGGCCGCCGTGCCCAGGCCGATCGTCAGCAGCCGGCGCACGACCTGGT  2100
      CTTCCGGCGGCACGGGTCCGGCTAGCAGTCGTCGGCCGCGTGCTGGACCA

                                  XhoI
2101  TGCCGTTGTCGACGAAGTGACGCCAGGCCTCGAGGAAGGACGGATCGTCG  2150
      ACGGCAACAGCTGCTTCACTGCGGTCCGGAGCTCCTTCCTGCCTAGCAGC

2151  TTGCGCGGCTGCGCTCCGCGACGCGGGTGCCGGCTCTCCTGCTTCCACAG  2200
      AACGCGCCGACGCGAGGCGCTGCGCCCACGGCCGAGAGGACGAAGGTGTC

2201  CGCGATGACGTTGAGGATCATCGTCGTCAGCGCCGCGCCCTGGATCACCT  2250
      GCGCTACTGCAACTCCTAGTAGCAGCAGTCGCGGCGCGGGACCTAGTGGA

2251  GGATCAGCTTGGCCGGCGTGAAGTCGGCCAGCAGCCAGCCGAAGAAGAAG  2300
      CCTAGTCGAACCGGCCGCACTTCAGCCGGTCGTCGGTCGGCTTCTTCTTC

2301  GCGCTGACGATGGTGCCGACCAGCAGCATCACGTACATCAGGCCGACGAC  2350
      CGCGACTGCTACCACGGCTGGTCGTCGTAGTGCATGTAGTCCGGCTGCTG
```

Figure 41 (continued)

```
2351  CTTCGGCCGCGCTTCGGGCTCGGCGAGGTCGGTGGCCAGCGCCAGGCCCA  2400
      GAAGCCGGCGCGAAGCCCGAGCCGCTCCAGCCACCGGTCGCGGTCCGGGT

2401  CCGTCTGGACGGTGTGCATGCCGGCACCGACGAGCAGAAAGGCGATGGCC  2450
      GGCAGACCTGCCACACGTACGGCCGTGGCTGCTCGTCTTTCCGCTACCGG

2451  GCGCCGAGCTGGCCGACCCAGGTGGGCAACTGGCCCGACTCGCCGCCGCC  2500
      CGCGGCTCGACCGGCTGGGTCCACCCGTTGACCGGGCTGAGCGGCGGCGG

2501  GGACAGCACCAGCAGCGCGAACGGCATCATCGCGAAGCCGCCGAACTGCA  2550
      CCTGTCGTGGTCGTCGCGCTTGCCGTAGTAGCGCTTCGGCGGCTTGACGT

2551  GCATCGTGCCCTTCCAGATGTAGGGCACGCGGCGCCAGCCCAGCGCCGAC  2600
      CGTAGCACGGGAAGGTCTACATCCCGTGCGCCGCGGTCGGGTCGCGGCTG

2601  TGGTGGTTGTCGGACTTGAAGCCGATCAGCGCCCGGAACGGCGCGAACAG  2650
      ACCACCAACAGCCTGAACTTCGGCTAGTCGCGGGCCTTGCCGCGCTTGTC

2651  CAGCGGCAGCGCGACCATCACCGCGACGATCGTCGCCGGCACGGCGAGCT  2700
      GTCGCCGTCGCGCTGGTAGTGGCGCTGCTAGCAGCGGCCGTGCCGCTCGA

2701  CGACGATCATCACCCGGTTGAGCGTGCCGACGAGCAGCGTCACCGCCATG  2750
      GCTGCTAGTAGTGGGCCAACTCGCACGGCTGCTCGTCGCAGTGGCGGTAC

2751  CCGACCGTGATCTGGAACAGCGACAGCCGCAGCAGCCGCGACAGCGGCAG  2800
      GGCTGGCACTAGACCTTGTCGCTGTCGGCGTCGTCGGCGCTGTCGCCGTC

2801  CTCGGGCGTCGCCGCGTCGGCGAAGGGCAGGTAGCGGGTGCCGAACCCCG  2850
      GAGCCCGCAGCGGCGCAGCCGCTTCCCGTCCATCGCCCACGGCTTGGGGC

2851  TCCAGACCTGCATCAGTTTGCGGCCAATCGTGCTCATCGGGTGACATGCT  2900
      AGGTCTGGACGTAGTCAAACGCCGGTTAGCACGAGTAGCCCACTGTACGA
                                  Start          r.b.s
                                  pucC

2901  CCGCGGCAGGTGGCCGGCATGGACAAACGCCGGAACCTAACCCGGAAGCC  2950
      GGCGCCGTCCACCGGCCGTACCTGTTTGCGGCCTTGGATTGGGCCTTCGG
r.b.s
```

Figure 41 (continued)

```
2951  GGCTGTCGAGCCTGCTTTTTGTCAAATCCAAACGGTGCCGCCACGCACCG  3000
      CCGACAGCTCGGACGAAAAACAGTTTAGGTTTGCCACGGCGGTGCGTGGC

3001  CGATGACGCCGGTCAGGATCCGCCACCCGGCGCACGACGCCCGGCGGGCA  3050
      GCTACTGCGGCCAGTCCTAGGCGGTGGGCCGCGTGCTGCGGGCCGCCCGT

3051  CAAGGCTTGCCTTGCGGGGCCGTCCCTGCCTGCGAGATCTCCCCCTGTCG  3100
      GTTCCGAACGGAACGCCCCGGCAGGGACGGACGCTCTAGAGGGGGACAGC

3101  ACCCCGCCCCCATCCGACGCCGCCCCCGCGCCCCGCCCTGCCGCCTGGCG  3150
      TGGGGCGGGGGTAGGCTGCGGCGGGGGCGCGGGGCGGGACGGCGGACCGC

3151  A
      T
```

Note: The start and stop codons of the pucB, pucA and pucC genes are marked, as are the potential ribosome binding sites lying just upstream of the start codons (these are labelled r.b.s.). XhoI restriction sites are marked, the recognition site by a horizontal line and the cut points by and arrows. These sites serve to delineate the sequence obtained from the two strain 151 clones. The p17T2.1X sequence extends from bp 1 to the XhoI site at 2130 bp, p17T3.8X sequence extends from the XhoI site to the end of the sequence shown. The left-hand end of each clones sequence lies at the T3 end of the pBluescript MCS (see Figure 17 in the Materials and methods), the right-hand end of p17T2.1X (at 2130 bp) thus lies at the T7 end of the MCS.

Figure 42

The DNA sequence obtained from the *Rv. gelatinosus* strain 149 clone pF6X2. The positions of the XhoI sites which delineate the clone are highlighted. Also marked are the start and stop codons of the pucBA genes and the stop codon of the pucC gene.

```
      ▼ XhoI
   1  CTCGAGCAGCAGGTGCCGGCACTGCTGGGCTACCAGTGGAAGGTTTCGGA   50
      GAGCTCGTCGTCCACGGCCGTGACGACCCGATGGTCACCTTCCAAAGCCT
          ▲

  51  CGAGCGCGCCGGCAATTTCGCGCGCCTGTTCTACCGCGCGCTGTTCGAAC  100
      GCTCGCGCGGCCGTTAAAGCGCGCGGACAAGATGGCGCGCGACAAGCTTG

 101  CCGGCACGCCGTCCTACCGATACCTCGAATACGCCTTCATGCGGGCACGG  150
      GGCCGTGCGGCAGGATGGCTATGGAGCTTATGCGGAAGTACGCCCGTGCC

 151  CGCCAGGCTTACGAGCAGGCGGTGCAGGAAGCCGAGGCGCGGCTGGCCAT  200
      GCGGTCCGAATGCTCGTCCGCCACGTCCTTCGGCTCCGCGCCGACCGGTA

 201  CGCCGACGCCGGGGCGCCGGCCGTCAGCGTCGACGATTCTCTGCGCGACC  250
      GCGGCTGCGGCCCCGCGGCCGGCAGTCGCAGCTGCTAAGAGACGCGCTGG

 251  ACTCCTGGATCTCGCCCGTGCTGGTGATGCAGATGGAATGAGGCCGCGGC  300
      TGAGGACCTAGAGCGGGCACGACCACTACGTCTACCTTACTCCGGCGCCG

 301  GACGGCACGCCGTGGTGTGACACCCACGCGCAGAACCCCCCTGTTCCAGG  350
      CTGCCGTGCGGCACCACACTGTGGGTGCGCGTCTTGGGGGGACAAGGTCC

 351  GGTGCGGGCACGCCAACGTGCTCAGCCTCGCGGCGGGCGCAGGATGGACG  400
      CCACGCCCGTGCGGTTGCACGAGTCGGAGCGCCGCCCGCGTCCTACCTGC

 401  TCAAGCGCCCGGCCTCCCTAGGAACACAGCAAGTGCTCACAGATGAGCCG  450
      AGTTCGCGGGCCGGAGGGATCCTTGTGTCGTTCACGAGTGTCTACTCGGC

 451  CGAGCGGCCGACTTGTCCGCCAATTTGTCATAGATCAGTTCCCGGGGAGA  500
      GCTCGCCGGCTGAACAGGCGGTTAAACAGTATCTAGTCAAGGGCCCCTCT

 501  TTTCACGTTGCCGGGGTGTCACGGCATGTTGACGCACCCCCCCTGTCAGC  550
      AAAGTGCAACGGCCCCACAGTGCCGTACAACTGCGTGGGGGGGACAGTCG
```

Figure 42 (continued)

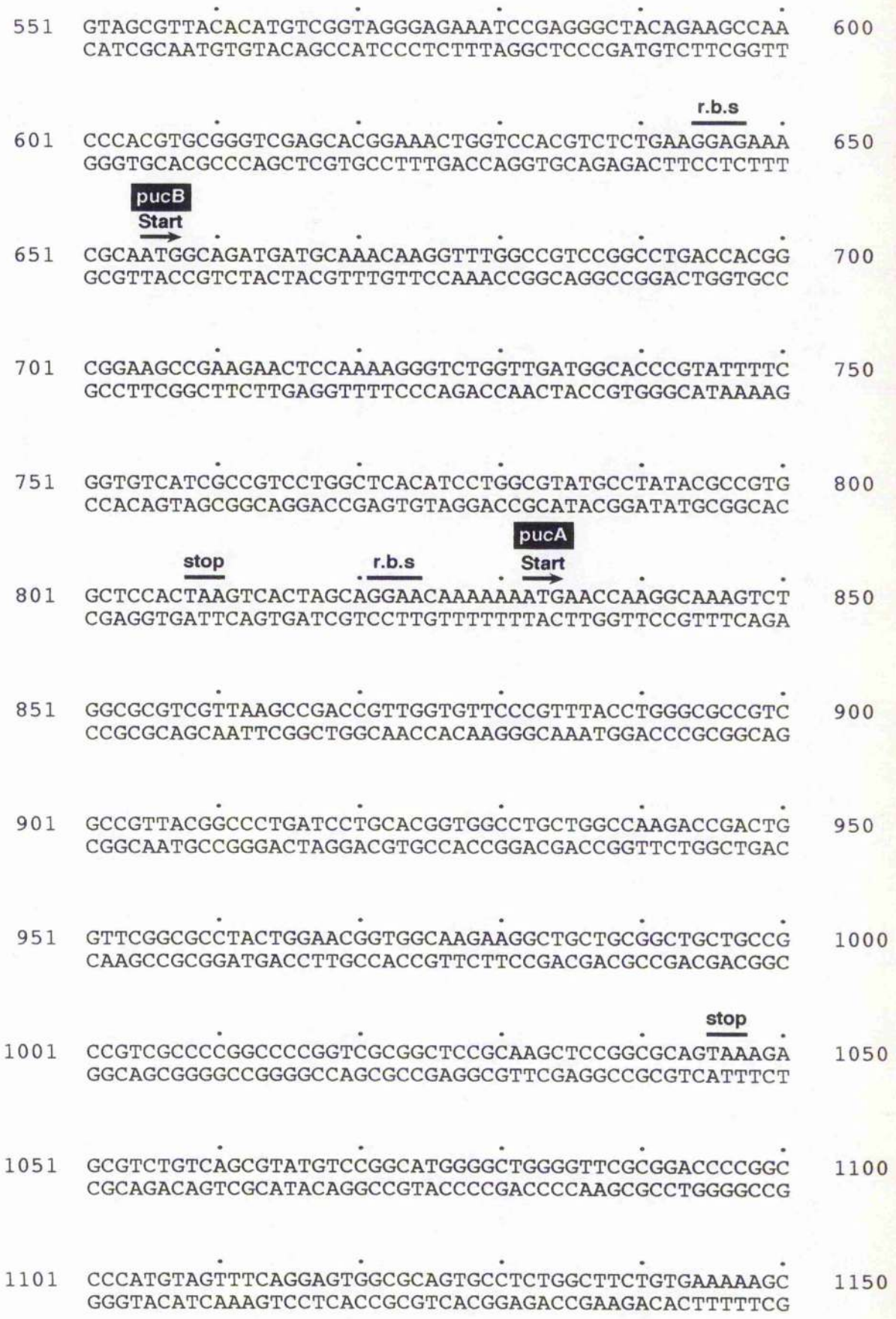


Figure 42 (continued)

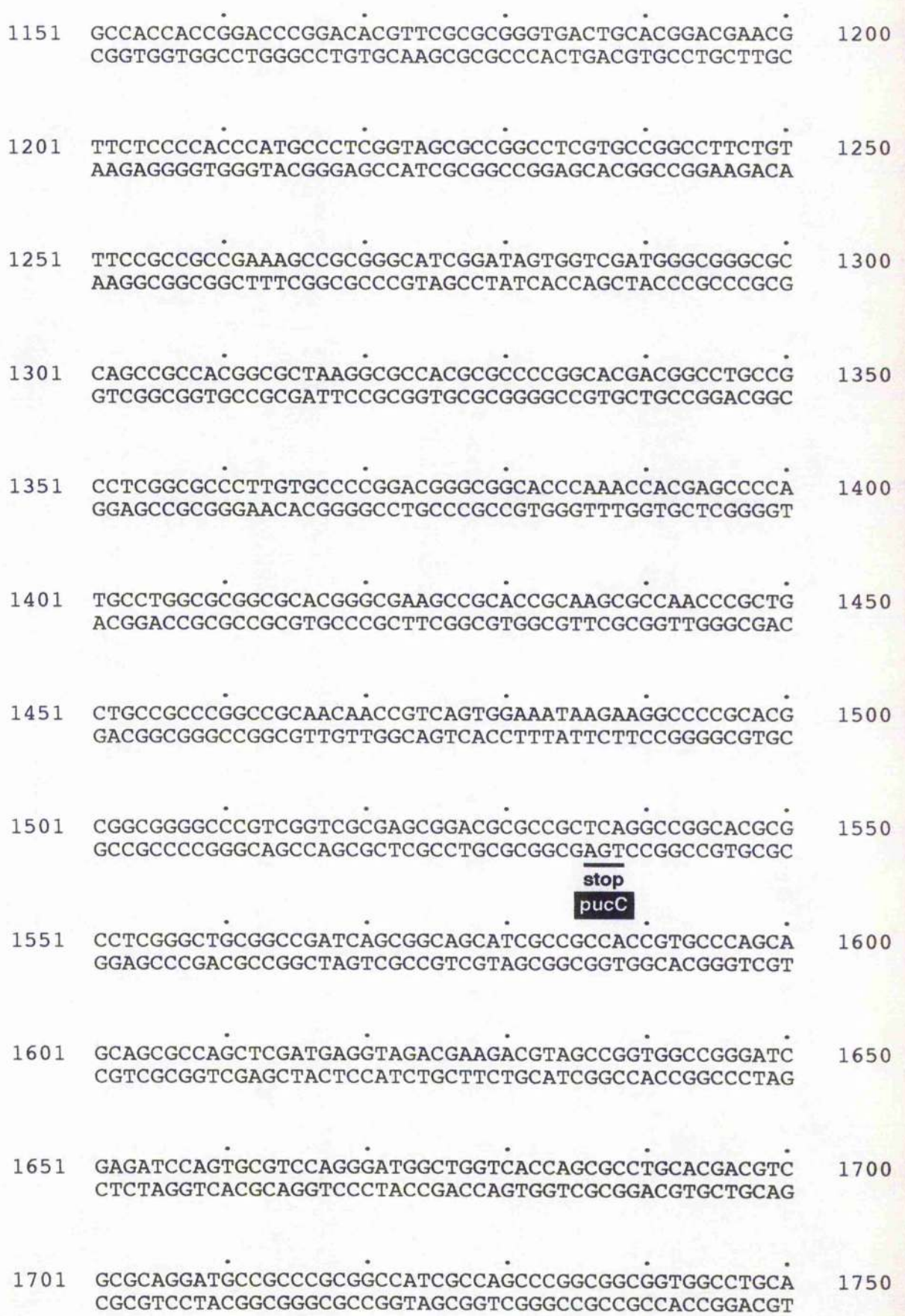


Figure 42 (continued)

```
1751  CGGCGCCCCAGGCGCCCAGCGCCAGGCCGGCCTGATTCGAAGGCGCGAGG  1800
      GCCGCGGGGTCCGCGGGTCGCGGTCCGGCCGGACTAAGCTTCCGCGCTCC

1801  TTCATCGTCGCCGTCAGCGTGCCGTGGCCGAACAGCCCGCCGCCGAAACC  1850
      AAGTAGCAGCGGCAGTCGCACGGCACCGGCTTGTCGGGCGGCGGCTTTGG

1851  GATCAGCGCCACGCCGGTGGCGAAGACGAAGGGCGAGTTCATCGGCGCCG  1900
      CTAGTCGCGGTGCGGCCACCGCTTCTGCTTCCCGCTCAAGTAGCCGCGGC

1901  CGATGATGACGAAGAAAAAGGCCGGGATGCCGACGAGCGCGCCGATGCAG  1950
      GCTACTACTGCTTCTTTTTCCGGCCCTACGGCTGCTCGCGCGGCTACGTC

1951  GCCATGCGGAACGGGTCCATGCCGCGGCTCAGCACCTTGGACGCCAGGCC  2000
      CGGTACGCCTTGCCCAGGTACGGCGCCGAGTCGTGGAACCTGCGGTCCGG

2001  GAAGCCGAACAGCCCGCCCAGCGCCAGCGTCGCCGTCAGCTTGGTCGTGT  2050
      CTTCGGCTTGTCGGGCGGGTCGCGGTCGCAGCGGCAGTCGAACCAGCACA

                                          ↓ XhoI
2051  CGCCGACGGTGAGCCCGAGGATCTGGCCGCCATAAGGCTCGAG  2093
      GCGGCTGCCACTCGGGCTCCTAGACCGGCGGTATTCCGAGCTC
                                                ↑
```

Note: The start and stop codons of the pucB and pucA genes are marked, as are the potential ribosome binding sites lying just upstream of the start codons (these are labelled r.b.s.). The pucC stop codon is also marked (only the 3′ end of this gene is present on this fragment). XhoI restriction sites are marked, the recognition site by a horizontal line and the cut points by and arrows. The left-hand end of the sequence lies at the T3 end of the MCS in pBluescript.

Conclusions

The previous results section detailed the construction of genomic libraries for strains 151 and 149 of *Rv. gelatinosus*, and the isolation of λ clones which hybridised with an oligonucleotide probe homologous to the pucA gene. This section described the successful subcloning and sequencing of an approximately 2.1 kb fragment from each strain. DNA sequencing studies confirmed that the pucBA genes were present on these fragments as expected, eliminating the possibility that the α consensus oligonucleotide was hybridising to something other than the pucA gene. The predicted amino acid sequence of the β-polypeptide was also shown to match that of the protein sequenced β-polypeptide from each strain, and since these differ by one amino acid this gave confirmation of the source of the cloned DNA. Having shown that the two cloned 2.1 kb fragments were isolated from the correct strains Southern analysis was carried out to confirm that the cloned fragments were representative of those in the original genome. The clearly hybridising 2.1 kb fragment visible in strain 151 genomic DNA (see Figure 35) matches that seen in the λ17T and p17T2.1X clones, suggesting that the cloned DNA corresponds to that present in the strain 151 genome. A similar experiment was not carried out for strain 149, the genomic DNA was however run and probed alongside that of strain 151. This showed a hybridising fragment slightly smaller than that of strain 151, since the sequenced fragment is 40 bp shorter it would appear to match that present in the genomic DNA. In further support are the strain 149 λ clones which form an overlapping set each containing the approximately 2.1 kb fragment (apart from one clone in which it is truncated). If there had been some change during the cloning process, such that the cloned stain 149 DNA was no longer representative of the genomic DNA, it is unlikely that all the clones would have remained so similar.

Interesting results were obtained when genomic DNA from the two strains of *Rv. gelatinosus* was probed with a fragment containing most of the *Rb. sphaeroides* pucC gene. This indicated that strain 151 contained the pucC gene and that it lay close to the pucBA genes as is the case for *Rb. sphaeroides* and *Rb. capsulatus*. It also indicated that much of the gene was situated on the 3.8 kb fragment lying adjacent to the 2.1 kb fragment in the strain 151 clone λ17T. For strain 149 a different pattern was observed, with several faintly hybridising bands apparent. It is possible that the multiple hybridising bands of 149 genomic DNA may represent binding of the probe to pucC homologues. At least two such genes have been identified in the *Rb. capsulatus* PGC known as orf477 and orf428, these show a degree of homology to the pucC gene of *Rb. capsulatus* (see the introduction for more details). Thus these bands may indicate the presence of such homologues in strain 149, however it is difficult to explain why these bands don't also appear in strain 151 genomic DNA, since this strain would presumably also posses these genes.

The difference between the pucC hybridisation patterns observed for the two strains may represent a difference between them, a difference which may explain the differential LH2 expression observed in strains 151 and 149. However the results obtained by probing with the *Rb. sphaeroides* pucC gene are somewhat confusing, and conflict to some extent with the DNA sequencing studies. The latter studies indicate

that the portion of the *Rv. gelatinosus* pucC located on the 2.1 kb fragments is almost identical in each strain, suggesting that the pucC gene is present in the same locus in both strain 149 and 151. However this does not exclude the possibility of a lesion in the strain 149 pucC gene further downstream from the pucBA genes.

In order to clarify the results obtained with the *Rb. sphaeroides* pucC probe it was decided that the strain 151 pucC gene should be isolated for use as a hybridisation probe. The isolation of the strain 151 pucC gene would both allow a more stringent hybridisation, and should provide a stronger hybridisation signal (because of the higher homology). A hybridisation experiment utilising the strain 151 pucC gene could thus clarify the results obtained with the *Rb. sphaeroides* pucC gene, and allow confirmation or elimination of the possibility that strain 149's apparent difference as regards pucC is responsible for its lower synthesis of LH2.

Results Part 4: Sequence Analysis

Introduction

This section of the results consists of analysis carried out on the sequenced DNA and predicted protein sequences extracted from the DNA. In particular comparisons are made between the two strains of *Rv. gelatinosus* studied and other species. I will first describe the DNA sequence and the features found within it, the open reading frames and potential transcription control elements, before moving onto the predicted protein sequences.

Results

DNA sequence analysis

Base composition

The sequenced DNA was analysed using the GCG Composition program to examine how the the base composition of the DNA compared with that measured by Willems *et al*, (1991), who found the two strains to be approximately 71% GC. The 3324 bp of sequenced strain 151 DNA were found to be 70% GC, whilst the 2093 bp of strain 149 sequence were 68% GC. The discrepancy between these figures and that obtained by Willems *et al* may reflect a local variation in the base composition around the puc operon. For instance, if one calculates the GC percentage of the 2032 bp strain 151 segment homologous to the sequence from strain 149, a slightly lower figure of 69% GC is found. It may however reflect differences between the strains used, particularly because there also appears to be a difference between strain 151 and strain 149 genomic DNA digested with XhoI. This can be seen in Figure 35 (opposite page 117), where the genomic DNA of strain 149 is situated lower on the gel and may thus of been cut more often by XhoI. This difference in the cut frequency could be due to a slight difference in the GC content of the genomic DNA.

Open reading frames present in the cloned fragments

The previous results sections have shown that the pucBA and pucC genes are present on the cloned fragments. The pucB and A reading frames were easily identified, because the DNA sequencing began within them and the protein sequence was known (the predicted protein sequence is described in the protein analysis section below). The open reading frame representing pucC was identified by comparing the *Rb. sphaeroides* pucC DNA sequence to the strain 151 DNA sequence. Having identified these expected genes it was also necessary to ask whether any other reading frames were present on the cloned fragments. To answer this question the strain 149 and 151 sequences were analysed using 'Testcode', a part of the GCG package which uses Fickett's Testcode statistic to plot a measure of the non-randomness of every third

base (see Fickett, 1982). This is based on an analysis of a large number of coding and non-coding DNA sequences, and was determined empirically to distinguish between the two types of sequence. The resulting plots can be seen in Figure 43 and 44 (overleaf). Testcode was run using a window size of 200 bp, which is thought to correctly predict 95% of ORFs. The upper and lower horizontal lines represent cut off points, above the top line the score is significant for coding regions whilst below the bottom line the score is significant for non-coding regions. The central region represents sequence for which no accurate prediction can be made.

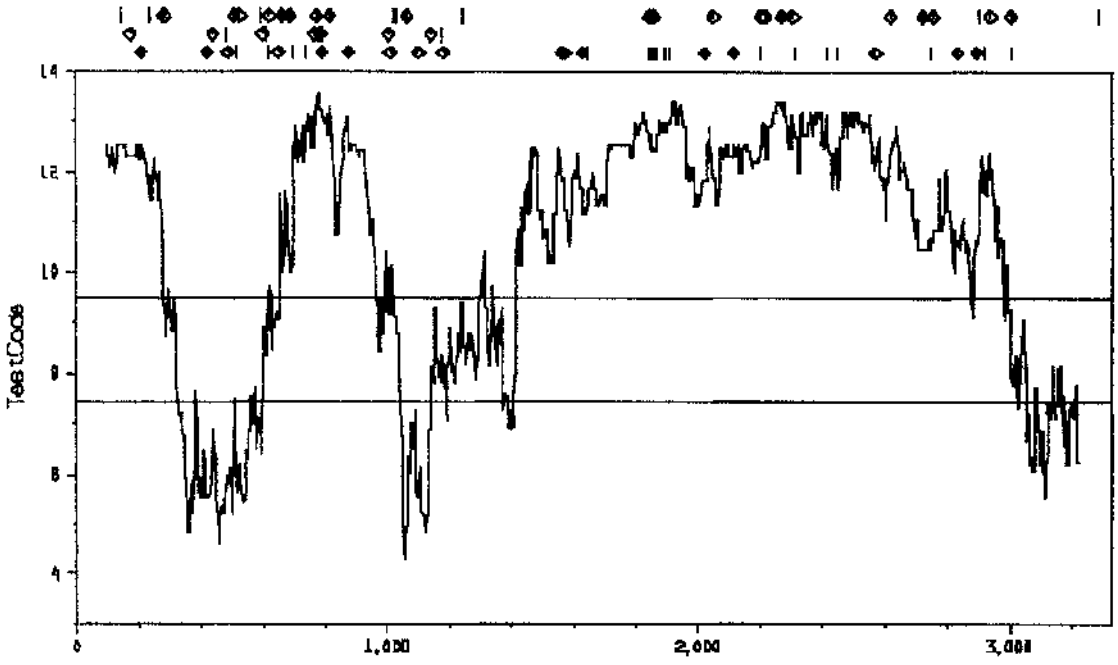
If one looks at the strain 151 sequence presented in the upper panel of Figure 43, it can be seen that there are predicted reading frames at about 0 - 300 bp, at 600 - 1000 bp, and at 1400 - 3000 bp. The third and largest one represents pucC and the central one the pucBA genes, Testcode fails to distinguish between the B and A genes because the window the statistic was calculated over is larger than the each gene. Testcode also clearly suggests the presence of a third ORF, upstream of pucB, this appears to be truncated however. The plot for the pucC gene is more convincing when it is in the correct sequence orientation, as can be seen in the lower panel of Figure 43. In this case the large coding region represented by pucC has a much less variable Testcode score, and also shows a sharper transition from non-coding to coding sequence and back. Figure 44 provides a similar picture for strain 149, the scale of the plots is expanded slightly because less sequence information was recorded for strain 149. The top panel once again shows three coding regions, representing the pucBA genes in the centre, the pucC gene to the right (this is truncated) and the other potential ORF to the left. The plot for the sequence in reverse orientation is at the bottom of the page, and also shows these three regions.

Translation of the DNA sequence into potential proteins yields only one ORF that is similar in the two strains. This is thus likely to represent the correct protein sequence, this putative protein fragment will henceforth be referred to as orf1. The peaks in the Testcode statistic upstream of pucB could perhaps be due to regulatory sequences in the DNA which one would expect to be present, since over six hundred base pairs upstream of the pucB gene was found to be involved in transcriptional regulation of the puc operon of *Rb. sphaeroides*. However further evidence that orf1 is a fragment of a real ORF, rather than conserved control regions upstream of pucB is presented below.

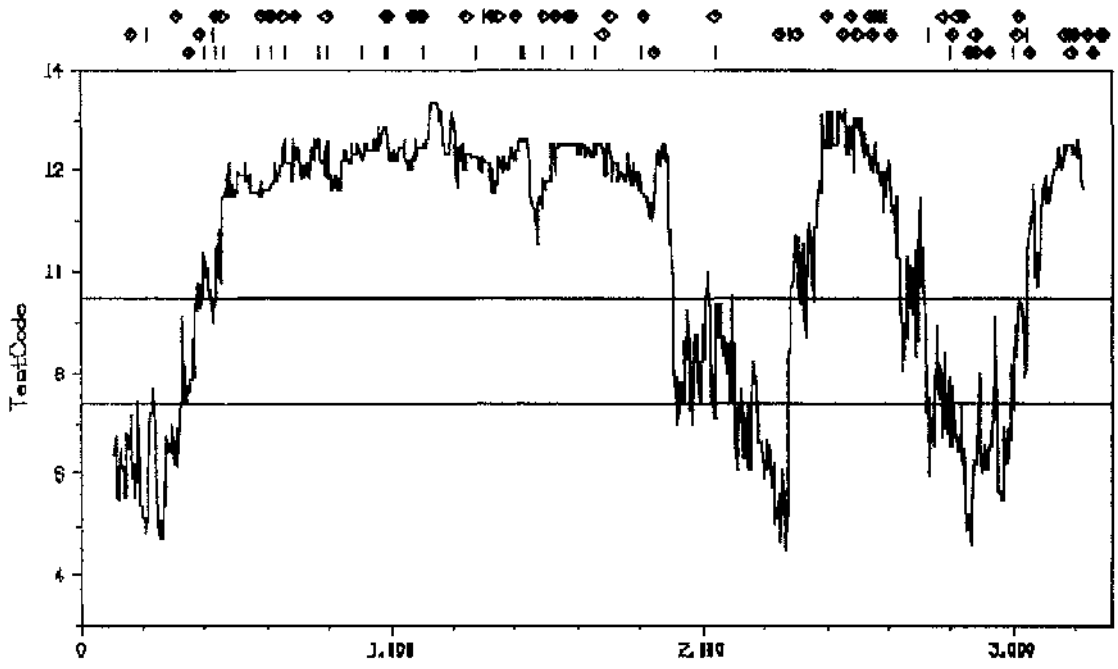
Figure 43

Testcode output for *Rv. gelatinosus* strain 151 sequence (p17T2.1X & p17T3.8x). The upper panel shows the sequence tested with pucBA in the 5′-3′ orientation, whilst the lower panel shows the sequence tested in the reverse orientation pucC running in the 5′-3′ direction.

pucBA 5′-3′



pucC 5′-3′

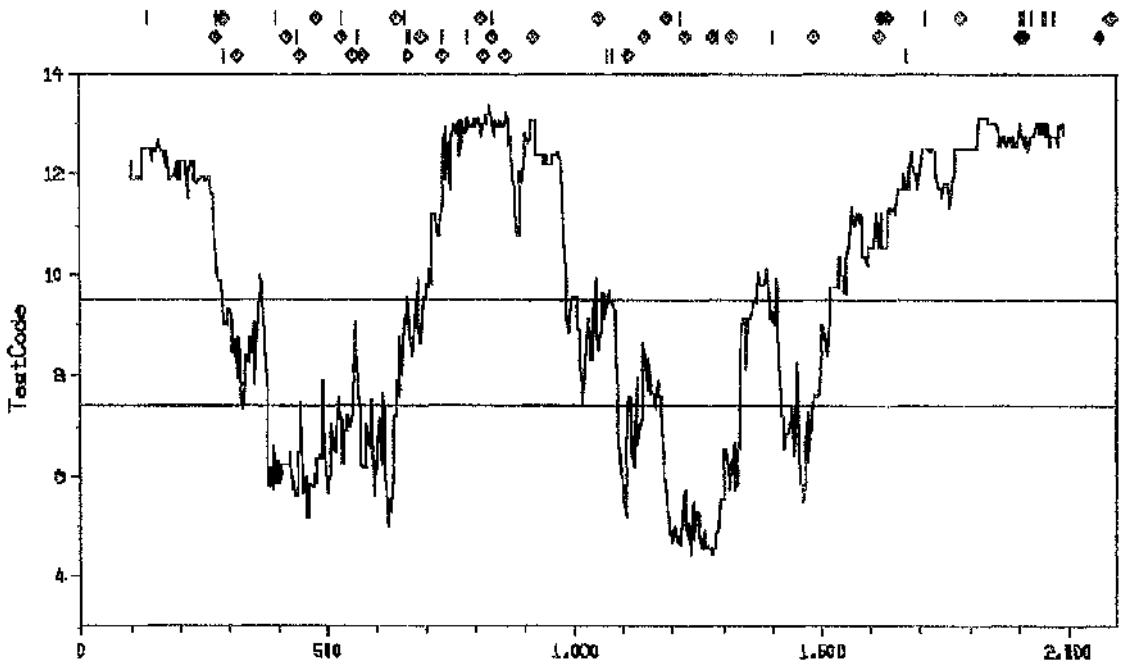


Note: The x axis represents the residue number. Residues above the upper line are significant, and represent possible protein coding regions. Those below the lower line are significant for non-coding regions whilst those between the two lines are of uncertain function. Diamonds above the frame are potential start sites, vertical lines potential stop sites.

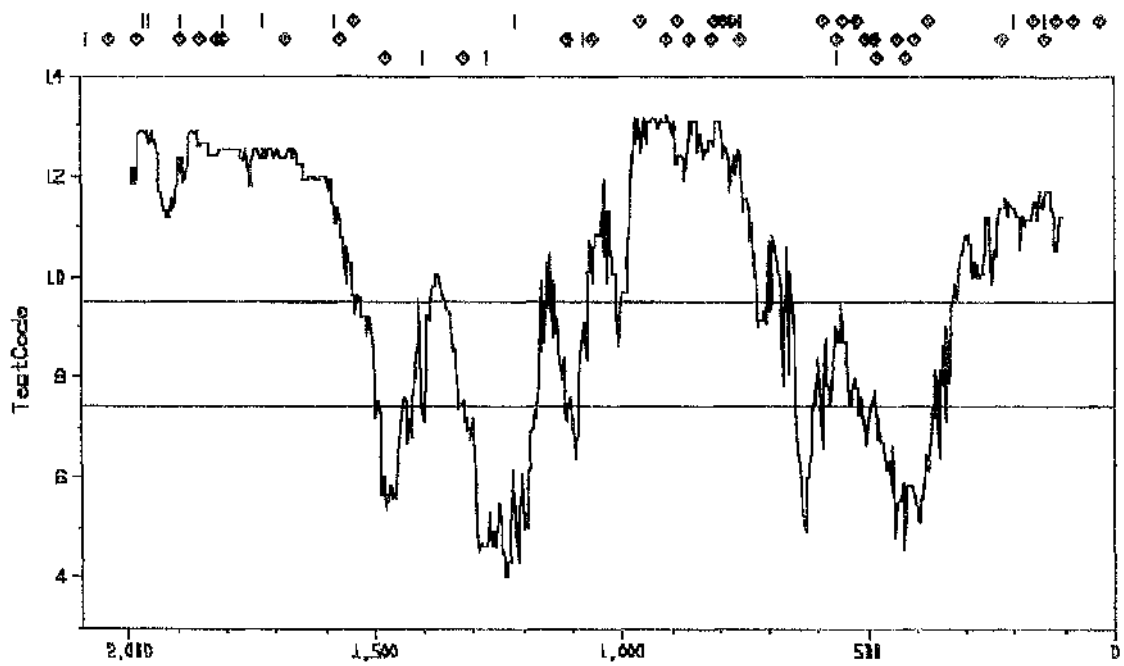
Figure 44

Testcode output for *Rv. gelatinosus* strain 149 sequence (pF6X2). The upper panel shows the sequence tested with pucBA in the 5´-3´ orientation, whilst the lower panel shows the sequence tested in the reverse orientation.

pucBA 5´-3´



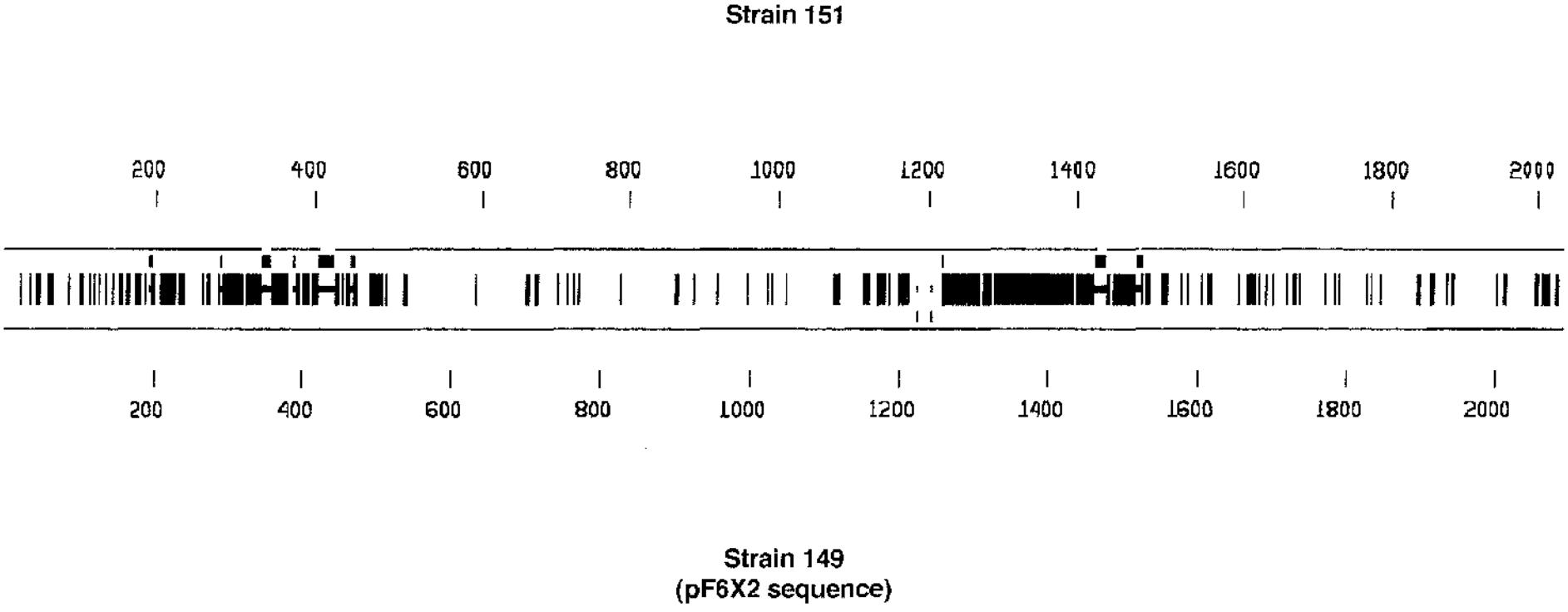
pucBA 3´-5´



Note: The x axis represents the residue number. Residues above the upper line are significant, and represent possible protein coding regions. Those below the lower line are significant for non-coding regions whilst those between the two lines are of uncertain function. Diamonds above the frame are potential start sites, vertical lines potential stop sites.

Figure 45

Output from the GCG Gapshow program. A schematic representation of the differences between the strain 151 and 149 DNA sequences equivalent to the strain 149 clone pF6X2 insert sequence. Strain 151 sequence is at the top, strain 149 at the bottom, the vertical black bars represent differences between the two strains. Breaks in the horizontal lines represent gaps inserted by the alignment program, as do vertical lines above and below the central block.



Comparison of strain 149 and 151 sequence

Figure 45 (opposite) illustrates the differences which occur between the strain 149 pF6X2 sequence and the corresponding sequence from strain 149. Since the sequence at either end of the pF6X2 insert is very homologous, the strain 151 sequence has been truncated at the sequence TTCGAG, which has clearly changed to the XhoI recognition sequence CTCGAG in strain 149 lying at the end of the strain 149 fragment. The truncated strain 151 sequence and the strain 149 sequence were aligned using the GAP program of the GCG suite, and the diagram shown generated from this alignment. The three sets of ORFs are visible in the diagram, where the sequence is more similar (and thus has less bars marking the difference). The 3′ region of the pucC gene is present to the right, from around the 1500 bp point to the end of the fragments. The pucBA genes occupy the most conserved region of the two fragments, in the 500 to 1000 bp segment, with orf1 occupying a less well conserved region over the first 300 bp. Between these conserved regions are areas with little apparent similarity and containing several deletions, indicating that these intergenic regions probably serve little functional purpose, although they could instead be responsible for functional differences between the two strains. It is interesting to note that the only deletion which occurs within a protein coding sequence is within orf1, and is exactly three base pairs, equivalent to single amino acid. This further supports the assertion that orf1 is part of a real protein.

The full GAP aligned sequence is presented in Figure 46 on the following pages. On closer inspection it becomes apparent that the conserved regions extend up and downstream of the protein coding sequences, to include potential promotors, repressor binding sites and downstream terminator sequences. These regulatory elements are marked on Figure 46 and described in more detail in the next section.

Figure 46

An alignment of the pF6X2 (bottom) and p17T2.1X (top, truncated) DNA sequence. The XhoI sites, start and stop codons of the open reading frames, potential promotors, repressor binding sites and transcription terminator sequences are marked. A note at the end of the figure describes these sites in more detail.

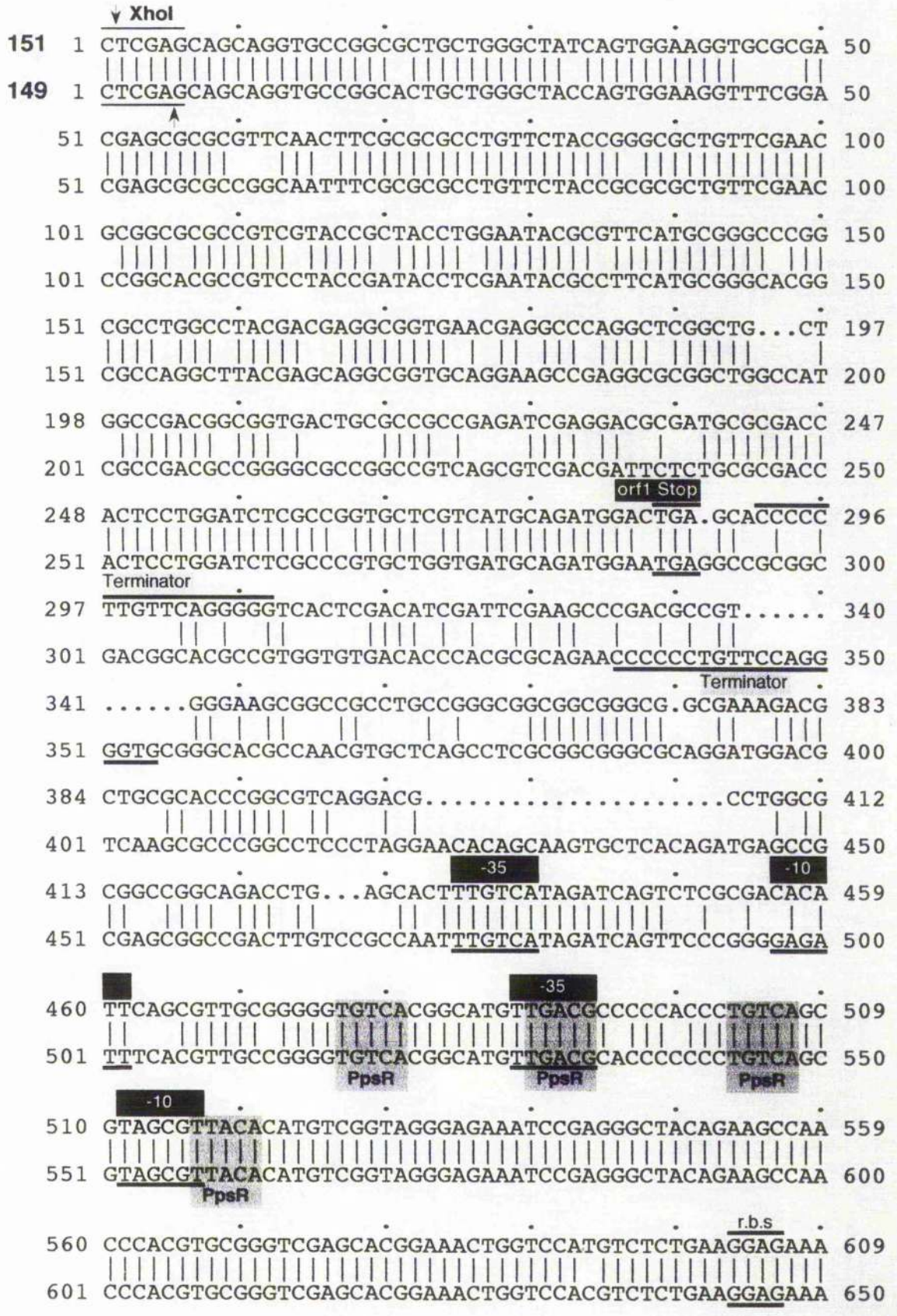


Figure 46 (continued)

Figure 46 (continued)

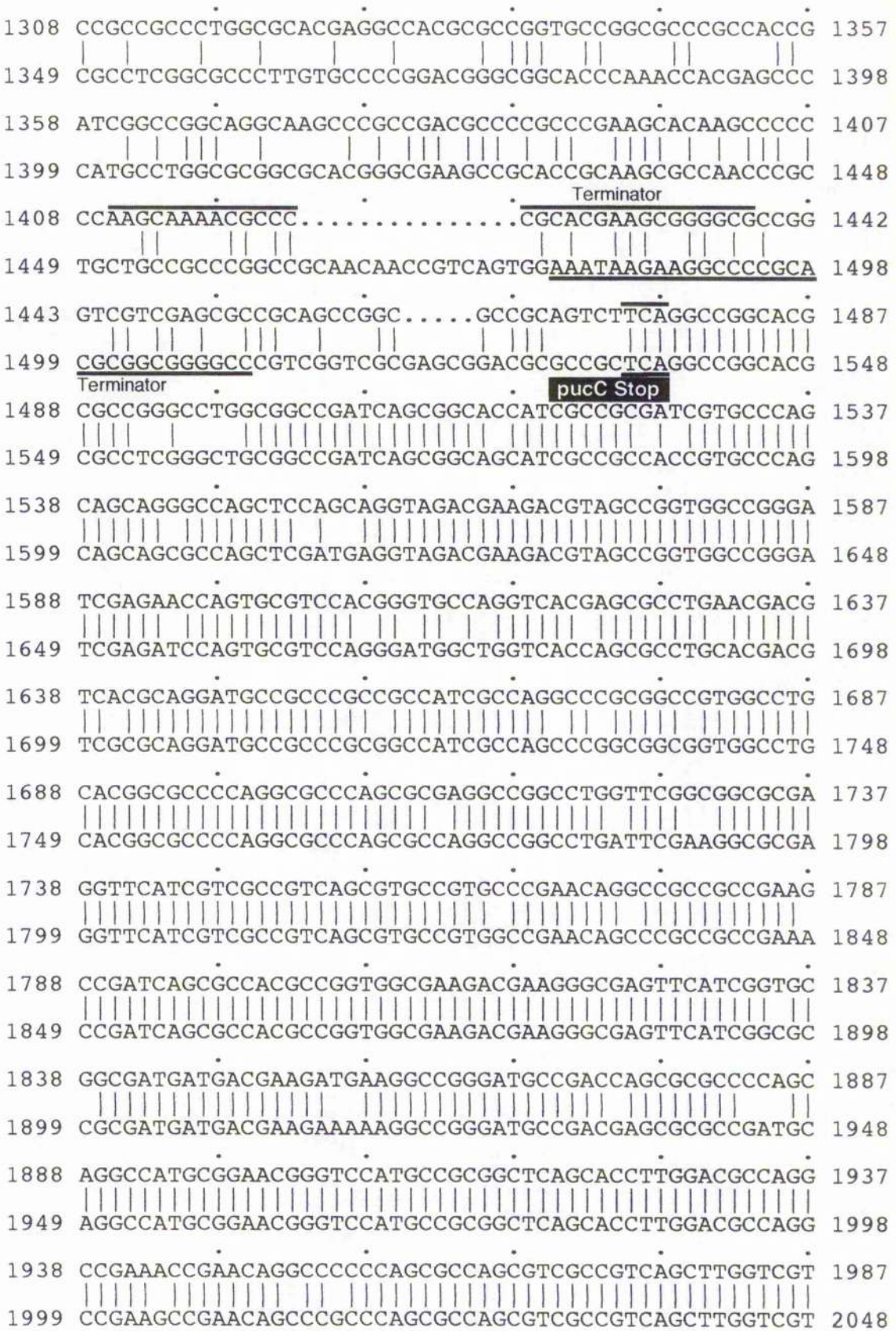


Figure 46 (continued)

```
1988 GTCGCCGACCGTCAGCGCCAGCACCTCGCCGCCATAGGGTTCGAG 2032
     ||||||||| || ||| | || | || |||||||||| || |||||
2049 GTCGCCGACGGTGAGCCCGAGGATCTGGCCGCCATAAGGCTCGAG 2093
                                          ^ XhoI
```

Parameters used for alignment:

Gap Weight:	5.000	Average Match:	1.000
Length Weight:	0.300	Average Mismatch:	0.000

GAP summary:

Quality:	1591.5	Length:	2095
Ratio:	0.783	Gaps:	11
Percent Similarity:	82.069	Percent Identity:	82.069

Note: The XhoI restriction sites lying at each end of the strain 149 fragment (lower sequence) are marked, the points at which XhoI cuts are marked by an arrow. The ribosome binding sites (r.b.s) and start/stop codons of the pucBA genes are marked with lines above and below the sequences and are labelled. The stop codons of the truncated pucC and putative orf1 genes are similarly marked. Sites with homology to the *E. coli* σ^{70} promotor -35 and -10 consensus sequences are marked as such. These are often overlapped by the PpsR repressor binding sites which are shaded in grey and labelled 'PpsR'. Potential transcription termination sites are marked with a line above or below and labelled 'terminator'.

Conserved elements outside of the open reading frames

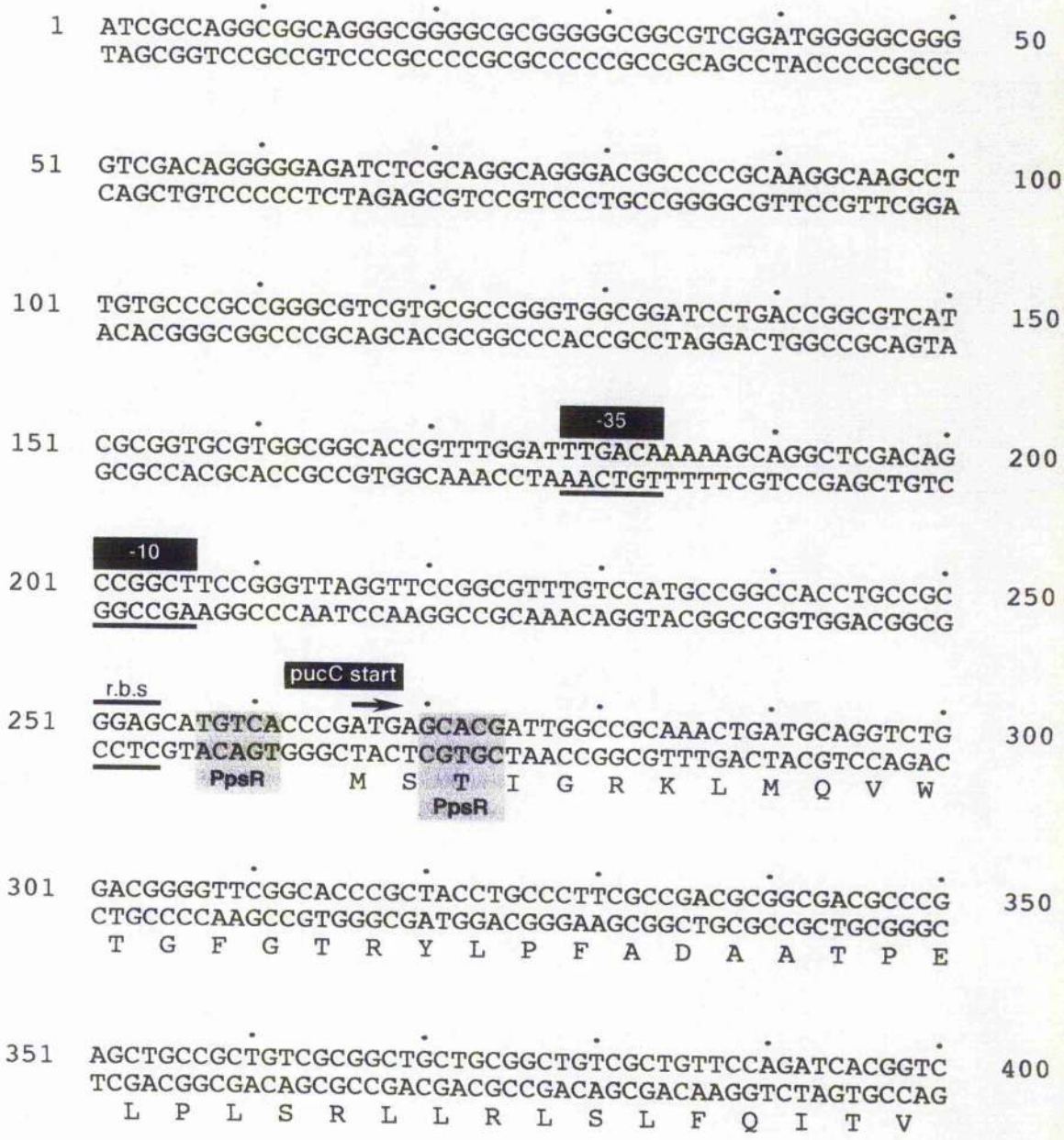
Upstream of the pucBA genes of both strains and the pucC gene of strain 151 are sequences similar to the σ^{70} consensus promotor sites of *E. coli*. Similar σ^{70} like promotor sites have been found upstream of photosynthesis genes in other photosynthetic bacteria. Like these other sites the -35 sequence is well conserved but the -10 sequence much less so. The sequences matching the promotor sites are marked on Figure 46 (previous pages) for the pucB upstream region, and on Figure 47 (overleaf) for the pucC region. Whilst only a single σ^{70} like promotor is seen upstream of pucC, both strain 149 and 151 have two such sites appearing in tandem upstream of the pucBA genes.

A search was made in the upstream sequences for elements showing homology to other known and suggested regulatory sequences which had been found in other bacteria, and are described in the introduction. A striking match to the consensus PpsR binding site, the oxygen regulated repressor of puc transcription in the *Rhodobacter* species was found. The consensus sequence $TGT(N)_{12}ACA$ is the core palindrome of a larger recognition sequence in *Rb. sphaeroides*. This sequence occurs in tandem overlapping the σ^{70} like promotor of the *Rb. sphaeroides* puc operon, a similar arrangement is seen in the strain 151 and 149 pucBA promotor like region. The repeat more proximal to pucB does not completely match the consensus however, although the difference is conserved between the two strains. Instead the more proximal element has the sequence $TGT(N)_{12}ACG$, suggesting that the consensus PpsR binding site of *Rv. gelatinosus* matches the sequence $TGT(N)_{12}ACR$ (where R means A or G). Since no PpsR binding sites were visible overlapping the putative strain 151 pucC σ^{70} like promotor, the sequence was searched for matches to the modified consensus defined above. A single match was found, straddling the start codon of the pucC gene. For those sites upstream of pucC the match to the PpsR binding sites of *Rb. sphaeroides* becomes even more startling when one considers the full binding site. This palindrome reads $TGTCA(N)_8TGACA$ and is repeated with a gap of 7 bp between the repeats. Both strain 149 and 151 have the near match $TGTCA(N)_8TGACG$ followed after a gap of 9 bp by $TGTCA(N)_8TTACA$. This means that only a single base differs in each repeat from the PpsR binding site of *Rb. sphaeroides*. The full site at the start of the strain 151 pucC gene is a little less similar, it reads $TGTCA(N)_8GCACG$. Thus if PpsR does bind at this point it would do so more weakly. No other significant matches to elements found upstream of photosynthesis genes in other species could be found, even when some mismatching was allowed.

The Shine & Delgarno ribosome binding site is also present just before the start codon of the pucB, A and C genes. The core sequence GGAG is 7 bp in front of the pucB gene of both strains whilst pucA has GGAA, also 7 bp before the ATG start codon. The ribosome binding site upstream of pucC is GGAG like that upstream of pucB, but is positioned 11 bp upstream of the start codon. This is well within the variation observed in the ribosome binding site position and may possibly reflect a regulatory function for ribosome binding site positioning.

Figure 47

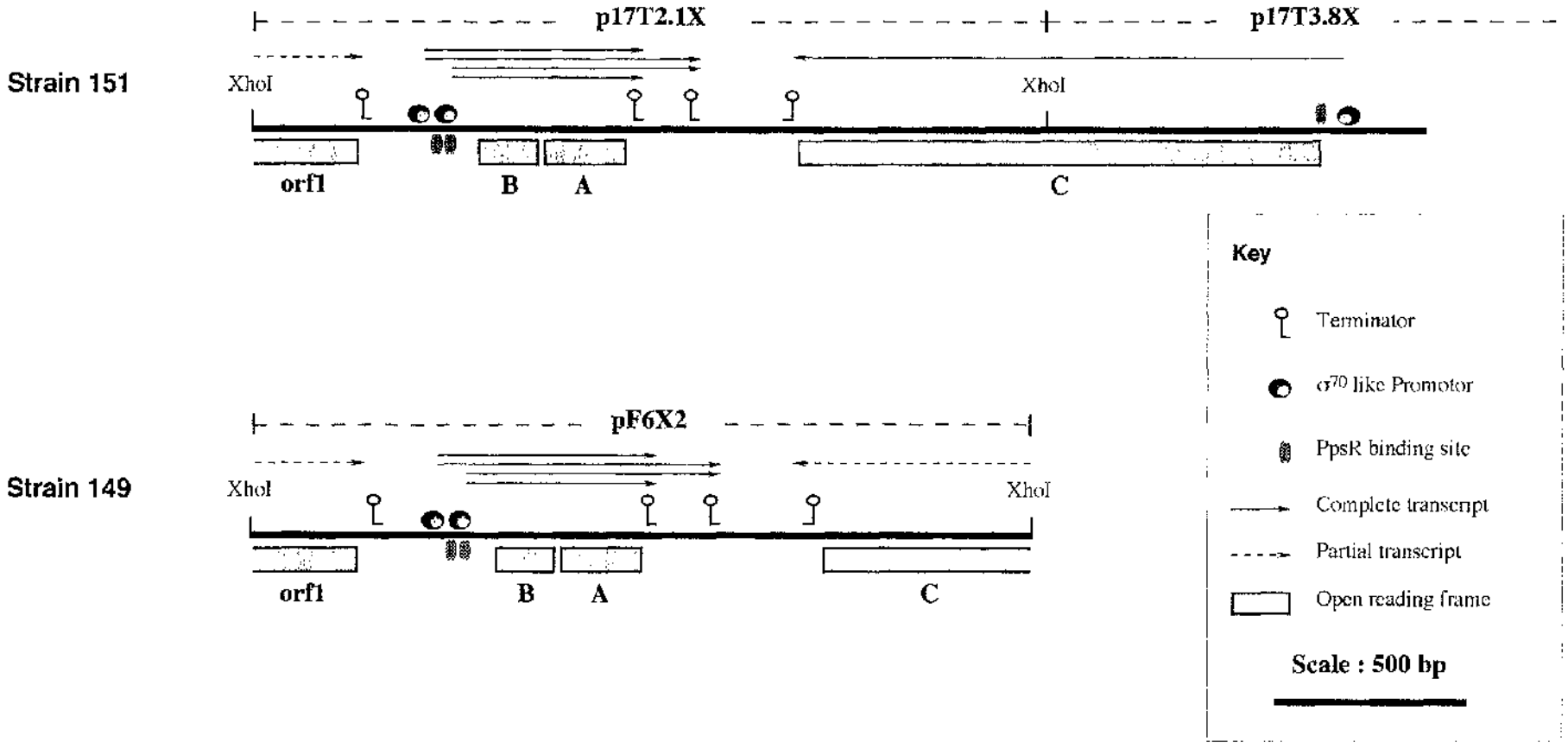
The sequence upstream from, and covering the 5′ end of, the pucC gene of *Rv. gelatinosus* strain 151 showing potential transcription control elements.



Note: The ribosome binding site (r.b.s) and start codon of the pucC gene are marked with lines above and below the sequences and are labelled. The sites with homology to the *E. coli* σ^{70} promotor -35 and -10 consensus sequences are marked as such. Whilst the potential PpsR repressor binding site is shaded in grey and labelled 'PpsR'. Also shown is the predicted protein sequence at the start of pucC.

Figure 48

A schematic view of the sequenced DNA fragments from *Rv. gelatinosus* strains 151 and 149, indicating the position of potential regulatory elements.



Note: The diagrams above indicate the position of potential regulatory elements relative to the sequenced fragments and the genes. The key above right describes the different elements. The potential transcripts are also marked, those that are not fully present in the sequenced DNA are dashed. The arrowheads indicate the direction that transcription would proceed.

A search for factor independent transcription terminators was carried out using the GCG program Terminator, which implements the algorithm of Brendel and Trifanov (1984). This searches potential terminators by comparing the sequence with a matrix derived from known terminators. These structures usually consist of a dyad symmetry element, capable of forming a stem & loop structure when transcribed into RNA, and a series of thymidine residues at which termination occurs. Brendel and Trifanov noted however that whilst the presence of these elements was sufficient for transcription termination, they were not necessary for termination suggesting other elements are involved. The sites found by the terminator program have been marked on Figure 46. Three potential terminators were identified by the Terminator program, a fourth has been marked immediately downstream of pucA, since this matches those structures found downstream of the puc and puf genes of the *Rhodobacter* species. Thus 2 potential terminators can be seen to lie downstream of pucA (with over 100 bp between them), a further element lies downstream of pucC and one lies downstream of the putative orf1. It should be noted however that the GAP alignment program has failed to align the orf1 and pucC terminator sequences correctly in the alignment presented in Figure 46, despite obvious homology between them. This highlights a problem inherent in the alignment program. GAP like most alignment programs tries to find the best *global* alignment for the whole sequence, this works well when the two sequences being compared are very similar. However when the similarity between two sequences gets lower the alignment often misses areas of *local* similarity which have biological significance. Thus in the region between the 3′ end of pucA and the 3′ end of pucC, and the region downstream of orf1 and before the pucBA promotor region where there is a great deal of difference between the two strains, GAP has favoured the global optimum over the 'biologically significant' alignment.

The potential regulatory elements described above are summarised in relation to the genes present on the sequenced fragments in Figure 48 (opposite). The potential RNA transcripts are also marked.

Protein sequence analysis

Comparative analysis

The predicted α- and β-polypeptide amino acid sequences extracted from the sequenced DNA are given at the top of Figure 49 (overleaf). This illustrates the fact that the two strains have α-polypeptides with identical sequence, and that only one amino acid differs between the strains as far as the β-polypeptide is concerned. This difference is visible as the two unboxed residues, an isoleucine in strain 149 is replaced by a valine in strain 151. Such a change to a homologous amino acid is unlikely to have a significant effect on the structure of LH2. The predicted protein sequence is an exact match to the polypeptide sequence obtained by protein sequencing of the strain 151 and strain 149 polypeptides (see Brunisholz *et al.*, 1994). Thus confirming the source of the cloned DNA fragments.

Also shown in Figure 49 is an alignment of the strain 151 and 149 putative orf1 proteins, identical residues are boxed whilst homologous residues are shaded. The fragment of the protein present on each of the clones has 77% identity and 91% homology. A search was made of the Swiss-Prot, PIR, Genbank and EMBL databases for proteins with homology to this fragment, but no significant matches were found. Thus it may be necessary to sequence the rest of the putative protein before a function can be assigned.

Shown in Figure 50 (following Figure 49) is an alignment of the predicted protein sequence of the strain 151 PucC protein with the fragment of the strain 149 PucC protein for which the DNA was sequenced. The identical residues are boxed and the homologous residues shaded. Over the 185 residue fragment the proteins are 93% identical and 99% homologous, suggesting that the protein is well conserved in the two strains.

Several other genes are known to be homologous to the pucC genes of *Rb. capsulatus* and *Rb. sphaeroides*. These include the *Rb. capsulatus* orf 477 protein, which appears to play a role in LH1 assembly analogous to that performed by pucC for LH2. A homologue of orf 477 has also been found in *Rs. rubrum*, and is known as G115 protein. This latter protein sequence has not been published as yet but a preliminary sequence has been placed in the PIR database, this preliminary sequence was used in the analysis described below.

The strain 151 PucC protein sequence was compared with these other proteins by way of a multiple alignment. The sequences were aligned using the GCG PileUp program which attempts to bind the best global alignment for a set of sequences. This works well for the pucC and orf 477 protein homologues as can be seen if on looks at the alignment displayed in Figure 51 (overleaf). Once again identical residues are boxed whilst the residues homologous to the primary sequence (*Rb. sphaeroides* PucC) are shaded. This shows quite clearly that these are very similar proteins at the level of amino acid sequence. The GAP program was used to generate pairwise alignments of the various proteins, in order to calculate the percentage identity and similarity. The strain 151 PucC gene is most similar to *Rb. sphaeroides* PucC at 78% (identity 58%), it is 73% similar to *Rb. capsulatus* PucC (57% identity) and *Rs. rubrum* orf G115 (58% identity), and is least similar 69% to *Rb. capsulatus* orf 477 (identity 50%).

Figure 49

Predicted protein sequences from strain 151 and 149 of *Rv. gelatinosus*, translated from the sequenced DNA.

β-polypeptides

```
149 beta   1  MADDANKVWPSGLTTAEAEELQKGLVDGTRIFGVI 35
151 beta   1  MADDANKVWPSGLTTAEAEELQKGLVDGTRVFGVI 35

149 beta  36  AVLAHILAYAYTPWLH                    51
151 beta  36  AVLAHILAYAYTPWLH                    51
```

α-polypeptides

```
149 alpha  1  MNQGKVWRVVKPTVGVPVYLGAVAVTALILHGGLL 35
151 alpha  1  MNQGKVWRVVKPTVGVPVYLGAVAVTALILHGGLL 35

149 alpha 36  AKTDWFGAYWNGGKKAAAAAAAAVAPAPVAAPQAPA 70
151 alpha 36  AKTDWFGAYWNGGKKAAAAAAAAVAPAPVAAPQAPA 70

149 alpha 71  Q                                   71
151 alpha 71  Q                                   71
```

Orf1 fragment amino acid sequence

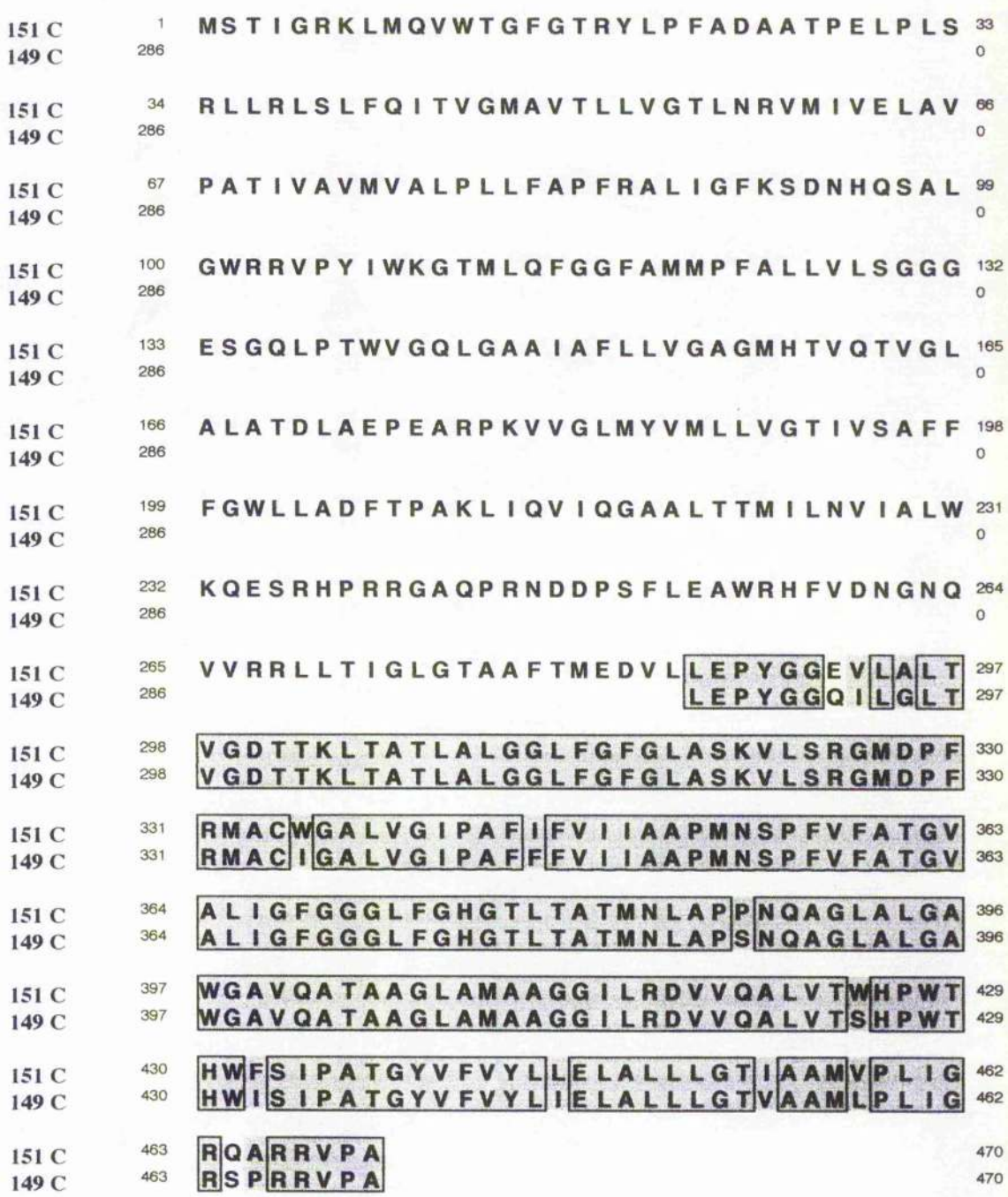
```
149 orf1   1  VPALLGYQWKVSDERAGNFARLFYRALFEPGTPSY 35
151 orf1   1  VPALLGYQWKVRDERAFNFARLFYRALFERGAPSY 35

149 orf1  36  RYLEYAFMRARRQAYEQAVQEAEARLAIADAGAPA 70
151 orf1  36  RYLEYAFMRARRLAYDEAVNEAQARL-LADGGDCA 69

149 orf1  71  VSVDDSLRDHSWISPVLVMQME               92
151 orf1  70  AEIEDAMRDHSWISPVLVMQMD               91
```

Note: Alignments of the predicted amino acid sequences of the LH2 β-polypeptide, LH2 α-polypeptide, and the fragment of the putative orf1 polypeptide present upstream of the pucB gene. The single letter amino acid code is used, identical residues are boxed in all sequences, additionally the orf1 sequences are shaded on the basis of homology. The GES homology scale was used.

Figure 50

Predicted PucC protein sequences from strain 151 and 149 of *Rv. gelatinosus*, translated from the sequenced DNA.

PucC proteins

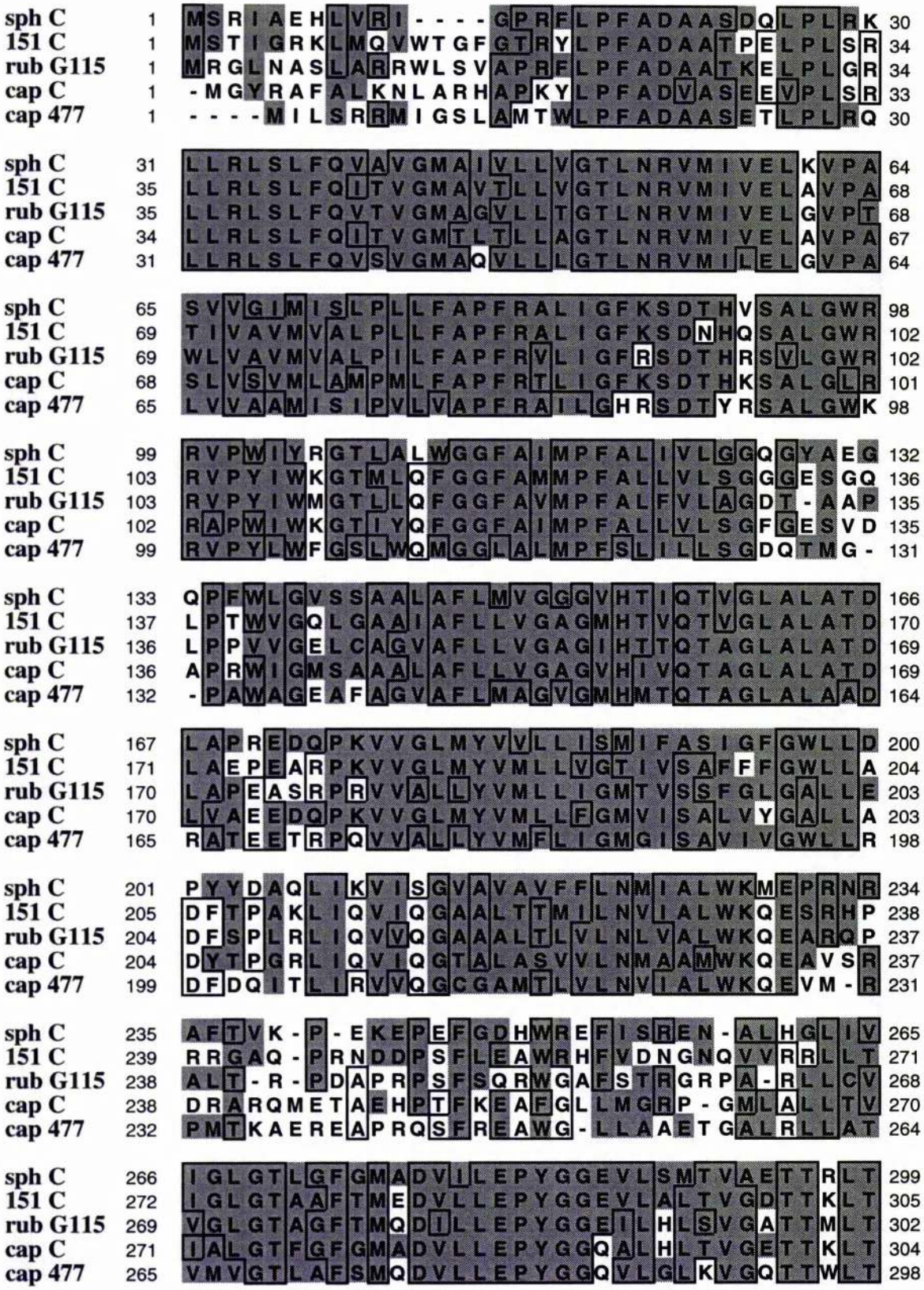


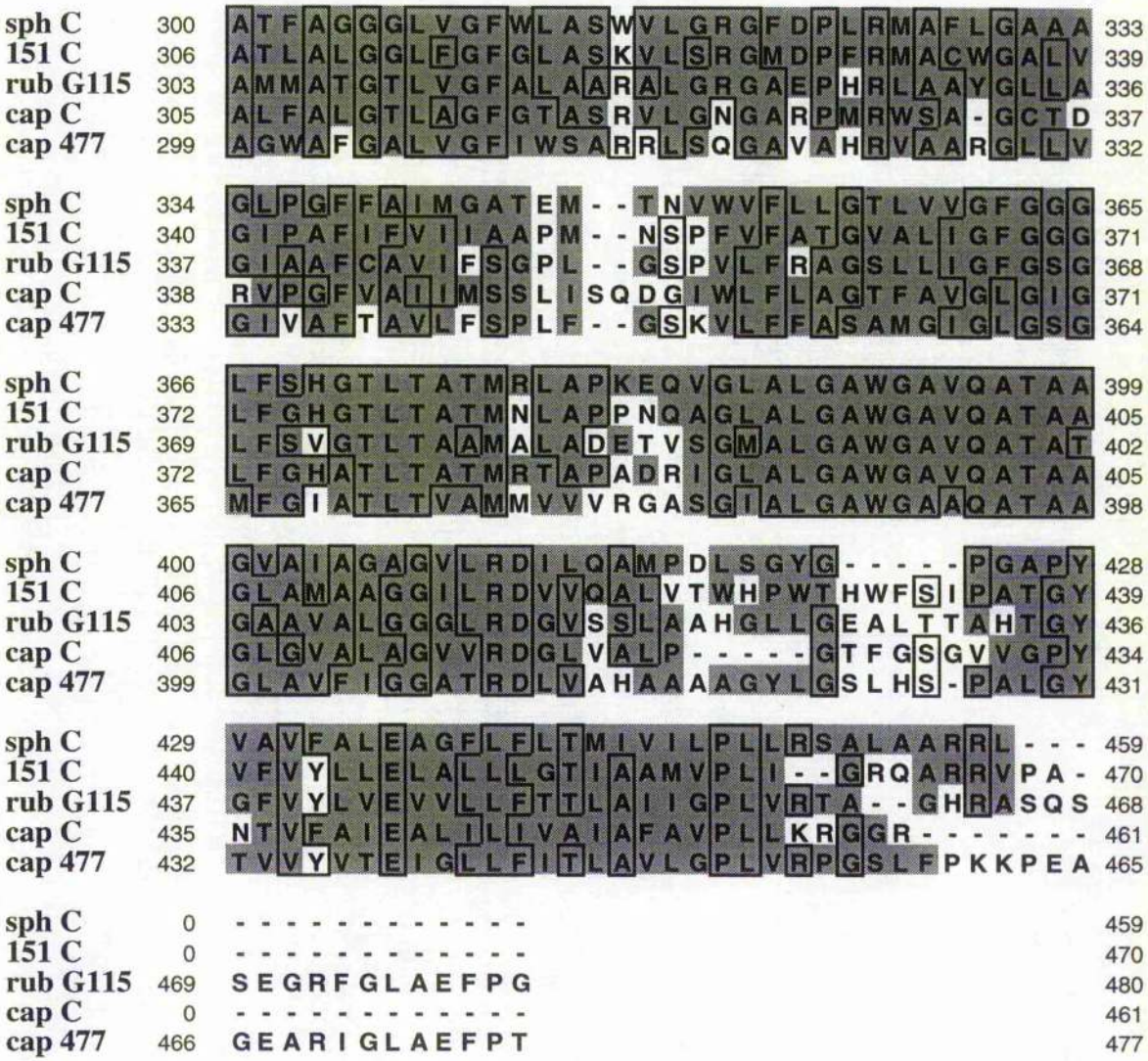
Note: Alignments of the predicted amino acid sequences of the PucC protein from strain 151 and strain 149. Only part of the strain 149 pucC coding region was sequenced. The single letter amino acid translation is used. Identical residues are boxed, homologous residues are shaded. The GES homology scale was used.

Figure 51

A multiple alignment of the *Rv. gelatinosus* PucC protein with PucC and homologous proteins from other species.

PucC proteins





Note: Alignments of the predicted amino acid sequences of the PucC protein from *Rv. gelatinosus* strain 151 (151 C), *Rb. sphaeroides* PucC (sph C), *Rb. capsulatus* PucC (cap C), *Rs. rubrum* G115 (rub G115), and *Rb. capsulatus* orf477 (cap 477). The sequences were aligned using the GCG PileUp program. The single letter amino acid translation is used. Identical residues are boxed, residues homologous to the top sequence (*Rb. sphaeroides*) are shaded. The GES homology scale was used.

Thus the strain 151 pucC gene is as similar to the *Rb. sphaeroides* pucC as the two *Rhodobacter* PucC proteins are to each other.

PucC protein structure

The prediction of protein secondary structure from primary sequence information is notoriously inaccurate, it does however work somewhat better for membrane proteins. This appears to be because the membrane environment imposes certain constrictions on the structure. Many membrane proteins have membrane spanning α-helices composed of hydrophobic amino acid residues, thus when the residue number is plotted against a measure of the hydrophobicity of each residue, it is often possible to identify runs of hydrophobic residues equivalent to transmembrane helices. Since it had been suggested that the PucC protein from *Rb. capsulatus* was a membrane protein, and that it contained 12 membrane spanning segments, it was of interest to compare the predicted structure with the strain 151 PucC protein.

Figure 52 (overleaf) shows two plots, at the top is a hydropathy plot using the scale of Kyte & Doolittle, (1982), and averaged over a window of 20 residues to highlight potential membrane spanning segments. This shows about 4 segments with a score over 1.6, such a score suggests transmembrane helices are present. The hydropathy plot also shows a hydrophilic section in the centre of the protein.

The lower plot shows a slightly different approach to predicting membrane spanning segments. This plot was calculated using the TMpred program, which makes a prediction of membrane-spanning regions and their orientation with regard to the cytoplasm. The algorithm is based on statistical analysis of TMbase, a database of naturally occurring transmembrane proteins of known structure (Hofmann & Stoffel, 1993). The prediction is made using a combination of several weight-matrices for scoring, only scores over 500 are considered significant for a transmembrane segment. This plot shows more clearly the potential membrane spanning segments, 11 of which are over the score 500 cutoff point. An extremely similar pattern was seen when the two other pucC proteins (Figure 53, following 52) and the orf477 and G115 proteins (Figure 54, opposite 53) were analysed using TMpred. The first 250 residues appear to be extremely homologous in the 5 proteins, each one apparently having 3 pairs of membrane spanning helices, separated by more hydrophilic segments. This section of the protein is delineated towards the centre of the protein by a large hydrophilic segment, a structural feature which appears in all 5 proteins. This section is also one of the more variable regions in the sequence alignments, suggesting that whilst the structure of this section has been conserved the sequence is less so. The remainder of the protein shows more variation. Two sets can be observed, one containing PucC from the two *Rhodobacter* species, whilst the second set contains the strain 151 PucC protein, orf477 and G115. The first set shows relatively clearly another six potential transmembrane segments, bringing the total to 12 and matching the experimentally determined number from *Rb. capsulatus* PucC. The second set shows one difference close to the central hydrophilic segment, a rise in the TMpred score that would be the next membrane spanning segment in the *Rhodobacter* pucC proteins barely rises above the level of significance in the second set of proteins.

Figure 52

Hydropathy plot and prediction of transmembrane helices for strain 151 PucC protein.

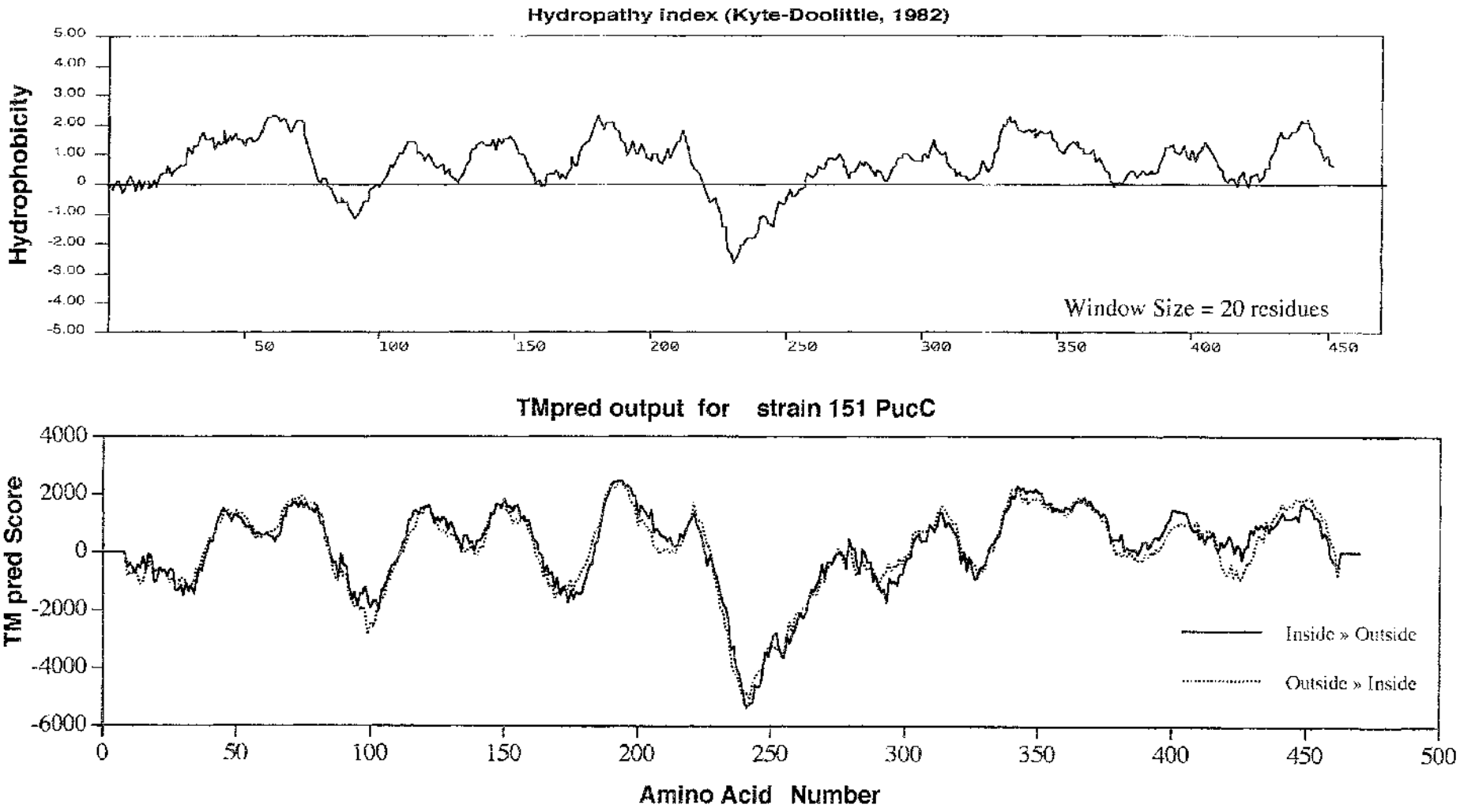


Figure 53

Prediction of transmembrane helices for PucC proteins.

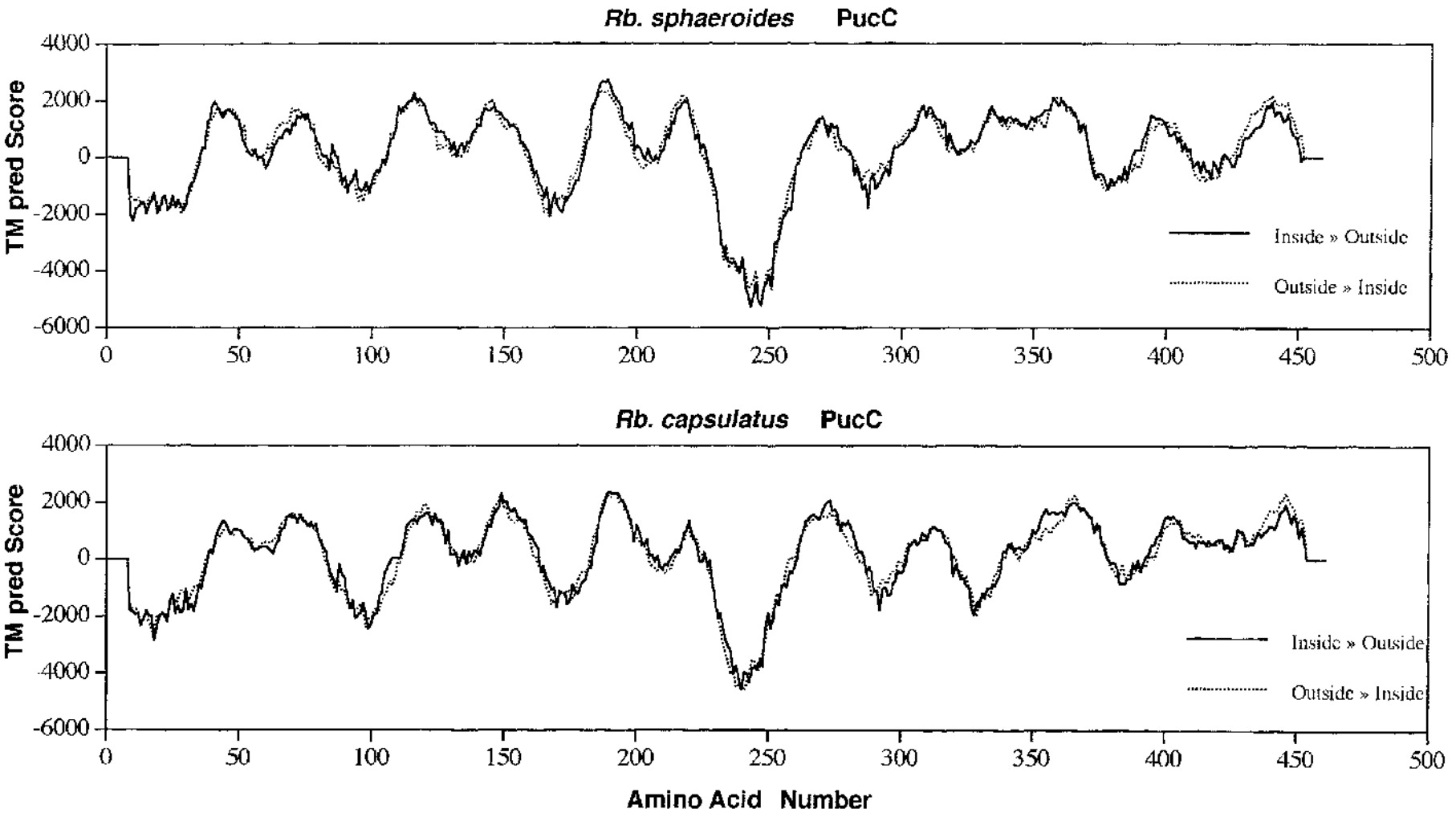
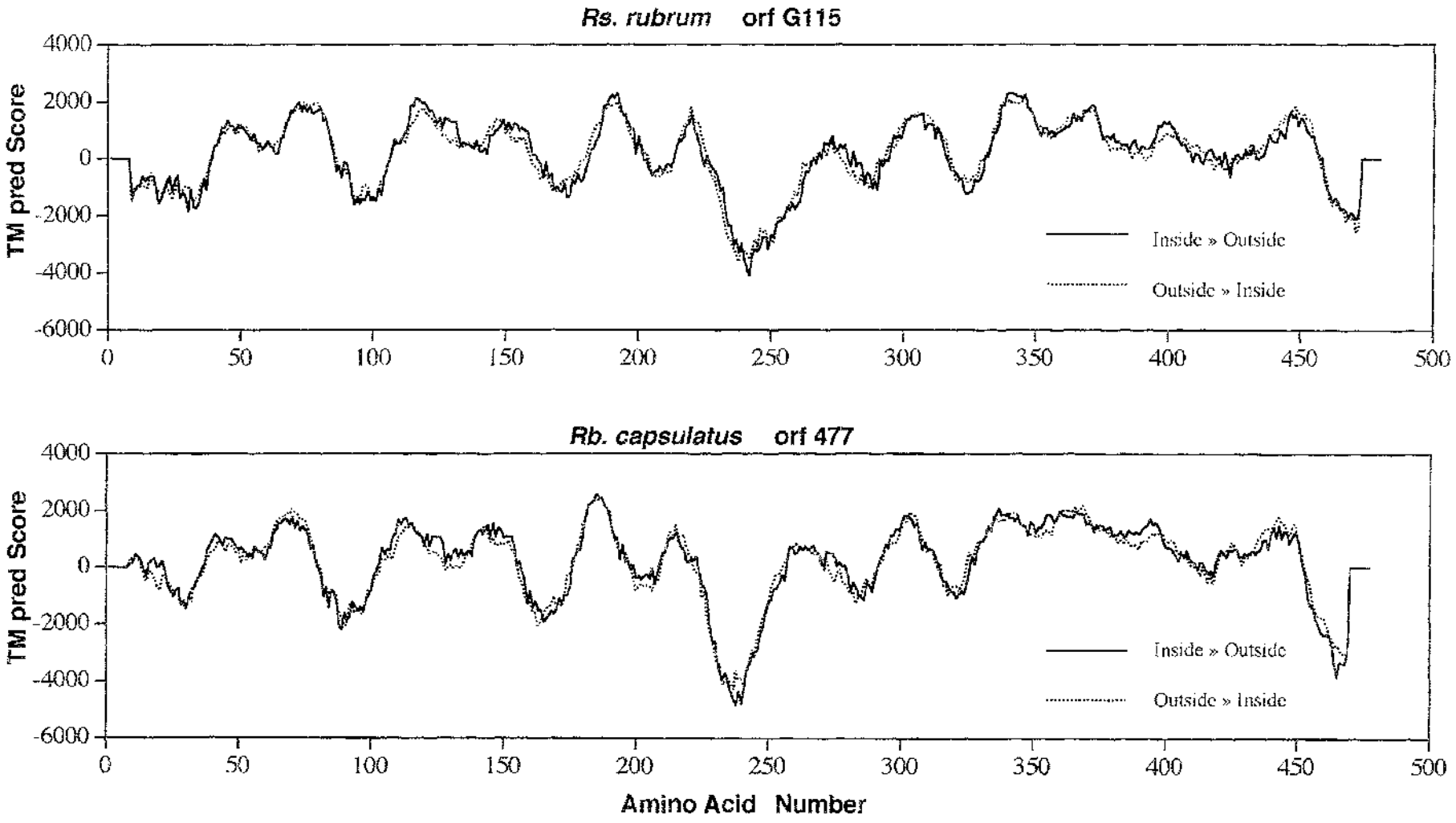


Figure 54

Prediction of transmembrane helices for G115 and orf 477 proteins.



A second difference is visible to the C-terminal end of the protein, where the membrane spaning segments of the *Rhodobacter* PucC proteins are much better defined. In contrast the other proteins have a much flatter plot in this region, although the score is well above the level of significance.

The Tmpred data also suggest that both the N and C terminii of the proteins lie in the cytoplasm. It is in the C-terminus that the orf477 and G115 proteins are differentiated from the PucC proteins, since both are slightly longer and apparently have a hydrophillic terminal region.

Conclusions

The predicted protein sequences derived from the sequenced DNA of *Rv. gelatinosus* strain 149 and 151 provide some interesting results. Firstly the predicted β- and α-polypeptide sequences confirm the protein sequenced polypeptide data. This shows that the only difference between the two strains resides in the beta polypeptides, where a single amino acid residue is differs. The difference proves useful as a way of confirming the source of the cloned DNA.

The strain 151 PucC protein appears to be 370 amino acids long and has a calculated molecular weight of around 50000, this is similar in size and weight to other PucC proteins which have been DNA sequenced. The PucC proteins of the two strains also appear to be highly homologous, at least over the 185 residues that were translated from the truncated strain 149 gene coding sequence, which was found to be 93% identical to that of strain 151. The strain 151 pucC gene was also shown to be very homologous, both in terms of amino acid sequence similarity and structure. The analysis of PucC and homologous proteins suggests that they have two main domains, each consisting of 6 membrane spanning segments and separated by a hydrophilic segment. The homology shown between the strain 151 PucC protein and the orf477 and G115 proteins, which are associated with LH1, could have two possible roots. Firstly since pucC and indeed the puc operon as a whole is thought to have arisen by gene duplication of an ancestral puf operon, it may be that the *Rv. gelatinosus* strain 151 PucC is closer to its ancestral LH1 equivalent than the PucC of the two *Rhodobacter* species. Alternatively the differences could simply represent functional changes in the proteins, tailoring them to specific LHCs. Further analysis of these proteins could prove useful in modelling, and allow the design of mutagenesis studies which could eventually elucidate the function of the proteins.

The orf1 protein also appears to have been conserved between the two strains, no significant homology was found to other proteins however. This perhaps suggests that orf1 is not part of a real protein, but instead is an artifact, however several features of the DNA sequence encoding it and downstream of it suggest otherwise. These include the fact that DNA sequence homology between the two strains of *Rv. gelatinosus* is high in this region, and only begins to break down when the the stop codon is reached. Downstream of the putative stop codon there is a conserved stem structure with homology to transcription terminators. These are usually found immediately downstream of genes, further suggesting that orf1 is a real protein coding sequence. The region around this terminator structure has been subject to a lot of change with several deletions/insertions, such that the strain 149 terminator is over 40 bp downstream of the orf1 stop codon, whilst that of strain 151 is 2 bp after the stop. This change around the terminator, whilst the terminator has changed little itself (2 bases differ between the strains), suggests that it has an important function to play.

Several other elements were identified in the DNA sequence obtained from strain 149 and 151. A sequence similar to the σ^{70} RNA polymerase subunit recognition site was identified upstream of pucC and two such elements were identified upstream of pucB. None of these potential promotors matches the *E. coli* σ^{70} consensus 100%, but neither do *E. coli* promotors. In a survey of 112 promotors in *E. coli* none matched the

consensus exactly (Collado-vides *et al.*, 1991). This survey also indicated that most repressor proteins bind in a position which covers part of the promotor site and physically prevents transcription, and also that they often bind in tandem. This has been found to be the case upstream of the *Rb. sphaeroides* puc operon where a σ^{70} like promotor sequence is blocked by the oxygen responsive repressor PpsR. Near palindromes with homology to the PpsR binding site were found in tandem covering the proximal σ^{70} like promotor of *Rv. gelatinosus* puc genes, much in the same way that they do in *Rb. sphaeroides*. A potential consensus for the *Rv. gelatinosus* puc repressor appears to be $TGT(N)_{12}ACG$, a sequence matching this is also found across the pucC start codon, where it may interfere with transcription. The tandem sites upstream of pucB probably allow dimerisation of the two repressor molecules, causing stronger repression than would be observed for the lone site at the start of pucC.

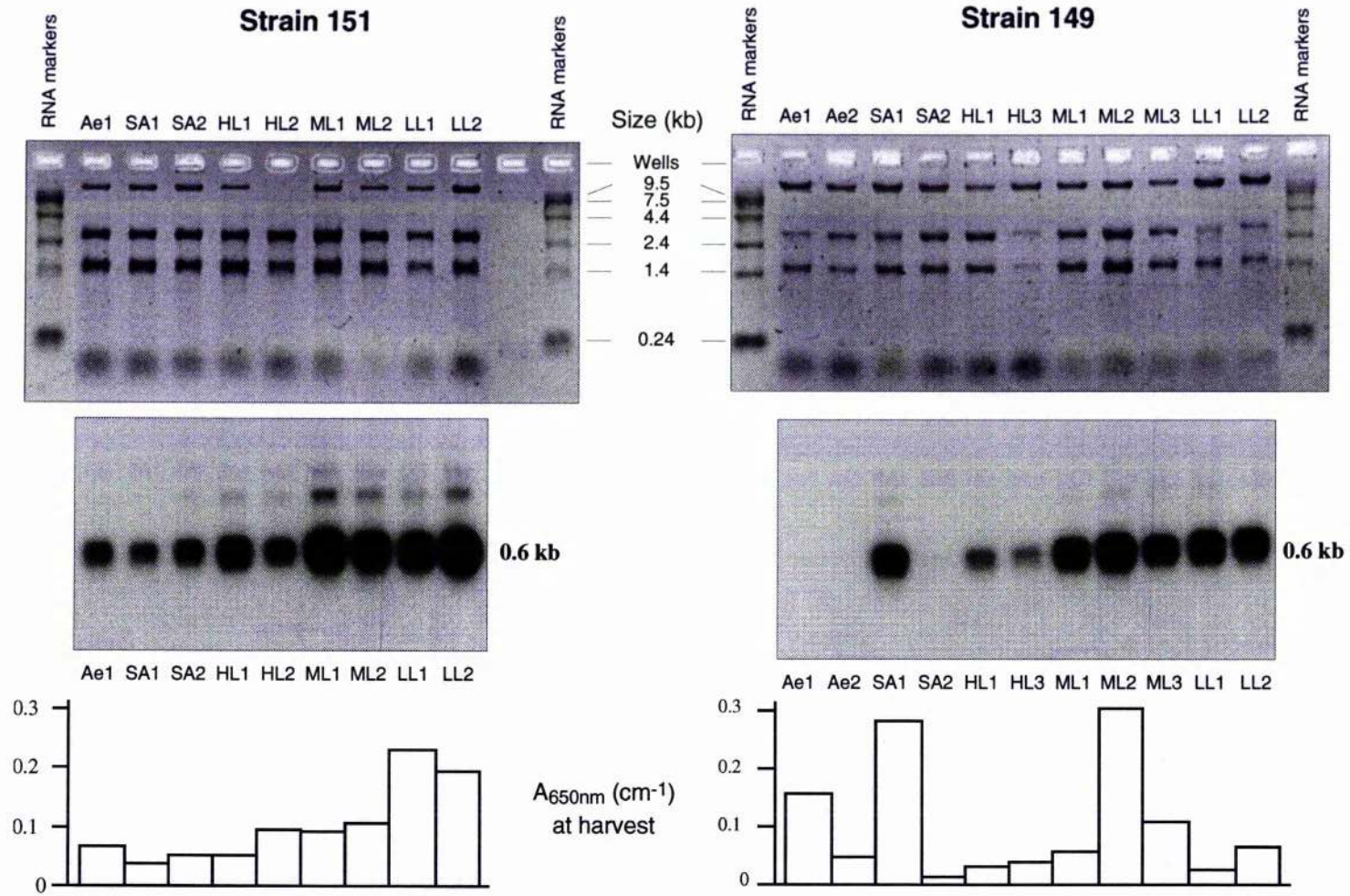
The presence of PpsR binding site appears to be a highly conserved feature in the upstream regions of photosynthesis genes, thus the protein will most likely be identified in *Rv. gelatinosus* at some point in the future. The tandem promotor elements are also not uncommon, Collado-Vides *et al.*, also found that 40% of E. coli promotors occurred in tandem. This presumably allows the bacterium to impose differential regulation on the two promotors, often one promotor is constitutive whilst the other requires activator proteins to function. One could speculate that in *Rv. gelatinosus* the promotor more proximal to pucB would be expressed constitutively in the absence of oxygen, whilst the distal promotor could be regulated by light. No other regulatory sites were identified upstream of pucB however, so its not clear how such differential regulation would be achieved. The use of different promotors would of course lead to different sized transcripts being produced, although the small difference in size may require the use of polyacrylamide gel electrophoresis, rather than agarose gel analysis.

Potential factor independent transcription terminators were also located in the intergenic region between pucA and pucC. One of these is oriented such that it would terminate transcription at the end of pucC, whilst the other two elements would act as terminators for pucBA. The two terminators downstream of pucA may have arisen as a result of selective pressure against read-through into the pucC gene. This would produce pucC anti-sense RNA which could interfere with the synthesis of PucC protein. If read-through occurred from the highly expressed pucBA genes, it could swamp transcription from the pucC promotor which is likely to operate at much lower efficiency (expression of the Rhodobacter pucC genes is much lower than their pucBA genes). Alternatively the RNA secondary structure element immediately downstream of the pucA gene may be involved primarily with the stability of the pucBA transcript, like that downstream of the *Rb. capsulatus* pucBA genes, rather than acting as a terminator (it could also do both).

A final point concerns the base composition of the sequenced DNA. This may indicate a difference between the two strains, or may simply result from local differences in DNA composition in the sequenced DNA. As more sequence information becomes available from these strains it should be possible to distinguish between these possibilities.

Figure 55

Agarose gels and associated Northern blots showing RNA samples isolated from strains 149 and 151 of *Rv. gelatinosus* cultured under a range of conditions.



Note: Pictured at the top are the two halves of a 1.4% agarose gel containing strain 151 (to the left) and strain 149 (to the right) RNA samples cultured under a variety of conditions. The abbreviated sample name is given above each lane, the main text has details of the full name. An RNA ladder was loaded at each side of each gel, the sizes of the fragments are given in the centre. Below the gels are Northern blots taken from each gel and probed with the pucBA genes from strain 151. The blots thus represent expression of the pucBA genes. The estimated size of the hybridising transcript is given at the side of the autoradiographs. At the very bottom are bar charts indicating the whole cell A_{650nm} at cell harvest (cm^{-1}).

Results Part 5:
Expression of the *Rv. gelatinosus* puc Genes

Introduction

Having confirmed in the first results section that there was a measurable difference in synthesis of LH2 between the two strains 149 and 151, and having cloned the pucBA genes from both strains, it was interesting to ask whether a difference between the two strains occurred at the level of puc gene transcription. To answer this question Northern analysis was carried out on samples of RNA isolated from *Rv. gelatinosus* cultured under the irradiance conditions described in the first results section. The results of the Northern analysis are presented below.

Also presented in this section are the results of a PCR experiment designed to isolate the pucBA genes from strain 151 for study in a heterologous expression system. The PCR product was also used as a probe extensively during the course of the strain 149 library construction, and for the subsequent Southern and Northern analysis.

Results

Northern analysis

RNA was extracted as described in the Materials and methods section "Isolation and purification of total RNA", from cells cultured under the HL, ML and LL irradiance levels described previously. Dark aerobic and dark semi-aerobic culture conditions were also used to investigate the effect of oxygen on puc gene expression in *Rv. gelatinosus*. At least 2 separate cultures were produced of each strain in each growth condition. Following the RNA isolation and purification samples were analysed on agarose gels and Northern blotted onto nylon membrane. These nylon filters were then probed with the pucBA genes isolated from strain 151 to examine the level of puc gene transcript in each sample. An aliquot was also removed from each sample before RNA isolation, and the absorption spectra of the cells recorded. Comparison with earlier spectra recorded from cells grown under the same culture conditions (as described in "Results 1"), confirmed that the same response to irradiance level had occurred. Whilst the absorption value at 650 nm extracted from the spectra provided a useful measure of the stage of growth that the cultures had reached at cell harvest. The cell density is important because at higher concentrations the cells could begin to shade each other and thus alter the irradiance level that they were effectively receiving.

The results of the Northern analysis are presented in Figure 55 (opposite) and the spectral data corresponding to the samples loaded on the gel is displayed on the following pages in Figure 56 for strain 151 and Figure 57 for strain 149. Figure 55 (opposite) pictures two halves of an agarose gel on which RNA samples from the two strains were loaded, below these are the autoradiographs of the subsequent northern blots probed with the strain 151 pucBA genes. At the very bottom of the figure are two charts showing the A_{650nm} at the time the cells were harvested.

Figure 56

Whole cell absorption spectra of strain 151 samples from which RNA was isolated. The spectra have been smoothed using a 7 point moving average and normalised to an absorbance at 650 nm of 1 per cm. The y axes have been removed from the central graphs for clarity, but are scaled to the same size allowing direct comparison.

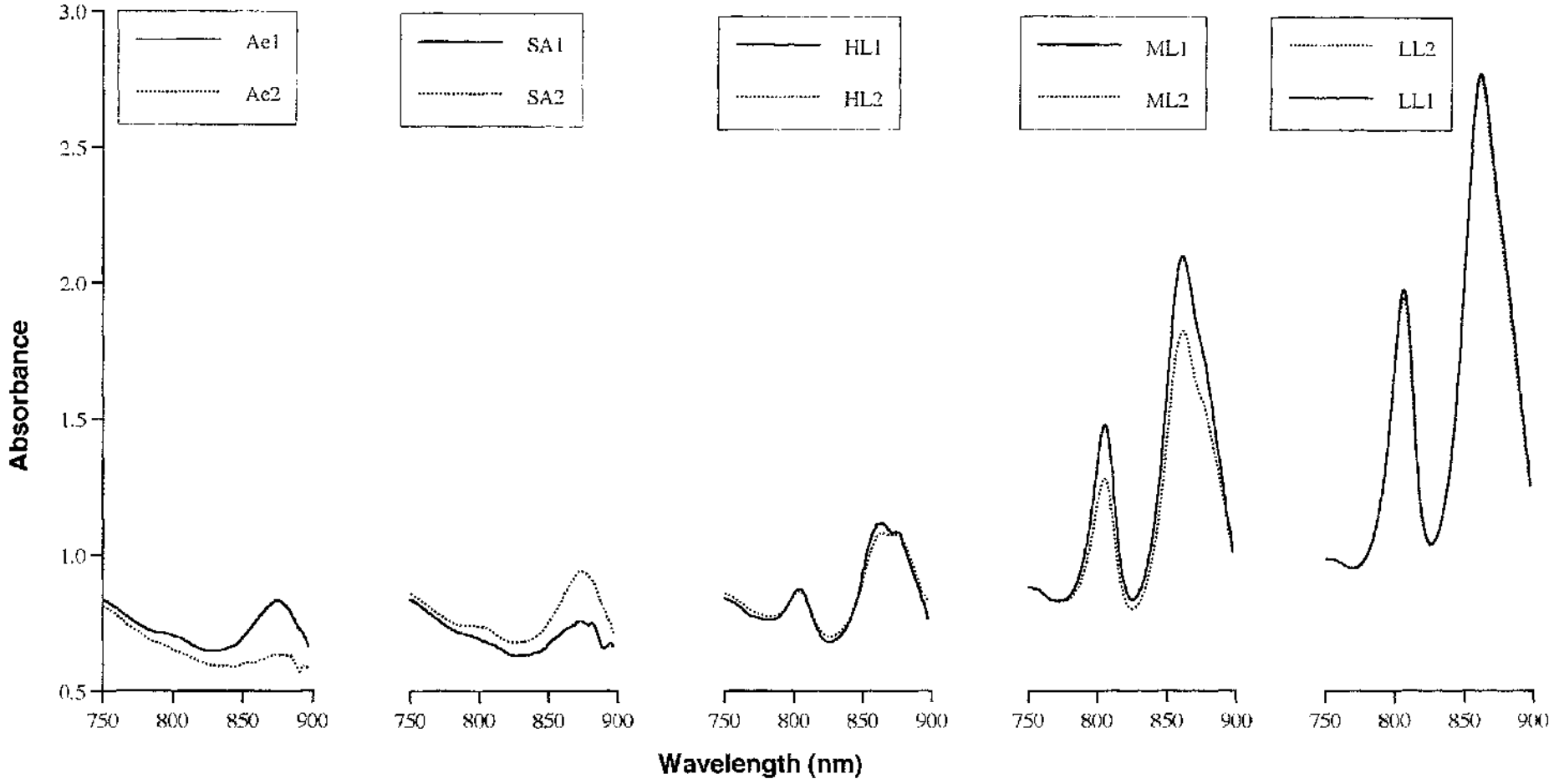
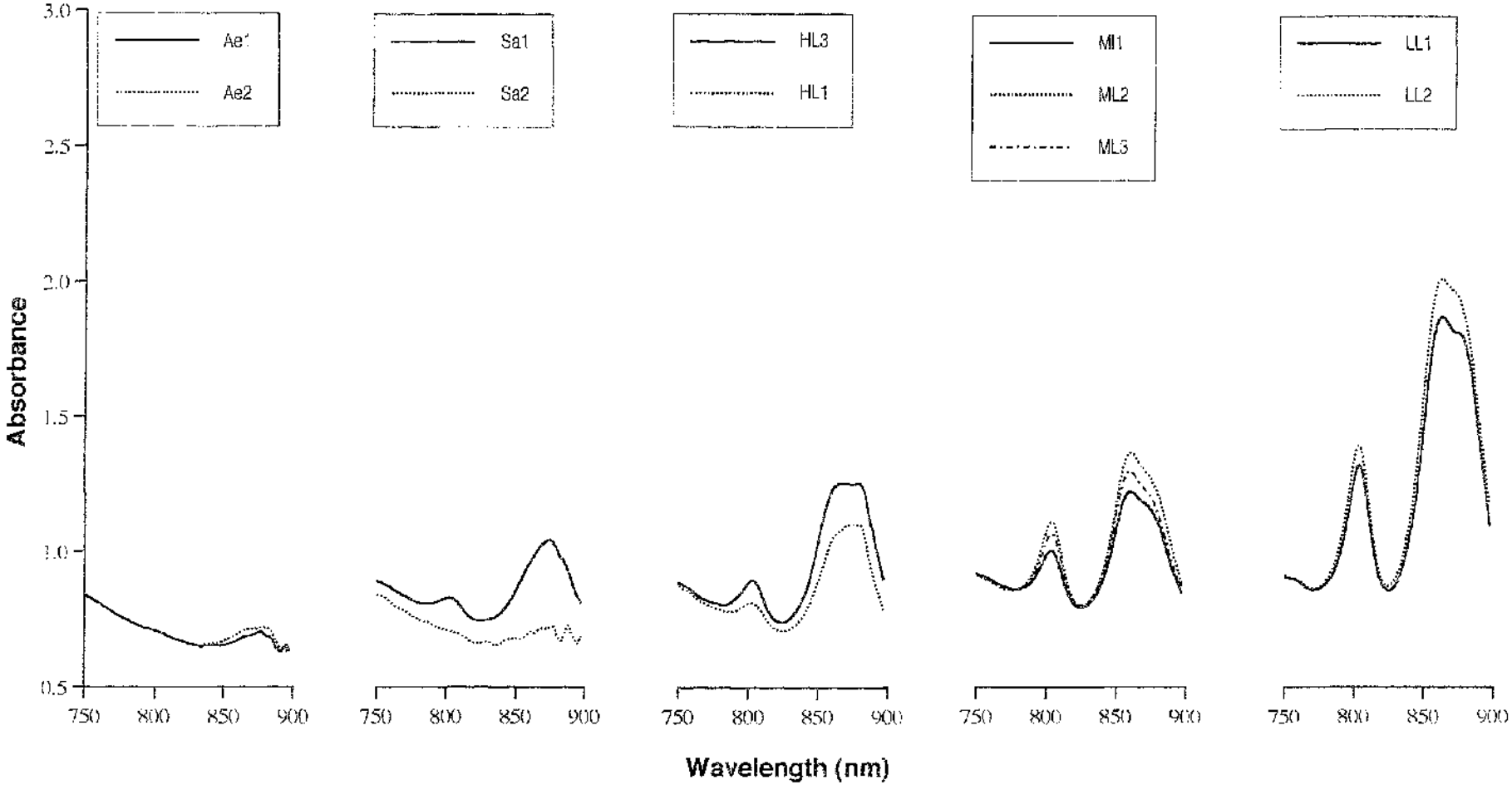


Figure 57

Smoothed whole cell absorption spectra recorded from the strain 149 samples used in RNA preparation. The samples are all normalised to an absorbance at 650 nm of 1 per cm. The y axes of the central graphs have been omitted for clarity but are scaled to the same size allowing direct comparison of all graphs pictured.



From the latter it is immediately apparent that there was a wide variation between the different cultures in terms of their cell density at harvest. Looking at the ribosomal RNA bands visible on the agarose gels it is also apparent that less RNA was loaded for the strain 149 samples that was loaded for the strain 151 samples (the rRNA bands of strain 149 are fainter). These two sources of variation suggest that the Northern hybridisation signals obtained for each sample must be interpreted with caution, and makes comparison of puc expression between the two strains difficult.

Despite the difficulty in comparing the expression of the strain 149 and 151 puc genes using the data collected, several useful observations can still be drawn from the data. Firstly if one considers the absorption spectra recorded from the HL, ML and LL samples (Figure 56 and Figure 57), it is apparent that they appear similar to those recorded from cells cultured under the same conditions and presented in the Results 1 section. The two strains respond as described previously, with the the LH2 peak at 861 nm increasing in size as the light intensity decreases. The difference between the two strains is also visible again, the strain 151 spectra always showing the LH2 peak at 861 nm dominating the LH1 peak at 881 nm. Whilst in the strain 149 spectra the LH1 peak is always visible, dominating that of LH2 at LL. The fact that the spectra of cells cultured at each irradiance level have an absorption spectra characteristic of that lighting condition, suggests that conditions which induce differential LH2 synthesis were again successfully established.

The Northern analysis part of the experiment appears to show that both strains posses a single puc transcript estimated to be about 600 bp in size. It also appears to show that puc expression in each strain is reduced in the presence of oxygen, since the hybridisation signals are much reduced in the aerobic and semi-aerobic cultures compared to the light grown cultures. If one accounts for the differences in RNA loaded onto the gels and ignores those cultures grown to an exceptionally high cell density (such as 149 ML2), it also appears that the level of puc transcription in both strains increases as the irradiance level is decreased. This transcriptional response thus mirrors the level of LH2 synthesised.

The absorption spectra recorded from cells cultured under aerobic and semi-aerobic conditions also allow some interesting observations to be made. Most importantly even cultures which were not grown to an excessively high cell density such as strain 151 Ae1 (A_{650nm} ~0.07 cm^{-1}), SA1 (~0.04 cm^{-1}) and SA2 (~0.05 cm^{-1}), produced measurable levels of LHC's. These samples also had visible levels of transcription as can be seen on the Northern blots. This either suggests that *Rv. gelatinosus* puc expression is repressed by oxygen to a lesser extent than that of other species, or that the conditions used did not cause a great enough aeration of the cultures. The semi-aerobic culture condition in particular could quickly be changed as the bacteria multiply and respire away any oxygen present. Although it is difficult to distinguish between these two possibilities, there may be a distinguishable difference between the two strains as regards the repression of LH2 synthesis by oxygen. Comparing the flat strain 149 Ae1 spectra with the strain 151 Ae1 spectra (which has a visible peak at 881 nm) it is clear that strain 151 has synthesised significantly more LHC than strain 149, in spite of the fact that 149 Ae1 was grown to over twice cell

density of 151 Ae1. The absorption spectra of the strain 151 Ae2 sample (grown to a lower cell density, A_{650nm} = 0.032 cm^{-1}) in contrast is similar to that of the strain 149 samples. This suggests that at high oxygen tension puc expression in both strains is repressed, but that strain 151 puc expression is derepressed at a higher tension than that of strain 149.

A second observation that can be made regarding the aerobic and semi-aerobic absorption spectra is that LH1 predominates the spectrum. This is true even in strain 151 where LH2 usually predominates the NIR absorption spectrum. These cultures are essentially undergoing a transition from aerobic to anaerobic growth however and thus represent a special case.

The difficulty in comparing the results obtained for each strain can be seen if one considers the duplicate cultures LL1 and LL2 from each strain, which were loaded towards the right-hand side of the gels in Figure 55, the hybridisation signal appears greater in strain 151 than strain 149. Considered in isolation this would suggest that strain 151 has a more abundant puc transcript than strain 149, an observation which would support the conclusion that strain 151 synthesises more LH2 than strain 149. However when one takes into account the extra strain 151 RNA that appears to have been loaded onto the gel, the difference between the strains is much less convincing. The result is further drawn into question by the difference in cell density at harvest, with the strain 151 cultures having been grown to approximately three times the density of the strain 149 cultures.

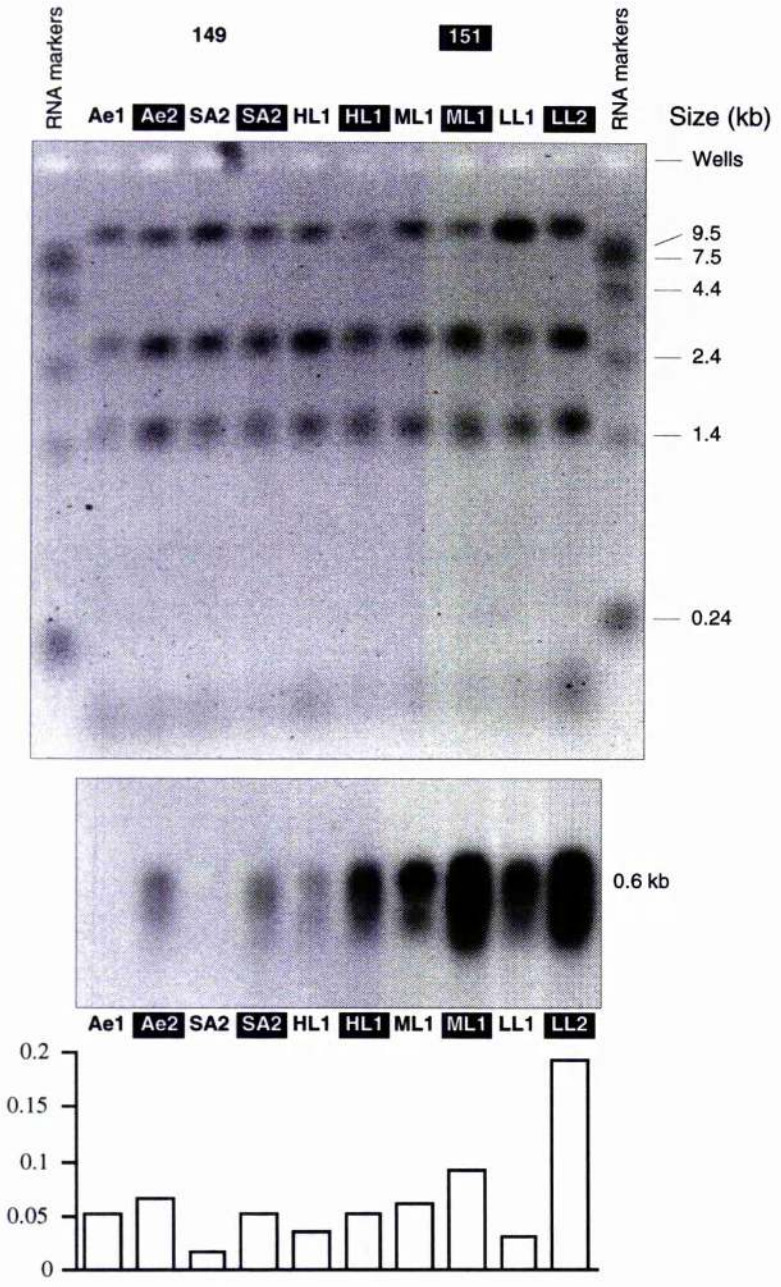
A similar picture is seen when the ML and HL culture conditions are examined. The strain 149 ML1 and ML3 samples both have a weaker hybridisation signal than the strain 151 ML1 and ML2 samples, but in each case less strain 149 RNA appears to have been loaded. Interestingly ML2 samples from both strains appear evenly loaded on the gel, and have apparently similar hybridisation signals, indicating a very similar level of puc expression. Once again the result can be questioned however, because the strain 149 culture was grown to almost three times the density of the strain 151 culture. Analysis of the HL cultures from each strain produces similarly inconclusive results.

The dark aerobic and semi-aerobic cultures are also difficult to compare across the strains. The strain 149 semi-aerobic 1 (SA1) sample appears to have an abnormally high hybridisation signal, this probably results from the high cell density to which the culture was grown. Once the SA cultures reached a certain density they could potentially respire away the oxygen present, creating an anaerobic environment, and allowing the expression (derepression) of the puc genes. This appears to have occurred in the strain 149 SA1 culture, thus this sample can be discounted from the comparison. The remaining dark grown cultures show the same pattern of less strain 149 RNA, and weaker strain 149 hybridisation signals (although no signal is visible in the strain 149 autoradiograph shown in Figure 55, a longer exposure did indicate a low level of hybridisation in the Ae1, Ae2 and SA2 samples).

A second attempt at comparing the RNA samples was made, running the equivalent samples from each strain adjacent to each other on an agarose gel. The resulting gel, autoradiograph and the A_{650nm} values at harvest are shown in Figure 58 (overleaf).

Figure 58

An agarose gel and associated Northern blot showing RNA samples isolated from strains 149 and 151 of *Rv. gelatinosus* cultured under a range of conditions.

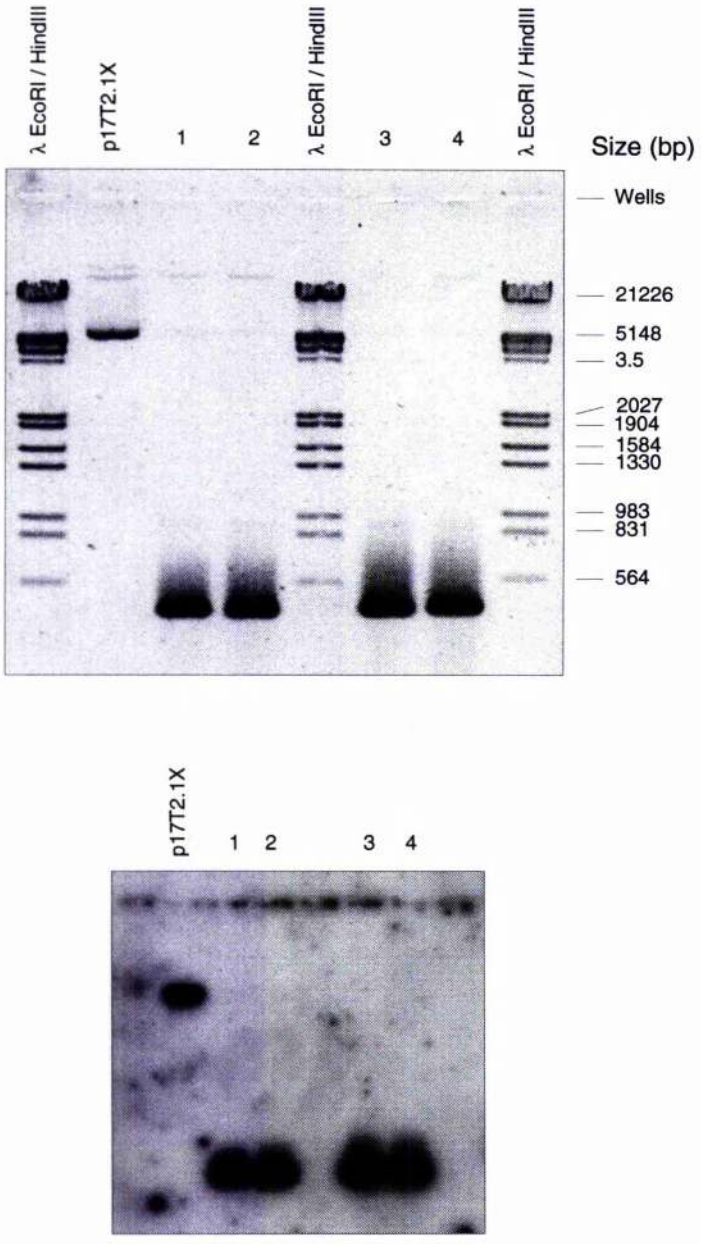


Note: Pictured at the top is a 1.4% agarose gel containing strain 151 (white text on black) and strain 149 (normal black text) RNA samples cultured under a variety of conditions. The abbreviated sample name is given above each lane with strain 151 as white text on black and strain 149 as normal black text, the main text has details of the full name. An RNA ladder was loaded at each side of the gel, the sizes of the fragments are given to the right. Below the gel is a Northern blot taken from the gel and probed with the pucBA genes from strain 151. The estimated size of the hybridising transcript is given at the side of the autoradiograph. At the very bottom is a chart indicating the whole cell A_{650nm} at cell harvest (cm^{-1}) for each of the samples loaded.

Again if one considers the autoradiograph in isolation it appears that strain 151 has a higher puc transcription level than strain 149 in every culture condition tested. And again the A_{650nm} values call this interpretation into question, since all the strain 151 cultures were grown to a higher density than the equivalent strain 149 cultures. It also seems likely that there is variation in the level of RNA loaded onto the gel, but this is difficult to see because unfortunately the photograph of the gel is out of focus. However this second gel does appear to support the conclusion that *Rv. gelatinosus* responds to the presence of oxygen by suppressing puc transcription, and also that more transcript is present in cultures grown at LL, than at ML or HL, since there is a visible increase in hybridisation signal as one looks from the left to the right side of the autoradiograph.

Figure 59

An agarose gel and associated Northern blot showing the PCR product obtained from the *Rv. gelatinosus* strain 151 clone p17T2.1X.



Note: Pictured at the top is a 1.4% agarose mini-gel containing the PCR products from 4 reactions designed to amplify the strain 151 pucBA genes. The template for the reaction was the plasmid p17T2.1X which has also been loaded on the gel for comparison. Samples 1 and 2 contained 220 ng of template and samples 3 and 4 110 ng. Samples 2 and 4 also contained more primer (0.75 μM), compared to 1 and 3 (0.5 μM). DNA markers were loaded, these were produced by digesting phage λ DNA with EcoRI and HindIII, the sizes of the fragments are given to the right (160 ng were loaded). Below the gel is a Southern blot taken from the gel and probed with the α consensus oligonucleotide.

PCR of the strain 151 pucBA genes and their expression in *Rb. sphaeroides*

As part of a collaborative project with Dr. Greg Fowler at the University of Sheffield a fragment of DNA carrying only the pucBA genes from strain 151 was isolated and cloned into an expression vector. This facilitated the expression of the *Rv. gelatinosus* pucBA genes in *Rb. sphaeroides*. The fragment carrying the strain 151 pucBA genes was isolated from p17T2.1X using the polymerase chain reaction. Two oligonucleotide primers, of 30 and 28 residues were synthesised. These were homologous to the DNA sequence up and downstream of the pucBA genes, but mutated the DNA outside of the gene reading frames to include a KpnI restriction site upstream of the genes and a BamHI site downstream. These introduced restriction sites enabled the PCR product to be inserted into a plasmid expression vector in the correct orientation.

The results of the PCR reaction are illustrated in Figure 59 (opposite). Four separate reactions were performed using different amounts of primer and template DNA, this appears to have had little effect on the reaction yield however. The size of the PCR product should be 448 bp, since the products run below the 564 bp λ marker the correct fragment appears to have been amplified. To further confirm that the pucBA genes were present on the PCR product the gel was Southern blotted and probed with the a consensus oligonucleotide. The resulting autoradiograph pictured at the bottom of Figure 59 (opposite) shows a strong hybridisation signal for the PCR product. The template plasmid p17T2.1X can also be seen to hybridise to the probe as would be expected.

The PCR product was cloned into an expression vector and transformed into a deletion mutant (DD13/G1) of *Rb. spheroides* which synthesised no LH2 or LH1-RC complexes by Dr. Fowler. This allowed the fluorescence emission and excitation spectra, plus the absorption spectra of the expressed strain 151 LH2 to be recorded without interference from the native *Rb. sphaeroides* LHCs. The spectra recorded confirmed that the LH2 complex synthesised was that of *Rv. gelatinosus* (with a different carotenoid present), as did western blot analysis using antibodies to the *Rb. sphaeroides* LH2 complex. The full results are presented in Fowler *et al.*, (1995).

Conclusions

In both strain 151 and 149 there appears to be a transcriptional response to the different irradiance levels used, and to the presence of oxygen. The hybridisation signals obtained for the LL grown cells were stronger than those obtained for ML, which in turn were stronger than those obtained for HL. These signals thus mirror the level of LH2 synthesised in each irradiance condition. They also suggests that *Rv. gelatinosus* responds to changes in light intensity in a manner similar to other purple bacterial species, which show a relative increase in the level of puc mRNA as the irradiance is decreased. Similarly the aerated cultures show a reduced hybridisation signal and thus reduced puc transcript in comparison to the light grown cultures, indicating that the presence of oxygen also suppresses the transcription of the puc genes in *Rv. gelatinosus*. The absorption spectra recorded from the aerated cultures also show little LH2 was synthesised, except when the cultures were 'overgrown'.

The aerated cultures which were grown to a higher cell density than intended also provide some useful insights, since they are undergoing a transition from aerobic to anaerobic growth. Both the hybridisation studies and the whole cell absorption spectra show that synthesis of LHCs has begun in these cultures, but although puc transcription is occurring LH1 dominates the absorption spectra. This suggests that LH1 is synthesised before LH2 during shift to anaerobiosis, an observation that has been made for other species of photosynthetic purple bacteria.

Although it is not clear if a difference exists between the two strains as far as puc transcription is concerned (from the Northern analysis experiment), the absorption spectra do show (as was illustrated in the first results section) that strain 151 synthesises more LH2 than strain 149. Thus these new spectra confirm the results obtained in the previous experiment. The spectra recorded form the aerated cultures hint at a further difference between the two strains. This can be seen if one considers the aerobic cultures (Ae1) of strain 151 and 149. The strain 151 Ae1 culture has both a measurable level of puc transcription and significant amounts of LHC, this should not be the case in the presence of oxygen. Thus one is drawn to the conclusion that this culture reached a cell density at which the cells began to respire more oxygen than could diffuse into the growth medium, causing a drop in oxygen tension and allowing expression of the puc and puf genes. Since this culture reached an A_{650nm} of 0.067 cm^{-1} and the strain 149 Ae1 culture was grown to 0.16 cm^{-1}, almost twice the density, one would expect strain 149 to have reacted in a similar manner, transcribing the puc and puf genes and synthesising LHCs. This does not appear to be the case, the strain 149 Ae1 absorption spectra shows little or no LHC and there is no visible puc transcript on the autoradiograph. An autoradiograph with longer exposure time did show a very faint hybridisation in the strain 149 Ae1 lane, but this appeared much less than the apparent difference in RNA loaded between the two strains. Thus one can tentatively conclude that strain 151 is able to transcribe photosynthesis genes and synthesise LHCs at higher oxygen tensions than strain 149. Further experimentation, in particular the growth of the two strains at several different but known oxygen tensions would be needed to confirm this observation.

It is not clear whether this observed strain difference in response to oxygen

tension is related to the different levels of LH2 synthesised under different irradiance conditions, or whether it instead forms an additional difference in the regulation of LH2 expression between the two strains.

A more solid conclusion that can be drawn from the northern analysis is that there is a single puc transcript approximately 600 bp in size. Although multiple and often larger transcripts have been observed in other species, the observed transcript correlates with the organisation of the the *Rv. gelatinosus* puc genes found during sequencing. A longer pucBA containing transcript could not be formed in *Rv. gelatinosus* because the pucC gene is present downstream of pucA in the opposite orientation. Read-through from the pucBA promotor would thus generate anti-sense RNA and interfere with expression of the pucC gene, a gene which appears to be essential for LH2 synthesis. However two potential promotors were identified upstream of pucB and two potential terminators downstream of pucA. If these promotors and terminators were being used by the bacteria one would expect to see some variation in the transcript size, this is not the case. It may be however that the resolution of the agarose gel does not allow one to distinguish between the different sized transcripts, and that the use of polyacrylamide gel electrophoresis would allow one to distinguish between them.

The results of the Northern analysis carried out to allow comparison of puc expression between *Rv. gelatinosus* strain 151 and strain 149 proved too variable to allow any solid comparisons to be made. This variation appears to stem from two sources, firstly from growth of cultures to widely varying cell densities and secondly because varying amounts of RNA were loaded onto the agarose gels. The point at which harvesting was carried out was judged by eye in the same manner used for the initial culturing of the two strains at different irradiance levels, as described in the Results Part 1 section. In the previous experiment the cultures were harvested at much more even cell densities, but judging the cell density by eye appears to have been less effective when harvesting the cells for RNA extraction. A more accurate method of determining the cell density would remove this source of variation. This could perhaps be achieved by measuring the absorbance at 650 nm of the culture within the bottle, or by removing aliquots from the culture and determining the absorbance during growth.

The low cell density at which cell harvest was carried out also caused problems. Although harvest at A_{650nm} of around 0.05 cm^{-1} to avoid shading worked well in terms of the absorption spectra recorded, each irradiance level inducing a characteristic absorption spectra, it meant that a small amount of RNA was obtained for each culture. This in turn did not allow accurate determination of RNA concentration for several cultures because a limited amount of sample was available. The RNA concentration of the samples was determined by spectroscopic methods, but since the absorbance at 260 nm was less than 0.1 for 5 of the 20 samples, their calculated RNA concentration may have been inaccurate. This could in part explain the variation in RNA loaded onto the agarose gels for Northern analysis.

The PCR reaction carried out to isolate the stain 151 pucBA genes worked well, producing a high yield. That this fragment contained the pucBA genes was shown by probing the PCR products on a blot with the α consensus oligonucleotide. The

fragment was used extensively as a probe for the pucBA genes of *Rv. gelatinosus* in both the strain 149 library construction and the Northern analysis.

The heterologous expression of the stain 151 pucBA genes in *Rb. sphaeroides* shows that it is possible to synthesise an active LH2 complex in a different species of photosynthetic bacterium. It further shows that the light harvesting complex can contain different carotenoid molecules and that the assembly process must be similar in the two species. This also provides a method of carrying out site directed mutagenesis studies on the *Rv. gelatinosus* LH2 complex without having to create deletion mutants of *Rv. gelatinosus*.

Discussion

The aims presented at the start of this thesis were to confirm the observation that *Rv. gelatinosus* strain 151 synthesised more LH2 complex than strain 149. To clone and sequence the puc genes from the two strains. And finally to examine expression of the puc genes in each strain to investigate whether differences in LH2 synthesis are reflected in different transcriptional levels. In this discussion I will first describe to what extent these aims were achieved and how the work could be extended or improved. Then move on to describe some of more general observations and conclusions which can be derived from the results presented.

The observation at the root of this thesis was the noting of a difference in the absorption spectra of the two strains 149 and 151, which appeared to show that strain 151 was synthesising more LH2 relative to RC-LH1 than strain 149. The first step was to confirm that there was a real difference between the two strains. The growth of the two strains at three different irradiance levels showed that there was indeed a difference between them. Strain 151 appears to synthesise approximately one third more LH2 than strain 149 at any given irradiance level. Interestingly the level of LH2 in both strains varied by a similar amount when they were cultured under the different irradiance levels. At ML both strains synthesised approximately 10% more LH2 than at HL, whilst at LL the there was a two-fold increase in the level of LH2. Thus the difference between the strains is not that 149 is less responsive to light changes, but that strain 151 synthesises a larger amount of LH2 relative to RC-LH1. There must clearly be a regulatory difference between the two strains which underlies this difference. The most obvious point for this difference to occur is at the transcriptional level, with lower expression of the puc genes from strain 149. However the difference may also occur at the post-transcriptional level, or may be a global regulatory difference.

Since a global regulatory difference would probably result in reduced expression of the LH1 complex as well as LH2, this possibility could perhaps be resolved by measuring the level of LH1 synthesised under various irradiance levels. This would prove difficult by absorption spectroscopy because LH2 tends to swamp absorbance from LH1 (particularly in strain 151). However one could also confirm the results by isolation and detergent solubilisation of chromatophore membranes, followed by discontinuous sucrose density gradient centrifugation which would separate LH1 and LH2. SDS-PAGE of chromatophore membranes could also prove useful, allowing one to visualise the physical amount of LH1 and LH2 present. This would also confirm the results of the absorption spectroscopy work by showing the quantitative difference in LH2 between the strains in a more physical manner.

As a preliminary to investigating puc gene expression genomic libraries were constructed from the two strains and the puc genes cloned and sequenced. This could also allow the identification of control elements in the DNA which were different between the two strains and thus potentially responsible of differential LH2 synthesis. The construction of genomic libraries did involve a few minor technical problems (such as the problem with mixing during Sau3AI digestion, and the lower efficiency achieved for the strain 151 library), but these did not prevent the isolation of λ clones and

subsequent cloning of the pucBA genes from each strain. The source of the cloned fragments was confirmed by using oligonucleotide primers with potential homology to the puc genes and surrounding region to perform DNA sequencing reactions, and by Southern analysis.

Full sequencing of the cloned pucBA carrying fragments highlighted a startling difference between *Rv. gelatinosus* and other species from which the puc operon has been cloned. This concerned the arrangement of the puc genes, in other species where the pucC gene has been found associated with the pucBA genes it usually lies in the same orientation and is transcribed from the same promotor (or promotors) as the pucBA genes (i.e. it lays in the same operon). DNA sequence from both strains indicates that the pucC gene is present in *Rv. gelatinosus*, and also that it lies downstream of pucBA, but in contrast to other species lies in the opposite orientation to pucBA and is thus part of a separate operon. This puc gene arrangement could be the result of genetic rearrangements, or could represent an ancestral arrangement of the puc genes which the other species have modified.

In contrast to the DNA sequencing Southern analysis using the *Rb. sphaeroides* pucC gene indicated the presence of pucC in strain 151, but not strain 149. The lack of a strongly hybridising fragment in strain 149 is puzzling, since a third of the gene was sequenced and should hybridise. Either the Southern hybridisation was flawed in some way or there is a difference between the two strains in the 5′ region of pucC (the 3′ region was sequenced in both strains). This is interesting since pucC is known to be intimately involved in the synthesis of LH2 in other species, if correct this difference could contribute to the reduced LH2 levels observed for strain 149. Probing strain 149 genomic DNA and the isolated strain 149 λ clones with a fragment carrying the pucC gene from strain 151 would probably clarify this point.

Potential promotor elements and repressor binding sites were identified upstream of the pucBA genes in both strains, these were highly conserved between the two strains. Thus it was not possible on the basis of sequence analysis alone to identify elements which could be involved in puc gene regulation and that differ between the strains. Further work, starting with the mapping of transcript 5′and 3′ end points to check that the identified promotors are being used, and followed by assays designed to identify DNA binding proteins which bind to the pucBA upstream region would necessary to identify any differences present.

Such work would have to follow comparative northern analysis of puc expression in the two strains, this was attempted as described in the Results 5 section, but the results proved to variable to allow comparison between the two strains. The variation appears to have stemmed form differences in cell harvesting time. This could probably be overcome if the growth of cultures was monitored, and harvest at a specific cell density more accurately achieved. A repeat of the northern analysis experiment could then show either that there was a reduced level of transcript in the strain 149 samples, or that the two strains had essentially the same level of transcription. If the latter proved to be the case then work would next have to focus on post-transcriptional control of LH2 expression.

Despite the difficulty in comparing the northern analysis results of strain 149

with those of strain 151, several useful observations could still be made regarding the data. Firstly it appears that *Rv. gelatinosus* puc expression is affected by oxygen in a manner similar to other species of photosynthetic bacteria, since aerobic culture reduced the levels of puc mRNA found. The same can be said with regard to the response to irradiance levels, as with other species the level of puc transcript increased as the irradiance level was decreased. This response clearly has a part to play in regulating the levels of LH2, since the levels of LH2 also increase as irradiance level is decreased.

The aerobic cultures of strain 151 and 149 which were (unintentionally) overgrown also showed an interesting response. At high cell density the cultures respire away the oxygen and begin to synthesise LHCs, but interestingly whilst an aerobic culture of strain 151 began puc transcription and to synthesise LH2 at an A_{650} of 0.067 cm^{-1}, a strain 149 culture grown to 0.16 cm^{-1} failed to begin LH2 synthesis. Since the 149 culture was at approximately twice the cell density it must have also have reached a point at which it began to respire away the oxygen, yet it did not respond in the same way as strain 151. This suggests a difference between the two strains in terms of the sensitivity to oxygen tension. Whilst it would be premature to conclude an absolute difference from this single replicate the point would warrant further investigation. This could be achieved by to monitoring and controlling oxygen concentration more tightly, and to culture both strains at a series of oxygen tensions to investigate the observed difference in response to oxygen tension. It could also be investigated by allowing an aerobic culture to overgrow, and removing aliquots for RNA isolation at intervals. The most important point would be to ensure that the oxygen tension would be maintained at the same level in each culture. It is not clear whether this difference is a separate one from the previously identified reduced LH2 level of strain 149, or whether it marks an additional difference.

Further interesting points concern the potential regulatory elements identified in the DNA sequence up and downstream of the pucBA and C genes. The most strikingly conserved feature between *Rv. gelatinosus* and other strains is the putative PpsR repressor binding sites which occur in the promotor region of the pucBA genes and at the start of the pucC gene. Similar sites have been found upstream of many other photosynthesis genes in *Rb. capsulatus*, *Rb. sphaeroides* and *Rps. palustris*. The identification of these sits in *Rv. gelatinosus* suggests that the protein is highly conserved in photosynthetic bacteria, probably because of its role in repressing synthesis of photosynthetic apparatus components in the presence of oxygen. There must be a strong selective pressure against the potentially lethal combination of oxygen, light and bacteriochlorophyll occurring.

The use of σ^{70} like promotors has also been seen in other photosynthetic bacteria. The presence of two promotors upstream of pucB produces some interesting possibilities, it presumably allows the bacterium to impose differential regulation on the two promotors. One could speculate that in *Rv. gelatinosus* the promotor more proximal to pucB would be expressed constitutively in the absence of oxygen, whilst the distal promotor could be regulated by light. Both promotors would be repressed by PpsR since it would bind over the proximal promotor. This would of course require experimental conformation, firstly by transcript end mapping to see if both promotors

were being used, and secondly to see if different promotors were utilised under different culture conditions.

Another conserved feature between *Rv. gelatinosus* and other photosynthetic bacteria is the presence of exceptionally large and stable stem-loop forming structures immediately downstream of the pucA genes. These appear to play a role in transcription termination and in stabilising the transcript against degradation by endonucleases.

The predicted protein sequence of the strain 151 pucC gene shows a great deal of homology to other pucC genes and to the homologous genes orf477 and G115. This family of proteins also shows some potential structural homology. It appears that the pucC protein has two main domains separated by a polar hydrophilic section, and 12 membrane spanning segments with its N and C terminii both located on the cytoplasmic side of the membrane. It is perhaps not surprising that the structure of these proteins is so well conserved when one considers the heterologous expression experiments. Since the pucC protein of *Rb. sphaeroides* appears to substitute for the native *Rv. gelatinosus* pucC and allow assembly of the *Rv. gelatinosus* LH2 in *Rb. sphaeroides*.

To summarise it appears that the molecular biology underlying synthesis of LH2 in *Rv. gelatinosus* shows an overall similarity to other photosynthetic bacteria, but also some notable differences. It has also been shown that there is a definite difference between the two strains with regard to LH2 synthesis, but further work is required to delineate at which stage of the synthesis process this difference arises.

Appendices

Appendix 1: Media used for the culture of *Rv. gelatinosus*

Modified from 'Semi-synthetic medium' as described in Subir, (1963).

1.1 c-Succinate Medium

per litre:

Base Concentrate*	20 ml
1M Di-potassium hydrogen orthophosphate	10 ml
1M Potassium di-hydrogen orthophosphate	10 ml
10% Ammonium sulphate solution	5 ml
1M Sodium succinate, pH 6.8	10 ml
Growth factors*	1 ml
Casamino acids	1 g

made up to 1 litre with deionised water.

* Constituents of these solutions are described below.

1.2 Base Concentrate

per litre:

nitrilotriacetic acid	10g
magnesium sulphate	14.45g
calcium chloride.$2H_2O$	3.4g
ammonium molybdate	9.25mg
ferrous sulphate.$7.H_2O$	99mg
nicotinic acid	50mg
Aneurine hydrochloride	25mg
Biotin	0.5mg
Metals 44*	50ml

* constituents are described in appendix 1.4

1.3 Growth factors

per litre:

Biotin	0.02g
Sodium hydrogen carbonate	0.5g

Make up to 1 litre with deionised water to dissolve, then add:

Nicotinic acid	1g
Aneurine hydrochloride	0.5g
4 Aminobenzoic acid	1g

and boil to dissolve.

1.4 METALS 44

Ethylenediaminetetra-acetic acid (EDTA)	250mg
Zinc sulphate	1095mg
Manganous sulphate.$4H_2O$	154mg
Copper sulphate.$5H_2O$	39.2mg
Cobaltous nitrate.$6H_2O$	24.8mg
Ferrous sulphate.$7H_2O$	500mg
Di-sodium tetraborate.$10H_2O$	17.7mg

Make up to 100 ml with deionised water and add 2 drops concentrated sulphuric acid to retard precipitation.

Appendix 2: Solutions used in RNA extraction

2.1 Solution D

Stock solution:

To 100g Guanidium thiocyanate in a glass bottle (as supplied by The Sigma chemicals company) add:

DEPC treated, Deionised water	117.2 ml
0.75M Sodium citrate pH 7.0	7.04 ml
10% Sarkosyl	10.6 ml

Heat to 65°C to dissolve. The stock solution is stable for 3 months at 4°C, before use 0.1% β–mercaptoethanol is added to produce solution D.

2.2 Water saturated phenol

1 kg of phenol (Fisons, AR grade) was melted in a water bath at 68°C, and 8–hydroxyquinoline added to 0.1% of the volume. 150 ml of sterile Tris-HCl pH 8.0 was added and mixed, then allowed to settle, following which the aqueous phase was discarded. This was repeated with the addition of a further 150 ml of Tris-HCl. The aqueous phase was then replaced with sterile water, and β-mercaptoethanol added to 0.2% (volume of phenol phase), before storage at 4°C in a foil covered glass bottle.

Appendix 3: Production of radiolabelled probes

3.1 5′ end labelling of oligonucleotides

Mix the following ingredients in a microfuge tube on ice, make up to 50 µl total volume with sterile analar water, and incubate for 30 minutes at 37°C.

Oligonucleotide	(200ng)	~1 µl
50mM $MgCl_2$		10 µl
1M Tris-HCl, pH 7.6		5 µl
50 mM DTT		5 µl
γ^{32}P-ATP	@ 5000 Ci µmol^{-1}	10µl
10 units T4 kinase		1µl

3.2 Random prime labelling

This was carried out using the following protocol, based on that of Feinberg & Voglestein, (1983). DNA solution should be boiled for 5 minutes and cooled on ice, before adding the following:

Reaction Mix*		10 µl
BSA (10 mg ml^{-1})		2 µl
DNA solution$^{\Delta}$		2 to 8 µl
α^{32}P-dCTP [ICN]	(50 µCi)	5 µl
Large Fragment polymerase I (Klenow) [Boehringer-mannheim]		1 µl
sterile analar water		to 50 µl total

*The constituents of this solution are given below.

$^{\Delta}$Typically, 50ng of DNA was used in this reaction.

The reaction was either allowed to proceed for 1 hour at 37°C, or left at room temperature overnight. Probes prepared using this protocol regularly have specific activities greater than 10^9 cpm/µg of DNA (MacKenzie, 1990).

3.3 Solutions for Reaction Mix

For 1 ml of Reaction mix add:

Solution A	200 µl
Solution B	500 µl
Solution C	300 µl

Solution A

Solution D	1µl
β-mercaptoethanol	18µl
100 mM dATP	5µl
100 mM dTTP	5µl
100 mM dGTP	5µl

Solution B

2M HEPES pH 6.6

Solution C

Hexadeoxynucleotides [Pharmacia] at 9000 units ml^{-1} in TE pH 8.0

Solution D

1M Tris-HCl pH 8.0

0.125M $MgCl_2$

Appendix 4: Genotypes of bacterial strains

4.1 KW251

KW251 F$^-$, *sup*E44, *gal*K2, *gal*T22, *met*B1, *hsd*R2, *mcr*B1, *mcr*A, [*arg*A81:Tn10] *rec*D1014

4.2 DH5α

DH5α *sup*E44, ΔlacU169, (φ80lacZΔM15), *hsd*R17, *rec*A1, endA1, gyrA96, thi-1, relA1

Appendix 5: Sau3AI digestion mixtures

5.1 Enzyme premix		5.2 DNA premix	
sterile analar water	44 µl	sterile analar water	190 µl
10x React 4 buffer [BRL]	5 µl	10x React 4 buffer	50 µl
Sau3AI [BRL] @6 u μl^{-1}	1 µl	genomic DNA @460 ng μl^{-1}	260 µl

The concentration of enzyme to use in the enzyme mix was calculated on the basis of the frequency at which Sau3AI cuts a random DNA sequence (of neutral base composition). This is once every 250 bp, since we require 20 kb fragments we only wish to cut 1/80 th of the potential sites (20000 ÷ 250 = 80). Since 1 unit of enzyme fully digests 1 µg of DNA in one hour, it can be calculated that the 6 units of Sau3AI used in the enzyme mix will cut 1/80 th of the sites in the 119 µg of DNA added in ~4.5 minutes. This provides a useful time scale around which the digestion experiment can be arranged.

5.3 Scaled up enzyme mix		5.4 Scaled up DNA mix	
sterile analar water	43 µl	sterile analar water	650 µl
10x 1-4-all buffer [Pharmacia]	5 µl	10x 1-4-all buffer [Pharmacia]	150 µl
Sau3AI [Pharmacia]@10 u μl^{-1}	2 µl	genomic DNA @460 ng μl^{-1}	650 µl

Although a different suppliers enzyme was used in the scaled up digest this is not a essential change, Sau3AI from any supplier could be used. The reader should also note that the volumes presented were used for both strains 151 and 149, but that the concentration of genomic DNA was slightly lower for strain 149 (405 ng μl^{-1}).

Appendix 6: Constituents of the Ligation Reactions

6.1 Genomic library ligation

genomic DNA (~20 kb fragments)@65 ng μl^{-1}	6µl
λGEM-11 Arms [Promega Corp.] @500 ng μl^{-1}	1µl
Ligase buffer (10x)*	1µl
ATP @ 10mM	1µl
T4 Ligase [BRL] @1 u μl^{-1}	1µl

*Constituents given in Appendix 6.4

6.2 Test insert ligation

pTIII test insert [Promega Corp.] @500 ng μl^{-1}	1µl
λGEM-11 Arms [Promega Corp.] @500 ng μl^{-1}	1µl
Ligase buffer (10x)*	1µl
ATP @ 10mM	1µl
T4 Ligase [BRL] @1 u μl^{-1}	1µl
Sterile Analar water	5µl

*Constituents given in Appendix 6.4

6.3 Ligation of λ arms

λGEM-11 Arms [Promega Corp.] @500 ng μl^{-1}	1µl
Ligase buffer (10x)*	1µl
ATP @ 10mM	1µl
T4 Ligase [BRL] @1 u μl^{-1}	1µl
Sterile Analar water	6µl

*Constituents given in Appendix 6.4

6.4 10x Ligase buffer

$MgCl_2$	100mM
Tris-HCl, pH 7.5	400mM
DTT	100mM
ATP	10µM
Bovine Serum Albumin	500 µg ml^{-1}

For each ligation reaction the DNA solutions, water and ligase buffer were mixed and incubated at 42°C for 1 hour then cooled on ice before the other ingredients were added, ligation was carried out at 18°C for 48 hours.

Appendix 7: SM buffer (for the storage of phage)

Taken from Kaiser & Murray, (1985).

7.1 SM buffer (per litre)

NaCl	5.8g	100mM
$MgSO_4 \cdot H_2O$	2g	10mM
1M Tris-HCl pH 7.5	50 ml	100mM

Appendix 8: BBL media

Taken from Kaiser & Murray, (1985). BBL media is less rich than LB and is used for phage growth on plates when the plaques need to be kept small.

8.1 BBL broth

Trypticase [BBL Microbiology systems, Becton-Dickson]	10g
NaCl	5g
Water to	1 litre

Adjust pH to 7.2.

8.2 BBL broth agar

BBL broth plus:

Bacto-Agar	10g

8.3 BBL broth top agarose

BBL broth plus:

Agarose (gel quality)	6.5g
$MgSO_4$	10mM

Appendix 9: Ligation for shotgun cloning of λ clones

9.1 The ligation reaction

Subcloning of	Strain 151	Strain 149	
Sterile Analar water			4.5µl
Insert DNA	550 ng	300 ng	1.5µl
pBluescript DNA	400 ng	150 ng	1µl
Ligase buffer (10x)*			1µl
ATP @ 10mM			1µl
T4 Ligase [BRL] @1 u µl^{-1}			1µl

*Constituents given in Appendix 6.4

Incubated at room temperature overnight. In each case an excess of insert over vector DNA was used, this should help to reduce the background of pBluescript containing no insert.

Appendix 10: One step competent cells and transformation

Taken from Chung *et al.*, (1989).

10.1 Protocol

Cells grown to a density represented by an A_{600} of 0.125 were harvested at 2000 rpm, 4°C. The pelleted cells were resuspended in a tenth volume of TSS (see Appendix 10.2) and placed on ice. To 100 µl of these cells in a cold microfuge tube was added 1 or 2 µl of the DNA to be transformed, and these then incubated on ice for 5 to 60 minutes. Following this incubation period 0.9 ml of LB broth (see Appendix A1 of the Handbook of molecular cloning (Sambrook *et al.*, 1989)) was added and the mixture incubated at 37°C for 2 hours. The cells were then plated out onto selective media.

10.2 TSS

LB broth
10% PEG 8000
5% DMSO
50 mM $MgCl_2$
pH 6.5

Appendix 11: Preparation of single stranded DNA

Adapted from USB Corp. instruction manual for their "Pre-sequencing kit for linear, double-stranded DNA". DNA was first prepared at a concentration of 0.3 μg μl^{-1} in sterile analar water.

11.1 Digestion with PvuII

DNA	6μl
Sequenase buffer [USB Corp]	3μl
PvuII [Gibco-BRL] @10 u μl^{-1}	0.5μl
sterile analar water	5.5μl

This mix was then incubated for 1 hour at 37°C. This provides enough template for a single sequencing reaction, and some to analyse the products of digestion. If more template was required this reaction and the one following were simply scaled up.

11.2 Digestion with T7 gene 6 exonuclease

To the completed reaction from Appendix 11.1 was added 15 units of T7 gene 6 exonuclease, the mix was then incubated for 30 minutes at 37°C for digestion, followed by incubation at 80°C for 15 minutes to inactivate the exonuclease. For sequencing 9 μl of this digested DNA (~ 1 μg) was mixed with 1 μl of primer.

Appendix 12: DNA sequencing gels

12.1 TTE buffer (glycerol tolerant gel buffer)

per litre of 20x buffer:

Tris Base	216g
Taurine	72g
$Na_2EDTA{\cdot}2H_2O$	4g

Add water to 1000 ml total and filter. Use at 1x strength.

References

Aagard, J. & Sistrom, W. R. (1972)
Control of Synthesis of Reaction Center Bacteriochlorophyll in Photosynthetic Bacteria.
Photochem. Photobiol. **15** :209-225

Adams, C.W., Forrest, M.E., Cohen, S.N. and Beatty, T. (1989)
Structural and functional analysis of transcriptional control of the *Rhodobacter capsulatus puf* operon.
J. Bacteriol. **171** :473-482

Alberti, M. (1991)
"*R. capsulatus* complete photosynthesis gene cluster"
Genbank Ac. Z11165

Allardet-Servant, A., Michaux-Charachon, S. Jumas-Bilak, E., Karayan, L. and Ramuz, M. (1993)
Presence of one linear and one circular chromosome in the *Agrobacterium tumifaciens* C58 genome
J. Bacteriol. **175** :7869-7874

Allen, J.P., Feher, G., Yeates, T.O., Komiya, H. and Rees, D.C. (1988)a
"Structure of the Reaction Center from *Rhodobacter sphaeroides* R-26 and 2.4.1", In "The photosynthetic bacterial reaction centre", Eds. Breton, J. and Vermeglio, A., NATO ASI Series A: vol 149 Life Sciences, Plenum Press, New York.
p 5-13

Allen, J.P., Feher, G. Yeates, T.O., Komiya, H. and Rees, D.C., (1988)b
Structure of the Reaction Center form *Rb. sphaeroides* R26 and 2.4.1.: Protein cofactor (quinones and Fe^{2+}) interactions.
Proc. Natl. Acad. Sci. USA **85** :8487-8491

Allen, J.P., Feher, G., Yeates, T.O., Komiya, H. & Rees, D.C. (1987)a
Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The cofactors
Proc. Natl. Acad. Sci. USA **84** :5730-5734.e

Allen, J.P., Feher, G., Yeates, T.O., Komiya, H. & Rees, D.C. (1987)b
Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The protein subunits.
Proc. Natl. Acad. Sci . USA **84** :6162-6166

Allen, J.P., Feher, G., Yeates, T.O., Rees, D.C., Deisenhofer, J., *et al.* (1986)
Structural homology of reaction centres from *Rhodopseudomonas sphaeroides* and *Rhodopseudomonas viridis* as determined by x-ray diffraction.
Proc. Natl. Acad. Sci. USA **83** :8589-8593

Armstrong, G.A. (1994)
Eubacteria show their true colors: Genetics of carotenoid pigment biosynthesis from microbes to plants
J. Bacteriol. **176** :4795-4802

Armstrong, G.A., Alberti, M., Leach, F. and Hearst, J.E. (1989)
Nucleotide sequence, organization, and nature of the protein products of the carotenoid biosynthesis gene cluster of *Rhodobacter capsulatus*
Mol. Gen. Genet. **216** :254-268

Armstrong, G.A., Cook, D.N., Ma, D., Alberti, M., Burke, D.H. and Hearst, J.E. (1993)
Regulation of carotenoid and bacteriochlorophyll biosynthesis genes and identification of an evolutionarily conserved gene required for bacteriochlorophyll accumulation
J. Gen. Microbiol. **139** :897-906

Arnoux, B., Ducruix, A., Astier, C., Picaud, M., Roth, M. and Reiss-Husson, F. (1990)
Towards the understanding of the function of *Rb. sphaeroides* Y wild type reaction center: gene cloning, protein and detergent structures in the three-dimensional crystals
Biochemie **72** :525-530

Ashby, M.K., Coomber, S.A. and Hunter, C.N. (1987)
Cloning, Nucleotide Sequence and Transfer of Genes for the B800-850 Light-harvesting Complex of *Rhodobacter sphaeroides*.
FEBS Lett. **213** :245-248

Babitske, P. and Kushner, S.R. (1991)
The arns (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of *Escherichia coli*
Proc. Natl. Acad. Sci. USA **88** :1-5

Barrett, S.J. (1995)
Personal Communication

Bauer, C., Buggy, J. and Mosley, C. 1993
Control of photosystem genes in *Rhodobacter capsulatus*
Trends in Genetics **9** (2) :56-60

Bauer, C.E. and Marrs, B.L. (1988)
Rhodobacter capsulatus puf operon encodes a regulatory protein (pufQ) for bacteriochlorophyll biosynthesis.
Proc. Natl. Acad. SCi. USA **85** :7074-7078

Bauer, C.E., Buggy, J.J., Yang, Z. and Marrs, B.L. (1991)
The superoperonal organization of genes for pigment biosynthesis and reaction center proteins is a conserved feature in *Rhodobacter capsulatus*: analysis of overlapping bchB and puhA transcripts.
Mol. Gen. Genet. **228** :438-444

Bauer, C.E., Young, D.A. and Marrs, B.L. (1988)
Analysis of the *Rhodobacter capsulatus puf* operon. Location of the oxygen-regulated promoter region and the identification of an additional *puf* -encoded gene.
J. Biol. Chem. **263** :4820-4827

Belanger, G. and Gingras, G. (1988)
Structure and expression of the puf operon messenger RNA in *Rhodospirillum rubrum*
J. Biol. Chem. **263** :7639-7645

Bélanger, G., Bérard, J., Corriveau, P., and Gingras, G. (1988)
The structural genes coding for the L and M subunits of the *Rhodospirillum rubrum* photoreaction center
J. Biol. Chem. **263** :7632-7638

Belasco, G., Beatty, J.T., Adams, C.W., van Gabain, A. and Cohen, S.N. (1985)
Differential expression of photosynthetic genes in *R. capsulata* results from segmental differences in stability within the polycistronic rxcA transcript.
Cell **40** :171-181

Belasco, J.G. and Chen, C-Y.A. (1988)
Mechanism of *puf* mRNA Degradation: The Role of an Intercistronic Stem-loop Structure.
Gene **72** :109-117

Bérard, J., Bélanger, G. and Gingras, G. (1989)
Mapping of puh messenger RNAs from *Rhodospirillum rubrum*: Evidence for tandem promoters
J. Biol. Chem. **264** :10897-10903

Bergström, H., Sundström, V., van Grondelle, R., Gillbro, T. and Cogdell, R.J. (1988)
Energy transfer dynamics of isolated B800-B850 and B800-B820 pigment-protein complexes of *Rhodobacter sphaeroides* and *Rhodopseudomonas acidophila*
Biochim. Biophys. Acta. **936** :90-98

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

Bissig, I., Brunisholz, R.A., Suter, F., Cogdell, R.J. and Zuber, H. (1988)
The complete amino acid sequences of the B800-850 antenna polypeptides from Rhodopseudomonas acidophila strain 7750
Z. Naturforsch **43** (c) :77-

Blumberg, D.D. (1987)
"Creating a ribonuclease free environment", in "Guide to Molecular Cloning Techniques", Berger, S.L. and Kimmel (Eds.), A.R., Academic Press Inc.
p20

Bollivar, D.W., Suzuki, J.Y., Beatty, J.T., Dobrowolski, J.M. and Bauer, C.E. (1994)
Directed mutational analysis of bacteriochlorophyll *a* biosynthesis in *Rhodobacter capsulatus*
J. Mol. Biol. **237** (5) :622-640

Bonen, L. and Doolittle, W.F. (1975)
On the Prokaryotic Nature of Red Algal Chloroplasts
Proc. Natl. Acad. Sci. USA **72** :2310-2314

Boucher, F., van der Rest, M. and Gingras, G. (1977)
Structure and function of carotenoid in the photoreaction center from *Rhodospirillum rubrum*
Biochim. Biophys. Acta **461** :339-357

Brendel, V. and Trifanov, E.N. (1984)
A computer algorithm for testing potential prokaryotic terminators
Nuc. Acid. Res. **12** (10) :4411-4427

Britten, R.J. and Davidson, E.H. (1987)
"Hybridisation strategy", in "Nucleic acid hybridisation: a practical approach", Eds. Hames, B.D. and Higgins, S.J., IRL Press.
p3-15

Brunisholz, R.A. and Zuber, H. (1988)
"Primary structure analyses of bacterial antenna polypeptides: Correlation of aromatic amino acids with spectral properties. Structural similarities with reaction center polypeptides", In "Photosynthetic Light Harvesting Systems", Eds. Scheer, H. and Schneider, S., Walterde Gruyter, Berlin
p103

Brunisholz, R.A., Jay, F., Suter, F. and Zuber, H. (1987)
"Structural studies on the light-harvesting polypeptides of *Rp. acidophila*", In "Progress in Photosynthesis", Ed. Biggins, J., Martinus Nijhoff, The Hague.
p13

Brunisholz, R.A., Suter, F. and Zuber, H. (1984)
The light-harvesting polypeptides of *Rhodospirillum rubrum.*
Hoppe-Seyler's Z. Physiol. Chem. **365** :375-

Brunisholz, R.A., Suter, F. and Zuber, H. (1994)
Structural and spectral characterisation of the antenna complexes of *Rhodocyclus gelatinosus*. Indications of a hairpin-like-arranged antenna apoprotein with an unusually high alanine content
Eur. J. Biochem. **222** :667-675

Buck, M., Miller, S., Drummond, M. and Dixon, R. (1986)
Upstream activator sequences are present in the promoters of nitrogen fixation genes
Nature **320** :374-378

Buggy, J.J., Sganga, M.W., Bauer, C.E. (1994)a
Characterization Of A Light-responding Transactivator Responsible For Differentially Controlling Reaction-center And Light-harvesting-1 Gene-expression In *Rhodobacter-capsulatus*
J. Bacteriol **176** (22) :6936-6943

Buggy, J.J., Sganga, M.W., Bauer, C.E. (1994)b
Nucleotide sequence and characterization of the *Rhodobacter capsulatus* hvrB gene: HvrB is an activator of S-adenosyl-L-homocysteine hydrolase expression and is a member of the LysR family.
J. Bacteriol **176** :61-69

Burke, D.H., Alberti, M. and Hearst, J.E. (1993)a
The *Rhodobacter capsulatus* chlorin reductase-encoding locus, bchA, consists of three genes, bchX, bchY, and bchZ
J. Bacteriol. **175** :2407-2413

Burke, D.H., Alberti, M. and Hearst, J.E. (1993)b
bchFNBH Bacteriochlorophyll synthesis genes of *Rhodobacter capsulatus* and identification of the third subunit of light-independent protochlorophyllide reductase in bacteria and plants
J. Bacteriol. **175** :2414-2422

Carpousis, A.J., Van Houwe, G., Ehretsmann, C. and Krisch, H.M. (1994)
Copurification of *E. coli* RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation
Cell **76** :889-900

Chang, C-H., El-Kabbani, O., Tiede, D., Norris, J. and Schiffer, M. (1991)
Structure of the membrane bound protein photosynthetic reaction centre from Rhodobacter sphaeroides.
Biochemistry **30** :5352-5360

Chang, C-H., Tiede, D. M., Tang, J., Smith, U., Norris, J. R. and Schiffer, M. (1986)
Structure of the *Rhodopseudomonas sphaeroides* R26 reaction center
FEBS Lett. **205** :82-86

Chen, C.A., Beatty, J.T., Cohen, S.N. and Belasco, J.G. (1988)
An Intercistronic Stem-Loop Structure Functions as an mRNA Decay Terminator Necessary but Insufficient for *puf* mRNA Stability.
Cell **52** :609-619

Chen, E.J. and Seeburg, P.H. (1985)
Supercoil sequencing: a fast and simple method for sequencing plasmid DNA
DNA **4** :165-170

Cheng, H.-P. and Lessie, T.G. (1994)
Multiple replicons constituting the genome of *Pseudomonas capacia* 17616
J. Bacteriol. **176** :4034-4042

Chomczynski, P. and Sacchi, N. (1987)
Single-Step Method of RNA Isolation by Acid Guanidium Thiocyanate-Phenol-Chloroform Extraction.
Anal. Biochem. **162** :166-159

Chory, J., & Kaplan, S. (1983)
Light dependant regulation of the synthesis of soluble and intracytoplasmic membrane proteins of *Rhodopseudomonas sphaeroides*
J. Bacteriol. **153** (1) :465-474

Choudhary, M., Mackenzie, C., Nereng, K.S., Sodergren, E., weinstock, G.M. and Kaplan, S. (1994)
Multiple Chromosomes In Bacteria - Structure And Function Of Chromosome-ii Of *Rhodobacter-sphaeroides*-2.4.1(t)
J. Bacteriol. **176** (24) :7694-7702

Chung, C.T., Niemela, S.L. and Miller, R.H. (1989)
One-step preparation of competant *Escherichia coli*: transformation and storage of bacterial cells in the same solution
Proc. Natl. Acad. Sci. USA **86** :2172-2175

Clark, W.G., Davidson, E. and Marrs, B.L. (1984)
Variation of levels of mRNA coding for antenna and reaction center polypeptides in *Rhodopseudomonas capsulata* in response to changes in oxygen concentration
J. Bacteriol. **157** :945-948

Clayton, B. J. & Clayton, R. K. (1978)
Properties of photochemical reaction centers purified from *Rhodopseudomonas gelatinosa*
Biochim. Biophys. Acta **501** :470-477

Cogdell, R.J. (1985)
"Carotenoid-Bacteriochlorophyll Interactions", In "Antenna's and Reaction Centres of Photosynthetic Bacteria.", Ed. Michel-Beyerle, M. E., Chemical Physics Series, volume 43
p53

Cogdell, R.J. (1986)
"Light-Harvesting Complexes in the purple photosynthetic bacteria", section 6.2, In "Photosynthesis III", Eds. Stachelin, C. A., Arntzen, C.
p252

Cogdell, R.J., and Crofts, A.R. (1978)
Analysis of the pigment content of an antenna pigment-protein complex from three strains of *Rhodopseudomonas sphaeroides*
Biochim. Biophys. Acta **502** :409-416

Cogdell, R.J. and Thornber, J.P. (1980)
Light-harvesting pigment-protein complexes of purple photosynthetic bacteria.
FEBS Lett. **122** (1) :1-8

Cogdell, R.J., and Thornber, J.P. (1979)
"The preparation and characterisation of different types of light-harvesting pigment-protein complexes from some purple bacteria", In "Chlorophyll Organisation and Energy Transfer in Photosynthesis", Ciba Foundation Symposium 61 (new series). Elsevier, North Holland
p61-79

Cogdell, R.J., Durant, I., Valentine, J., Lindsay, J.G. and Schmidt, K. (1983)
The Isolation and Partial Characterisation of the Light-Harvesting Pigment-protein Complement from *Rhodopseudomonas acidophila* strain 10050
Biochim. Biophys. Acta **722** :427-435

Cogdell, R.J., Hipkins, M.F., MacDonald, W. and Truscott, G.T. (1981)
Energy transfer between the carotenoid and the bacteriochlorophyll within the B800-850 light-harvesting pigment-protein complex of *Rhodopseudomonas sphaeroides*
Biochim. Biophys. Acta **634** :191-202

Cogdell, R.J., Lindsay, J.G, Reid, G.P. and Webster, G.D. (1980)
A Comparison of the Constituent Polypeptides of the B800-850 Light-harvesting Pigment-protein Complex from *Rhodpseudomonas sphaeroides*.
Biochem. Biophys. Acta **591** :312-320

Cohen-Bazire, G. & Kunisawa, R. (1960)
Some observations on the synthesis and function of the photosynthetic apparatus in *Rhodospirillum rubrum*.
Proc. Natl. Acad. Sci. USA **46** :1543-1553

Cohen-Bazire, G. N., Pfennig, N. & Kunisawa, R. (1964)
The Fine Structure of Green Bacteria
J. Cell. Biol. **22** :207-255

Collado-Vides, J., Magasanik, B. and Gralla, J.D. (1991)
Control site location and transcriptional regulation in *Escherichia coli*
Microbiol. Rev. **55** :371-394

Colonna-Romano, S., Arnold, W., Schluter, A., Boistart, P., Puhler, A. and Priefer, U.B. (1990)
An Fnr-like protein encoded in *Rhizobium leguminosarum* biovar *viciae* shows structural and functional homology to Rhizobium meliloti FixK
Mol. Gen. Genet. **223** :138-147

Cook, D.N., Armstrong, G.A. and Hearst, J.E. (1989)
Induction of anaerobic gene expression in *Rhodobacter capsulatus* is not accompanied by a local change in chromosomal supercoiling as measured by a novel assay
J. Bacteriol. **171** :4836-4843

Coomber, S.A., Chaudhri, M., Connor, A., Britton, G. and Hunter, C.N. (1990)
Localised transposon Tn5 mutagenesis of the photosynthetic gene cluster of *Rhodobacter sphaeroides*
Mol. Microbiol. **4** :977-989

Coomber, S.A., Jones, R.M., Jordan, P.M. and Hunter, C.N. (1992)
A Putative Anaerobic Coproporphyrinogen III Oxidase in *Rhodobacter sphaeroides*. I. Molecular Cloning, Transposon Mutagenesis and Sequence Analysis of the Gene.
Mol. Microbiol. **6** :3159-3169

Creighton, S., Hwang, J.-K., Warshel, A., Parson, W.W. & Norris, J.R. (1988)
Simulating the dynamics of the primary charge separation process in bacterial photosynthesis
Biochemistry **27** :774-781

Crofts, A. R. and Wraight, C. A. (1983)
The Electrochemical Domain of Photosynthesis
Biochim. Biophys. Acta **726** :149-185

Cullen, P.J., FosterHartnett, D., Gabbert, K.K. and Kranz, R.G. (1994)
Structure and expression of the alternative sigma factor, RpoN, in *Rhodobacter capsulatus*; physiological relevance of an autoactivated nifU2-rpoN superoperon
Mol. Microbiol. **11** :51-65

Debus, R. J., Feher, G. and Okamura, M. Y. (1985)
LM Complex of Reaction Centers from *Rhodopseudomonas sphaeroides* R26: Characterisation and Reconstitiution with the H Subunit
Biochemistry **24** :2488-2500

Debus, R. J., Feher, G. and Okamura, M. Y. (1986)
Iron depleted reaction centres from *Rhodopseudomonas sphaeroides* R26.1: characterization and reconstitution with Fe^{2+}, Mn^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}.
Biochemistry **25** :2276-2287

Dehoff, B.S., Lee, J.K., Donohue, T.J., Gumport, R.I. and Kaplan, S. (1988)
In vivo anlaysis of puf operon expression in *Rhodobacter sphaeroides* after deletion of a putative intercistronic transcription terminator
J. Bacteriol. **170** :4681-4692

Deisenhofer, J. and Michel, H. (1989)
The photosynthetic reaction center from the purple bacterium *Rhodopseudomonas viridis*.
Science **245** :1463-1473

Deisenhofer, J. and Michel, H. (1991)a
High-Resolution Structures of Photosynthetic Reaction Centers
Annu. Rev. Biophys. Chem. **20** :247-266

Deisenhofer, J. and Michel, H. (1991)b
Structures of Bacterial Photosynthetic Reaction Centers
Annu. Rev. Cell Biol. **7** :1-23

Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. (1985)
Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3Å resolution.
Nature **318** :618-624

Deisenhofer, J., Epp, O., Sinning, I., Michel, H. (1995)
Crystallographic Refinement at 2.3-Å Resolution and Refined Model of the Photosynthetic Reaction-Center from *Rhodopseudomonas viridis*
J. Mol. Biol. **246** (3) :429-457

Dexter, D.L. (1953)
J. Chem. Phys. **21** :836-850

Dickerson, R. E. (1980)a
"Structural conservatism in proteins over three billion years: cytochrome with a touch of collagen." In "Biomolecular structure, conformation, function and evolution", vol 1, Srinivasan (Ed.), Pergamon Press, Oxford.
p227-249

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

Dracheva, S.M., Drachev, L.A., Konstantinov, A.A., Semenov, A.Y., Skulachev, V.P., *et al.* (1988)
Electrogenic steps in the redox reactions catalyzed by photosynthetic reaction centre complex from *Rhodopseudomonas viridis*
Eur. J. Biochem. **171** :253-264

Dressler, K., Umlauf, E., Schmidt, S., Hamm, P., Zinth, W. and Buchanan, S. (1991)
Detailed studies of the subpicosecond kinetics in the primary electron transfer of reaction centres of *Rhodopseudomonas viridis*
Chem. Phys. Letts. **183** :270-276

Drews, G. (1985)
Structure and functional organisation of light-harvesting complexes and photochemical reaction centres in membranes of phototrophic bacteria.
Microbiol. Rev. **49** (1) :59-70

Drews, G. (1991)
"Regulated development of the photosynthetic apparatus in anoxygenic bacteria", In "The Photosynthetic Apparatus", Eds. Bogorad, L. and Vasil, I.K., Academic Press, San Diego, CA.
p113-148

Drews, G. & Oelze, J. (1981)
Organisation and Differentiation of Membranes of Phototrophic Bacteria.
Adv. Microb. Physiol. **22** :1-92

Ehretsmann, C.P., Carpousis, A.J. and Krisch, H.M. (1992)
Specificity of *Escherichia coli* endoribonuclease RNase E: in vivo and in vitro analysis of mutants in a bacteriophage T4 mRNA processing site
Genes Dev. **6** :149-159

Eraso, J.M., and Kaplan, S. (1995)
Oxygen-insensitive synthesis of the photosynthetic membranes of *Rhodobacter sphaeroides*: A mutant histidine kinase
J. Bacteriol. **177** (10) :2695-2706

Eraso, J.M., and Kaplan, S. (1994)
prrA, a putative response regulator involved in oxygen regulation of photosynthesis gene expression in *Rhodobacter sphaeroides*
J. Bacteriol. **176** :32-43

Ermler, U., Michel, H. and Schiffer, M. (1994)
Structure and Function of the Photosynthetic Reaction Center from *Rhodobacter sphaeroides*
J. Bioen. Biomem. **26** (1) :5-15

Evans, M.B, Hawthornthwaite, A.M. and Cogdell, R.J. (1990)
Isolation and characterisation of the different B800-850 light-harvesting complexes from low- and high-light grown cells of *Rhodopseudomonas palustris* strain 2.1.6
Biochim. Biophys. Acta. **1016** (1) :71-76

Farchaus, J.W. and Oesterhelt, D. (1989)
A *Rhodobacter sphaeroides* puf L, M and X deletion mutant and complementation in *trans* with a 5.3kb puf operon shuttle fragment
EMBO J. **8** :47-54

Farchaus, J.W., Barz, W.P., GrÜnberg, H. and Oesterhelt, D. (1992)
Studies on the expression of the pufX polypeptide and its requirement for photoheterotrophic growth in *Rhodobacter sphaeroides*
EMBO J. **11** :2779-2788

Feher, G., Allen, J. P., Okamura, M. Y. and Rees, D. C. (1989)
Structure and function of bacterial photosynthetic reaction centres
Nature **339** :111-116

Feher, G. and Okamura, M. Y. (1978)
"Chemical composition and properties of reaction centers.", In "The Photosynthetic Bacteria", Eds. Clayton, R. K. and Sistrom, W. R., Plenum Press, New York.
p349-386

Feher, G., Hoff, A.J., Isaacson, R.A. and Ackerson, L.C. (1975)
Endor Experiments on Chlorophyll and Bacteriochlorophyll *in vitro* and in the photosynthetic unit
Ann. N.Y. Acad. Sci. **244** :239-259

Feick, R.G., Fitzpatrick, M. and Fuller, R.C. (1982)
Isolation and characterization of cytoplasmic membranes and chlorosomes from the green bacterium *Chloroflexus aurantiacus*
J. Bacteriol. **159** :905-915

Feinberg, A.P. and Vogelstein, B. (1983)
A technique for labelling DNA restriction endonuclease fragments to high specific activity
Analytical Biochem. **132** :6-13

Fickett, J.W. (1982)
Recognition of protein coding regions in DNA sequences
Nucl. Acids Res. **10** :5303-5318

Fidai, S., Hinchigeri, S.B. and Richards, W.R. (1994)
Association of protochlorophyllide with the PufQ protein of *Rhodobacter capsulatus*
Biochem. Biophys. Res. Comm. **200** (3) :1679-1684

Firsow, N.N. and Drews, G. (1977)
Differentiation of the intracytoplasmic membrane of *Rhodopseudomonas palustris* by variation in oxygen partial pressure or light intensity
Arch. Microbiol. **115** :299-306

Flemming, G.R., Martin, J.L. and Breton, J. (1988)
Rates of primary electron transfer in photosynthetic reaction centres and their mechanistic implications
Nature **333** :190-192

Fonstein, M., Koshy, E.G., Nikolskaya, T., Mourachov, P. and Haselkorn, R. (1995)
Refinement Of The High-resolution Physical And Genetic-map Of *Rhodobacter-capsulatus* And Genome Surveys Using Blots Of The Cosmid Encyclopedia
EMBO J. **14** (8) :1827-1841

Fonstein, M., Zheng, S. and Haselkorn, R. (1992)
Physical map of the genome of *Rhodobacter capsulatus* SB1003
J. Bacteriol. **174** (12) :4070-4077

Ford, R.C. 1992
Photosynthetic Membrane Proteins
Current Opinion in Structural Biology **2** :527-533

FosterHartnett, D. and Kranz, R.G. (1994)b
The *Rhodobacter capsulatus* glnB gene is regulated by NtrC at tandem rpoN- independent promoters
J. Bacteriol. **176** :5171-5176

FosterHartnett, D., Cullen, P.J., Gabbert, K.K. and Kranz, R.G. (1993)
Sequence, genetic, and lacZ fusion analyses of a nifR3-ntrB-ntrC operon in *Rhodobacter capsulatus*
Mol. Microbiol. **8** :903-914

FosterHartnett, D., Cullen, P.J., Monika, E.M., Kranz, R.G. (1994)a
A new type of NtrC transcriptional activator
J. Bacteriol. **176** :6175-6187

Fowler, G.J.S, Gardiner, A.T., Mackenzie, R.C., Barrett, S.J., Simmons, A.E., Westerhuis, (1995) W.H.J., Cogdell, R.J. and Hunter, C.N.
Heterologous expression of genes encoding bacterial light-harvesting complexes in *Rhodobacter sphaeroides*
J. Biol. Chem. **270** (40) :1-8

Fowler, G.J.S., Sockalingum, G.D., Robert, B. and Hunter, C.N. 1994
Blue shifts in bacteriochlorophyll absorbance correlate with changed hydrogen bonding patterns in light-harvesting 2 mutants of *Rhodobacter sphaeroides* with alterations at alpha -Tyr-44 and alpha -Tyr-45
Biochem. J. **299** (3) :695-700

Fowler, G.J.S., Vischers, R.W., Grief, G.G., van Grondelle, R. and Hunter, C.N. 1994
Genetically modified photosynthetic antenna complexes with blue-shifted absorbtion band
Nature **335** :848-850

Frischauf, A.M., Lehrach, H., Poustka, A., and Murray, N. (1983)
Lambda replacement vectors carrying polylinker sequences
J. Mol. Biol. **170** :827-842

Fritsch, J., Rothfuchs, R., Rauhut, R. and Klug, G. (1995)
Identification of an mRNA element promoting rate-limiting cleavage of the polycistronic puf mRNA in *Rhodobacter capsulatus* by an enzyme similar to RNase E
Mol. Microbiol. **15** (6) :1017-1029

Fukushima, A., Matsuura, K., Shimada, K. and Satoh,T. (1988)
Reaction center-B870 pigment protein complexes with bound cytochromes c-555 and c-551 from *Rhodocyclus gelationosus.*
Biochim. Biophys. Acta. **933** :399-405

Gall, A. (1994)
Purification, Characterisation, and Crystallisation from a Range of Rhodospirillineae Pigment-protein Complexes
Ph.D. Thesis, Glasgow University p10-12

Gardiner, A.T. (1992)
"Peripheral Antenna Complexes from *Rhodopseudomonas acidophila* : Structure, Function and Genetic Manipulation", Ph.D. Thesis, University of Glasgow.

Gardiner, A.T., Cogdell, R.J. and Shinichi, T. (1993)
The effect of growth conditions on the light-harvesting apparatus in *Rhodopseudomonas acidophila*
Photosynth. Res. **38** :159-167

Gardiner, A.T., MacKenzie, R.C., Barrett, S.J., Kaiser, K. and Cogdell, R.J. (1992)
"The genes for the peripheral antenna complex apoproteins from *Rhodopseudomonas acidophila* 7050 from a multigene family", In "Research in Photosynthesis". vol. 1, Ed. Murata, N., Kluwer Academic Publishers, Netherlands.
p77-80

Garí, E., Toledo, J.C., Gibert, I. and Barbé, J. (1992)
Nulceotide sequencce of the methoxyneurosporene dehydrogenase gene from *Rhodobacter sphaeroides*: comparison with other bacterial carotenoid dehydrogenases
FEMS. Microbiol. Letts. **93** :103-108

Gest, H. (1993)
Photosynthetic And Quasi-photosynthetic Bacteria
FEMS Microbiol. Letts. **112** (1) :1-6

Gibson, L.C.D., McGlynn, P., Chaudhri, M. and Hunter, C.N. (1992)
A Putative Anaerobic Coproporphyrinogen III Oxidase in *Rhodobacter sphaeroides*. II. Analysis of a region of the genome encoding hemF and the puc operon.
Molec. Microbiol. **6** :3171-3186

Gicquel-Sanzey, B. and Cossart, P. (1982)
Homologies between different procaryotic DNA-binding proteins and between their sites of action
EMBO J. **1** (5) :591-595

Gilbert, W. (1976)
"Starting and stopping sequences for the RNA polymerase", in "RNA Polymerase", Eds. Losick, R. and Chamberlin, M., Cold Spring Harbor Laboratory
p193-205

Gillbro, T. and Cogdell, R.J. (1989)
Chem. Phys. Lett. **158** :312-316

Gillbro, T., Cogdell, R.J. and Sundström, V. (1988)
Energy transfer from carotenoid to bacteriochlorophyll a in the B800-820 antenna complexes from *Rhodopseudomonas acidophila* strain 7050
FEBS Lett. **235** :169-172

Giuliano, G., Pollack, D., Stapp, H. and Scolnik, P.A. (1988)
A genetic-physical map of the *Rhodobacter capsulatus* carotenoid biosynthesis gene cluster
Mol. Gen. Genet. **213** :78-83

Goedheer, J.C. (1959)
Energy Transfer between carotenoid and bacteriochlorophyll in chromatophores of purple bacteria
Biochim. Biophys. Acta **35** :1-8

Golecki, J.R., Ventura, S. and Oelze, J. (1991)
the architecture of unusual membrane tubes in the B800-850 light-harvesting bacteriochlorophyll deficient mutant 19 of *Rhodobacter sphaeroides*
FEMS Microbiol. Lett. **77** :335-340

Gomelsky, M. and Kaplan, S. (1995)
Genetic evidence that PpsR from *Rhodobacter sphaeroides* 2.4.1 functions as a transcriptional repressor of *puc* and *bchF* expression
J. Bacteriol. **177** (6) :1634-1637

Gong, L., Lee, J.K. and Kaplan, S. (1994)
The Q gene of *Rhodobacter sphaeroides*: Its role in puf operon expression and spectral complex assembly
J. Bacteriol. **176** (10) :2946-2961

Gray, K.A., Farchaus, J.W., Wachtveitl, J., Breton, J. and Oesterhelt, D. (1990)
Initial Characterisation of Site-Directed Mutants of Tyrosine M210 in the Reaction Centre of *Rhodobacter sphaeroides*
EMBO J. **9** :2061-2070

Hagemann, G.E. and Fadros, M. (1994)
"Isolation, characterisation and gene organisation of the unusual variable antenna complex (LHII) of *Rhodovulum sulfidophilus*". In "Abstracts of the VIII international symposium on phototrophic prokaryotes", Eds. Tedioli, G., Ventura, S. and Zannoni, D.
p69

Hamm, P., Zurek, M., Mantele, W., Meyer, M., Scheer, H. and Zinth, W. (1995)
Femtosecond infrared spectroscopy of reaction centers from *Rhodobactersphaeroides* between 1000 and 1800 cm-1
Proc. Natl. Acad. Sci. USA **92** (6) :1826-1830

Harashima, K. and Nakada, H. (1983)
Carotenoids And Ubiquinone In Aerobically Grown Cells Of An Aerobic Photosynthetic Bacterium, *Erythrobacter* Species Och-114
Agric. Biol. Chem. **47** (5) :1057-1063

Hattori, M. and Sakaki, Y. (1986)
Dideoxy sequencing method using denatured plasmid templates
Anal. Biochem. **152** :232-238

Hawthornthwaite, A.M. and Cogdell, R.J. (1991)
"Bacteriochlorophyll binding proteins", In "Chlorophylls", Ed. Scheer, H., CRC Press, Boca Raton.
p493-528

Hayashi, H., Nakano, M., and Morita, S. (1982)
Comparative studies of Protein Properties and Bacteriochlorophyll Contents of Bacteriochlorophyll-Protein Complexes from Spectrally Different Types of *Rhodopseudomonas palustris*.
J. Biochem. **92** :1805-1811

Hayashi, H., Noguchi, T. and Tasumi, M. (1989)
Studies on the interrelationship among the intensity of a raman marker band of carotenoids, polyene chain structure, and efficiency of transfer from carotenoids to bacteriochlorophyll in photosynthetic bacteria
Photochem. Photobiol. **49** :337-343

Heinmeyer E-A., and Schmidt, K. (1983)
Changes in carotenoid biosynthesis caused by variation of growth conditions in cultures of *Rhodopseudomonas acidophila* strain 7050
Arch. Microbiol. **134** :217-221

Henikoff, S., Haughn, G.W., Calvo, J.M. and Wallace, J.C. (1988)
A large family of bacterial activator proteins
Proc. Natl. Acad. Sci. USA **85** :6602-6606

Hess, S., Visscher, K.J., Pullerits, T., Sundstrom, V., Fowler, G.J.S. and Hunter, C.N. 1994
Enhanced rates of subpicosecond energy transfer in blue-shifted light harvesting LH2 mutants of *Rhodobacter sphaeroides*
Biochemistry (USA) **33** (27) :8300-8305

Hiraishi, A., and Ueda, Y. (1994)
Intrageneric Structure Of The Genus *Rhodobacter* - Transfer Of *Rhodobacter-sulfidophilus* And Related Marine Species To The Genus *Rhodovulum* Gen-nov
Int. J. Syst. Bacteriol. **44** (1) :15-23

Hiraishi, A., Umta, K., Satoh, T. (1995)
A New Genus Of Marine Budding Phototrophic Bacteria, *Rhodobium* Gen-nov, Which Includes *Rhodobium-orienstis* Sp-nov And *Rhodobium-marinum* Comb-nov
Int. J. Syst. Bacteriol. **45** (2) :226-234

Hofmann, K. and Stoffel, W. (1993)
TMBASE - A database of membrane spanning protein segments
Biol. Chem. Hoppe-Seyler **374** :166

Hubner, P., Masepohl, B., Klipp, W. and Bickle, T.A. (1993)
nif gene expression studies in *Rhodobacter capsulatus*: ntrC-independent repression by high ammonium concentrations
Mol. Microbiol. **10** (1) :123-132

Hunter, C.N., Ashby, M.K. and Coomber, S.A. (1987)
Effect of oxygen on levels of mRNA coding for reaction-centre and light-harvesting polypeptides of *Rhodobacter sphaeroides*
Biochem. J. **247** :489-492

Hunter, C.N., Hundle, B.S., Hearst, J.E., Lang, H.P., Gardiner, A.T., Takaichi, S. and Cogdell, (1994)
R.J.
Introduction of new carotenoids into the bacterial photosynthetic apparatus by combining the carotenoid biosynthetic pathways of *Erwinia herbicola* and *Rhodobacter sphaeroides*
J. Bacteriol. **176** :3692-3697

Hunter, C.N., McGlynn, P., Ashby, M.K., Burgess, J.G. and Olsen, J.D. (1991)
DNA sequencing and complementation/deletion analysis of the bchA-puf operon region of *Rhodobacter sphaeroides*: in vivo mapping of the oxygen-regulated puf promoter
Mol. Microbiol. **5** :2469-2661

Hunter, C.N., Pennoyer, J.D., Sturgis, J.N., Farrely, D. and Niederman, R.A. (1988)
Oligomerisation states and associations of light-harvesting pigment-protein complexes of *Rhodobacter sphaeroides* as analysed by lithium dodecyl sulphate-polyacrylamide gel electrophoresis
Biochemistry **27** :3459-3467

Iba, K. and Takamiya, K. (1989)
Action spectra for inhibition by light of accumulation of bacteriochlorophyll and carotenoid during aerobic growth of photosynthetic bacteria
Plant Cell Physiol. **30** :471-477

Imhoff, J. F. & Trüper, H. G. (1989)
"Purple nonsulfur bacteria (Rhodospirillaceae Pfennig and Trüper 1971, 17^{AL})". In "Bergey's Manual of Systematic Bacteriology" vol. 3, Ed. Staley, J.T., Bryant, M. P., Pfennig, N. and Holt, J. G., The Williams & Wilkins Co., Baltimore.
p1658-1682

Imhoff, J. F., Trüper, H. G. & Pfennig, N. (1984)
Rearrangement of the species and genera of the phototrophic "purple nonsulfur bacteria"
Int. J. Syst. Bacteriol. **34** :340-343

Inoue, K., Kouadio, J.L.K., Mosley, C.S. and Bauer, C.E. (1995)
Isolation And In-vitro Phosphorylation Of Sensory Transduction Components Controlling Anaerobic Induction Of Light-harvesting And Reaction-center Gene-expression In *Rhodobacter-capsulatus*
Biochemistry **34** (2) :391-396

Jirsakova, V. and Reiss-Husson, F. (1993)
Isolation And Characterization Of The Core Light-harvesting Complex B875 And Its Subunit Form, B820, From *Rhodocyclus gelatinosus*
Biochim. Biophys. Acta **1183** (2) :301-308

Jirsakova, V., Reiss-Husson, F. (1994)
A Specific Carotenoid Is Required For Reconstitution Of The *Rubrivivax-gelatinosus*- B875 Light-harvesting Complex From Its Subunit Form B820
FEBS Lett. **353** (2) :151-154

Jones, M.R., Fowler, G.J.S., Gibson, L.C.D., Grief, G.G., Olsen, J.D., Crielaard, W. & (1992)
Hunter, C.N.
Mutants of Rhodobacter sphaeroides lacking one or more pigment-protein complexes and complementation with RC, LH1 and LH2 genes.
Molecular Microbiol. **6** (9) :1173-1184

Jones, R. and Haselkorn, R. (1989)
The DNA sequence of the *Rhodobacter capsulatus ntrA*, *ntrB*, and *ntrC* gene analogues required for nitrogen fixation.
Mol. Gen. Genet. **215** :507-516

Kaiser, K. and Murray, N.E. (1985)
"The use of phage Lambda replacement vectors in the construction of representative genomic DNA libraries", In "DNA Cloning Volume 1: a pratical approach.", (Ed.) Glover, D.M., IRL Press.
p1-47

Kaplan, S. & Arntzen, C. J. (1982)
"Photosynthetic Membrane Structure and Function" In "Photosynthesis: Energy Conversion by Plants and Bacteria", volume 1, Ed. Govindjee, Academic Press (London).
p65-151

Karrasch, S., Bullough, P.A. and Ghosh, R. (1995)
The 8.5 Å projection map of the light-harvesting complex I from Rhodospirillum rubrum reveals a ring composed of 16 subunits.
EMBO J. **14** (4) :631-638

Kiley, P.J. and Kaplan, S. (1987)
Cloning, DNA sequence, and expression of *Rhodobacter sphaeroides* light harvesting B800-850a and B800-850b genes
Microbio. Rev. **169** :3268-3275

Kiley, P.J., Donohue, T.J., Havelka, W.A. and Kaplan, S. (1987)
DNA sequence and *in vitro* expression of the B875 light-harvesting polypeptides of *Rhodobacter sphaeroides*
J. Bacteriol. **169** :742-750

Kiley, P.J., Varga, A. and Kaplan, S. (1988)
Physiological and structural analysis of light-harvesting mutants of *Rhodobacter sphaeroides*
J. Bacteriol. **170** :1103-1115

Kirmaier, C. and Holten, D. (1987)
Primary photochemistry of reaction centres from the photosynthetic purple bacteria.
Photosynthesis Res. **13** :225-260

Klug, G. (1993)
Regulation Of Expression Of Photosynthesis Genes In Anoxygenic Photosynthetic Bacteria
Arch. Microbiol. **159** :397-404

Klug, G. (1991)a
Endonucleolytic degradation of puf mRNA in *Rhodobacter capsulatus* is influenced by oxygen
Proc. Natl. Acad. Sci. USA **88** :1765-1769

Klug, G. (1991)b
A DNA sequence upstream of the *puf* operon of *Rhodobacter capsulatus* is involved in its oxygen-dependent regulation and functions as a protein binding site.
Mol. Gen. Genet. **226** :167-176

Klug, G., Adams, C.W., Belasco, J., Doerge, B. and Cohen, S.N. (1987)
Biological Consequences of Segmental Alterations in mRNA Stability : Effects of Deletion of the Intercistronic Hairpin Loop Region of the *Rhodobacter capsulatus* puf Operon.
EMBO J. **6** :3515-3520

Klug, G. and Cohen, S.N. (1990)
Combined actions of multiple hairpin loop structures and sites of rate limiting endonucleolytic cleavage determine the differential degradation rates of individual segments within polycistronic puf operon mRNA
J. Bacteriol. **172** :5140-5146

Klug, G. and Drews, G. (1984)
Construction of a gene bank of *R. capsulata* using a broad host range cloning sysem.
Arch. Microbiol. **139** :319-325

Klug, G,. Kaufmann, N., Drews, G. (1985)
Gene expression of pigment binding proteins of the bacterial photosynthetic apparatus: Transcription and assembly in the membrane of *Rhodopseudomonas capsulata*
Proc. Natl. Acad. Sci. USA. **82** 6485-6489

Knox, R.S. (1977)
"Photosynthetic efficiency and exciton transfer and trapping", In "Topics in Photosynthesis", vol. 2, Ed. Barber, J., Elsevier, Amsterdam
p55

Kramer, H.J.M., Pennoyer, J.D., van Grondelle, R., Westerhuis, W.H.J., Niederman, R.A. and (1984)
Amesz, J.
Lowptemperature optical properties and pigment organization of th B875 light-harvesting Bacteriochlorophyll-protein complex of purple photosynhtetic bacteria
Biochim. Biophys. Acta **767** :335-344

Kyte, J. and Doolittle, R.F. (1982)
A simple method for diplaying the hydrophobic character of a protein
J. Mol. Biol. **157** :105-132

Lang, H.P. and Hunter, C.N. (1994)
The relationship between carotenoid biosynthesis and the assembly of the light-harvesting LH2 complex in *Rhodobacter sphaeroides*
Biochem. J. **298** :197-205

Lang, H.P., Cogdell, R.J., Gardiner, A.T. and Hunter, C.N. (1994)
Early Steps In Carotenoid Biosynthesis - Sequences And Transcriptional Analysis Of The crtI And crtB Genes Of *Rhodobacter-sphaeroides* And Overexpression And Reactivation Of crtI In *Escherichia-coli* And *Rhodobacter-sphaeroides*
J. Bacteriol. **176** :3859-3869

Lang, H.P., Cogdell, R.J., Takaichi, S. and Hunter, C.N. (1995)
Complete DNA sequence, specific Tn5 insertion map, and gene assignment of the carotenoid biosynthesis pathway of *Rhodobacter sphaeroides*
J. Bacteriol. **177** (8) :2064-2073

Laussermair, E. and Oesterhelt, D. (1992)
A system for site-specific mutagenesis of the photsynthetic reaction centre in *Rhodopseudomonas viridis*
EMBO J **11** :777-783

LeBlanc, H. and Beatty, J.T. (1994)
"Topological arrangement of pucC, a membranous protein required for the presence of LHII complex in *Rhodobacter capsulatus*", In "Abstracts of the VIII international symposium on phototrophic prokaryotes", Eds. Tedioli, G., Ventura, S. and Zannoni, D.
p70

LeBlanc, H.N. and Beatty, J.T. (1993)
Rhodobacter capsulatus puc operon: Promoter location, transcript sizes and effects of deletions on photosynthetic growth
J. Gen. Microbiol. **139** :101-109

Lee, J.K. and Kaplan, S. (1992)
Cis-acting Regulatory Elements Involved in Oxygen and Light Control of *puc* Operon Transcription in *Rhodobacter sphaeroides*.
J. Bacteriol. **174** :1146-1157

Lee, J.K., DeHoff, B.S., Donohue, T.J., Gumport, R.I. and Kaplan, S. (1989)a
Transcriptional analysis of puf operon expression in *Rhodobacter sphaeroides* 2.4.1 and an intercistronic transcription terminator mutant
J. Biol. Chem. **264** :19354-19365

Lee, J.K., Kiley, P.J., and Kaplan, S. (1989)b
Posttranscriptional control of *puc* operon expression of B800-850 light-harvesting complex of *Rhodobacter sphaeroides*.
J. Bacteriol. **171** :3391-3405

Lee, J.K., Wang, S., Eraso, J.M., Gardner, J. and Kaplan, S. (1993)
Transcriptional regulation of puc operon expression in *Rhodobacter sphaeroides*. Involvement of an integration host factor-binding sequence
J. Biol. Chem. **268** (32) :24491-24497

Leibl, W., Sinning, I., Ewald, G., Michel, H. and Breton, J. (1993)
Evidence that serine L223 is involved in the proton transfer pathway to Q_B in the photosynthetic reaction center of *Rhodopseudomonas viridis*
Biochemistry **32** :1958-1964

Liebetanz, R., Hornberger, U. and Drews, G. (1991)
Organization of the genes coding for the reaction-centre L and M subunits and B870 antenna polypeptides alpha and beta from the aerobic photosynthetic bacterium *Erythrobacter* species OCH114
Mol. Microbiol. **5** (6) :1459-1468

Lilburn, T.G., Haith, C.E., Prince, R.C. and Beatty, J.T. (1992)
Pleiotropic effects of pufX gene deletion on the structure and function of the photosynthetic apparatus of *Rhodobacter capsulatus*
Biochim. Biophys. Acta **1100** :160-170

Lundberg, U., von Gabain, A. and Melefors, Ö. (1990)
Cleavages in the 5′ region of the ompA and bla mRNA control stability: studies with an *E. coli* mutant altering mRNA stability and a novel endoribonuclease
EMBO J. **9** :2731-2741

Ma, D., Cook, D.N., O'Brien, D.A. and Hearst, J.E. (1993)
Analysis of the promoter and regulatory sequences of an oxygen-regulated bch operon in *Rhodobacter capsulatus* by site-directed mutagenesis
J. Bacteriol. **175** (7) :2037-2045

MacKenzie, R.C. (1990)
"Cloning and sequencing the genes encoding the light harvesting (LHII) polypeptides of *Rhodopseudomonas acidophila* strain 7050", Ph.D. Thesis, University of Glasgow.

Mackie, G.A. (1992)
Secondary structure of the mRNA for ribosomal protein S20
J. Biol. Chem. **267** :1054-1061

Mantele, W., Steck, A., Becker, T., Wacker, T., Welte, W., Gad'on, N. and Drews, G. (1988)
"Pigment-Protein Complexes of *Rhodopseudomonas palustris*", In "the Photosynthetic Bacterial Reaction Center"., Eds. by Breton, J. and Vermeglio, A., Plenum Publishing Coporation.
p33-39

Margulies, M. M. (1991)
Sequence Similarities Between Photosystems I and II. Identification of a Photosystem I Reaction Centre Transmembrane Helix that is Similar to Transmembrane Helix IV of the D2 subunit of Photosystem II and the M Subunit of the Nonsulfur Purple and Flexible Green Bacteria.
Photosynth. Res. **29** :133-147

Marrs, B. (1981)
Mobilisation of the genes for photosynthesis from *Rhodopseudomonas capsulata* by a promiscuous plasmid
J. Bacteriol. **146** :1003-1012

Martin, F.H. and Tinoco, I., Jr. (1980)
DNA-RNA hybrid duplexes containing oligo (dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription
Nucl. Acids Res. **8** :2295-2299

Mason, P.J. and Williams, J.G. (1987)
"Hybridisation in the analysis of recombinant DNA", in "Nucleic acid hybridisation: a practical approach", Eds. Hames, B.D. and Higgins, S.J., IRL Press.
p113-137

McClure, W.R. (1985)
Mechanism and control of transcription initiation in prokaryotes
Ann. Rev. Biochem. **54** :171-204

McDermott, G., Prince, S.M., Freer, A.A., Hawthornthwaite-Lawless, A.M., Papiz, M.Z., (1995)
Cogdell, R.J. and Isaacs, N.W.
Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria
Nature **374** (6 April) :517-521

McGlynn, P. and Hunter, C.N. (1993)
Genetic analysis of the bchC and bchA genes of *Rhodobacter sphaeroides*
Mol. Gen. Genet. **236** :227-234

McGlynn, P. and Hunter, C.N. (1992)
Isolation and Characterisation of a Putative Transcription Factor Involved in the Regulation of the *Rhodobacter sphaeroides puc* BA Operon
J. Biol. Chem. **267** :11098-11103

McGlynn, P., Hunter, C.N., Jones, M.R. (1994)
The *Rhodobacter sphaeroides* PufX protein is not required for photosynthetic competence in the absence of a light harvesting system
FEBS Lett. **349** (3) :349-353

Meinkoth, J., and Wahl, G. (1984)
Hybridisation of Nucleic Acids Immobilised on Solid Supports.
Anal. Biochem. **138** :267-284

Melefors, Ö. and von Gabain, A. (1991)
Genetic studies of cleavage-initiated mRNA decay and processing of ribosomal 9s RNA show that the *Escherichia coli* ams and rne loci are the same
Mol. Microbiol. **5** :857-864

Michaux-Charachon, S., Paillisson, J., Carles-Nurit, M.J., Bourg, G., Allardet-Servant, A., and Ramuz, M. (1993)
Presence of two independent chromosomes in the *Brucella melitensis* 16M genome
J. Bacteriol. **175** :701-705

Michel, H., Epp, O. and Deisenhofer, J. (1986)
Pigment-protein interactions in the photosynthetic reaction centre from *Rhodopseudomonas viridis*
EMBO J. **5** :2445-2451

Michel-Beyerle, M.E., Plato, M., Deisenhofer, J., Michel, H., Bixon, M. and Jortner, J. (1988)
Undirectionally of charge separation in reaction centers of photosynthetic bacteria.
Biochim. Biophys. Acta **932** :52-70

Miller, K.R. (1982)
Three dimensional structure of a photosynthetic membrane
Nature **300** :53-55

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

Monger, T.G. and Parson, W.W. (1977)
Singlet-triplet fusion in *Rhodpseudomonas sphaeroides*
Biochim. Biophys. Acta **460** :393-407

Mosley, C.S., Suzuki, J.Y. and Bauer, C.E. (1994)
Identification And Molecular-genetic Characterization Of A Sensor Kinase Responsible For Coordinately Regulating Light-harvesting And Reaction-center Gene-expression In Response To Anaerobiosis
J. Bacteriol. **176** (24) :7566-7573

Nagarajan, V., Parson, W.W., Gaul, D. and Schenck, C. (1990)
Effect of Specific Mutations of Tyrosine-(M)210 on the Primary Photosynthetic Electron-Transfer Process in *Rhodobacter sphaeroides*
Proc. Natl. Acad. Sci. USA **87** :7888-7892

Nagashima, K.V.P., Matsuura, K., Ohyama, S. and Shimada, K. (1994)
Primary structure and transcription of genes encoding B870 and photosynthetic reaction center apoproteins from *Rubrivivax gelatinosus*
J. Biol. Chem. **269** :2477-2484

Neidle, E.L. and Kaplan, S. (1993)
5-Aminolevulinic acid availability and control of spectral complex formation in hemA and hemT mutants of *Rhodobacter sphaeroides*
J. Bacteriol. **175** :2304-2313

Newbury, S.F., Smith, N.H., Robinson, E.C., Hiles, I.D. and Higgins, C.F. (1987)
Stabilisation of translationally active mRNA by prokaryotic REP sequences
Cell **48** :297-310

Nishimura, M. and Takamiya, A. (1965)
Analysis of light-induced bacteriochlorophyll absorbance change and fluorescence emission in purple bacteria
Biochim. Biophys. Acta **120** :34-44

Noguchi, T., Hayashi, H. and Tasumi, M. (1990)
Factors controlling the efficiency of energy transfer from carotenoids to bacteriochlorophyll in purple photosynthetic bacteria.
Biochim. Biophys. Acta. **1017** :280-290

Norris, J.R., Scheer, H. and Katz, J.J. (1975)
Models for Antenna and Reaction Center Chlorophylls
Ann. N.Y. Acad. Sci. **244** :260-280

Oberto, J., Drlica, K., and Rouvière-Yaniv, J. (1994)
Histones, HMG, HU, IHF - *Même combat*
Biochimie **76** :901-908

Oelze, J. & Drews, G. (1972)
Membranes of Photosynthetic Bacteria
Biochem. Biophys. Acta **265** :209-239

Ogden, R.C. and Adams, D.A. (1987)
"Electrophoresis in agarose and acrylamide gels", in "Guide to Molecular Cloning Techniques", Berger, S.L. and Kimmel (Eds.), A.R., Academic Press Inc.
p61

Ohad, I. & Drews, G. (1982)
"Biogenesis of Photosynthetic Apparatus in Prokaryotes and Eukaryotes." In "Photosynthesis", vol. 2, Ed. Govindjee, Academic Press.
p89-140

Okamura, K., Mitsunori, F., Ito, O., Takamiya, K.I. and Nishimura, M. (1986)
Photophosphorylation And Oxidative-phosphorylation In Intact-cells And Chromatophores Of An Aerobic Photosynthetic Bacterium, *Erythrobacter* Sp Strain Och114
J. Bacteriol. **168** (3) :1142-1146

Okamura, K., Takamiya, K. and Nishimura, M. (1985)
Photosynthetic Electron-transfer System Is Inoperative In Anaerobic Cells Of *Erythrobacter* Species Strain Och-114
Arch. Microbiol. **142** (1) :12-17

Okamura, M.Y. and Feher, G. (1992)
Proton transfer in reaction centres from photosynthetic bacteria
Rev. Biochem. **61** :861-896

Otsuka, J., Miyachi, H., & Horimoto, K. (1992)
Structure model of core proteins in photosystem I inferred from the comparison with those in photosystem II and bacteria; an application of principle component analysis to detect the similar regions between distantly related families.
Biochim. Biophys. Acta **1118** :194-210

Oxender, D.L., (1979)
Attenuation in the *E. coli* tryptophan operon: role of RNA secondary structure involving the tryptophan codon region
Proc. Natl. Acad. Sci. USA **76** :5524-5528

Paddock, M.L., Rongey, S.H., Feher, G. and Okamura, M.Y. (1989)
Pathway of Proton Transfer in Bacterial Reaction Centers: Replacement of Glutamic Acid 212 in the L Subunit by Glutamine Inhibits Quinone (Secondary Acceptor) Turnover
Proc. Natl. Acad. Sci. USA **86** :6602-6606

Parkinson, J.S. (1993)
Signal transduction schemes of bacteria
Cell **73** :857-871

Pearlstein, R.M. (1982)
Exciton migration and trapping in photosynthesis
Photochem. Photobiol. **35** :835-844

Pearlstein, R.M. (1985)
"Exciton state and energy transfer in bacterial membranes. The role of pigment-protein cyclic unit structures", In "Antennas and Reaction Centers of Photosynthetic Bacteria", Chemical Physics Series vol. 42, Ed. Michel-Beyerle, M.E., Springer, berlin
p53

Penfold, R.J. and Pemberton, J.M. (1994)
Sequencing, chromosomal inactivation, and functional expression in escherichia coli of ppsR, a gene which represses carotenoid and bacteriochlorophyll synthesis in *Rhodobacter sphaeroides*
J. Bacteriol. **176** :2869-2876

Penfold, R.J. and Pemberton, J.M. (1991)
A Gene from the Photosynthetic Gene Cluster of *Rhodobacter sphaeroides* Induces Trans Suppression of Bacteriochlorophyll and Carotenoid levels in *Rhodobacter sphaeroides* and *Rhodobacter capsulatus*.
Curr. Microbiol. **23** :259-263

Peters, J. and Drews, G. (1983)
The B870 pigment-protein complex of *Rhodopseudomonas capsulata* contains two different pigment-binding polypeptides
FEMS Microbiol. Lett. **17** :235-237

Pfennig, N. (1989)
"Green Bacteria". In "Bergey's Manual of Systematic Bacteriology" vol. 3, section 18, Ed. Staley, J.T., Bryant, M. P., Pfennig, N. and Holt, J. G., The Williams & Wilkins Co., Baltimore.
p1682-1683

Pfennig, N. and Trüper, H.G. (1973)
"The Rhodospirillales (Phototrophic or Photosynthetic Bacteria)", In "Handbook of Microbiology", vol 1, Eds. Laskin, A.I. & Lechevalier, H.A., CRC Press, Cleveland, Ohio.
p17-27

Pfennig, N. & Trüper, H. G. (1989)
"Anoxygenic Phototrophic Bacteria". In "Bergey's Manual of Systematic Bacteriology" vol. 3, section 18, Ed. Staley, J.T., Bryant, M. P., Pfennig, N. and Holt, J. G., The Williams & Wilkins Co., Baltimore.
p1635-1637

Phillips-Jones, M.K. and Hunter, C.N. (1994)
Cloning and nucleotide sequence of regA, a putative response regulator gene of *Rhodobacter sphaeroides*
FEMS Microbiol. Lett. **116** :269-276

Picorel, R., Belanger, G. and Gingras, G. (1983)
Antenna holochrome B880 of *Rhodospirillum rubrum* S1. Pigment, phospholipid, and polypeptide composition
Biochemistry **22** :2491-2497

Platt, T. (1986)
Transcription termination and the regulation of gene expression
Annu. Rev. Biochem. **55** :339-372

Platt, T. (1981)
Termination of transcription and its regulation in the tryptophan operon of *E. coli*
Cell **24** :10-23

Pollich, M., Jock, S., Klug, G. (1993)
Identification of a gene required for the oxygen-regulated formation of the photosynthetic apparatus of *Rhodobacter capsulatus*
Mol. Microbiol. **10** :749-757

Ponnampalam, S.N., Buggy, J.J., and Bauer, C.E. (1995)
Characterization of an aerobic repressor that coordinately regulates bacteriochlorophyll, carotenoid, and light harvesting-II expression in *Rhodobacter capsulatus*
J. Bacteriol. **177** (11) :2990-2997

Rees, D.C., Komiya, H., Yeates, T.O., Allen, J.P. & Feher, G. (1989)
The Bacterial Reaction Centre as a Model for Membrane Proteins.
Annu. Rev. Biochem. **58** :607-33

Reiss-Husson, F., Arnoux, B., Ducruix, A., Steck, K., Mantele, W., Schiffer, M. and Chang, C. (1990)
H.
"Spectroscopic and Structural Studies of Crystallized Reaction Centres from Wild Type *Rb. sphaeroides* Y", In "Molecular Biology of Membrane-bound Complexes in Phototrophic Bacteria.", Eds. Drews, G. and Dawes, EA., Plenum Press, London.

Rich, P. R. (1984)
Electron and Proton Transfers Through Quinones and Cytochrome *bc* Complexes
Biochim. Biophys. Acta **768** :53-79

Robert, B. and Lutz, M. (1985)
Structures of antenna complexes of several Rhodospirillales from their resonance raman spectra
Biochem. Biophys. Acta **807** :10-

Ronson, C.W., Nison, B.T. and Ausubel, F.M. (1987)
Conserved domains in bacterial regulatory proteins that respond to environmental stimuli.
Cell **49** :579-581

Saegesser, R.S.R. (1992)
Submitted (04-MAY-1992) to the EMBL/GenBank/DDBJ databases, AC N°. X65976
Unpublished

Sambrook, J., Fritsch, E.F. and Maniatis, T. (Eds.) (1989)
"Molecular cloning: A laboratory manual - Second edition", Cold Spring Harbour Laboratory Press.

Sanger, F., Niklen, S. and Coulson, A.R. (1977)
DNA sequencing with chain-terminating inhibitors
Proc. Nat. Acad. Sci. USA **74** :5463-5467

Sauer, K. (1986)
"Photosynthetic light reactions – physical aspects." In "Photosynthesis III: photosynthetic membranes." vol. 19. Eds. Staehelin, L.A. and Arntzen, C.J.
p85-96

Sauer, K. and Austin, L.A. (1978)
Bacteriochlorophyll-protein complexes from the light harvesting antenna of photosynthetic bacteria
Biochemistry **17** :2011-2019

Schuddekopf, K., Henneeke, S., Liese, U., Kutsche, M. and Klipp, W. (1993)
Characterization of anf genes specific for the alternative nitrogenase and identification of nif genes required for both nitrogenases in *Rhodobacter capsulatus*
Mol. Microbiol. **8** :673-684

Sganga, M.W., Aksamit, R.R., Cantoni, G.L. and Bauer, C.E. (1992)
Mutational and nucleotide sequence analysis of S-adenosyl-L-homocysteine hydrolase from *Rhodobacter capsulatus*
Proc. Natl. Acad. Sci. USA **89** :6328-6332

Sganga, M.W. and Bauer, C.E. (1992)
Regulatory Factors Controlling Photosynthetic Reaction Center and Light-harvesting Gene Expression in *Rhodobacter capsulatus*.
Cell **68** :945-954

Shiba, T. (1991)
Roseobacter-litoralis Gen-nov, Sp-nov, And *Roseobacter-denitrificans* Sp-nov, Aerobic Pink-pigmented Bacteria Which Contain Bacteriochlorophyll-*a*
Syst. Appl. Microbiol. **14** (2) :140-145

Shiba, T. (1987)
O_2 Regulation Of Bacteriochlorophyll Synthesis In The Aerobic Bacterium *Erythrobacter*
Plant Cell Physiol. **28** (7) :1313-1320

Shiba, T. and Simidu, U. (1982)
Erythrobacter-longus Gen-nov, Sp-nov, An Aerobic Bacterium Which Contains Bacteriochlorophyll-*a*
Int. J. Syst. Bacteriol. **32** (2) :211-217

Shiba, T., Harashima, K. (1986)
Aerobic Photosynthetic Bacteria
Microbiol. Sci. **3** (12) :376-378

Shimada, H., Iba, K. and Takamiya, K. (1992)
Blue-Light Irradiation Reduces the Expression of *puf* and *puc* Operons of *Rhodobacter sphaeroides* under Semi-Aerobic Conditions.
Plant Cell Physiol. **33** :471-475

Shimada, H., Ohta, H., Masuda, T., Shioi, Y. and Takamiya, K. (1993)
A putative transcription factor binding to the upstream region of the puf operon in *Rhodobacter sphaeroides*
FEBS Letts. **328** :41-44

Shimada, K., Hayashi, H. and Tasumi, M. (1985)
Bacteriochlorophyll-protein Complexes Of Aerobic-bacteria, *Erythrobacter-longus* And *Erythrobacter* Species Och-114
Arch. Microbiol. **143** (3) :244-247

Shiozawa, J.A., Csiszár, K. and Feick, R. (1990)
"Preliminary studies on the operon coding for the reaction center polypeptides in *Chloroflexus aurantiacus*", In "Molecular biology of membrane-bound complexes in phototrophic bacteria", Eds. Drews, G. and Dawes, E.A., Plenum press, New York
p11-18

Shiozawa, J.A., Welte, W., Hodapp, N. and Drews, G. (1982)
Studies on the size and composition of the isolated light-harvesting B800-850 pigment-protein complex of *Rhodopseudomonas capsulata*
Arch. Biochem. Biophys. **213** :473-485

Southern, E.M. (1975)
Detection of specific DNA sequences amoung DNA fragments separated by gel electrophoresis.
J. Mol. Biol. **98** :503-517

Stackebrandt, E., Murray, G. E. & Trüper, H. G. (1988)
Proteobacteria classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives."
Int. J. Syst. Bacteriol. **38** :321-325

Staley, J.T., Bryant, M. P., Pfennig, N. & Holt, J. G. (Eds.) (1989)
"Bergey's Manual of Systematic Bacteriology", volumes 1-4, The Williams & Wilkins Co., Baltimore.

Staley, J. T. & Krieg, N. R. (1989)
"Classification of prokaryotic organisms: an overview", In "Bergey's Manual of Systematic Bacteriology" vol. 3, Ed. Staley, J.T., Bryant, M. P., Pfennig, N. and Holt, J. G., The Williams & Wilkins Co., Baltimore.
p1601-1607

Stern, M.J., Ames, G. F.-L., Smith, N.H., Robinson, E.C. and Higgins, C.F. (1984)
Repetative extragenic palindromic sequences: a major component of the bacterial genome
Cell **37** :1015-1026

Subir, K.B. (1963)
'Media for anaerobic growth of photosynthetic bacteria", in "Bacterial photosynthesis", Gest, H., Pietro, A.S. and Vernon, L.P. (Eds.), The Antioch press, Yellow Springs Ohio.

Suwanto, A. and Kaplan, S. (1992)
Chromosome transfer in *Rhodobacter sphaeroides*: Hfr formation and genetic evidence for two unique circular chromosomes.
J. Bacteriol. **174** (4) :1135-1145

Suwanto, A. and Kaplan, S. (1989)a
Physical and Genetic Mapping of the *Rhodobacter sphaeroides* 2.4.1 Genome : Genome Size, Fragment Identification and Gene Localisation.
J. Bacteriol. **171** (11) :5840-5849

Suwanto, A. and Kaplan, S. (1989)b
Physical and Genetic Mapping of the *Rhodobacter sphaeroides* 2.4.1 Genome : Presence of Two Unique Circular Chromosomes.
J. Bacteriol. **171** (11) :5850-5859

Tabor, S. and Richardson, C.C. (1989)
Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by *in vitro* mutagenesis
J. Biol. Chem. **264** :6447-6458

Tadros, M.H. (1990)
In "Molecular biology of membrane-bound complexes in photosynthetic bacteria", Eds. Drews, G. and Dawes, E.A., Plenum Press, New York.
p19-31

Tadros, M.H. and Waterkamp, M.K. (1989)
Multiple copies of the coding regions for the light-harvesting B800-850 α- and β-polypeptides are present in the *Rhodopseudomonas palustris* genome
EMBO J. **8** :1303-1308

Tadros, M.H., Katsiou, E., Hoon, M.A., Yurkova, N. and Ramji, D.P. (1993)
Cloning of a new antenna gene cluster and expression analysis of the antenna gene family of *Rhodopseudomonas palustris*.
Eur. J. Biochem. **217** :867-875

Takahashi, E. and Wraight, C.A. (1992)
Proton and Electron Transfer in the Acceptor Quinone Complex of *Rhodobacter sphaeroides* Reaction Centers: Characterization of Site-Directed Mutants of the Two Ionizable Residues, Glu^{L212} and Asp^{L213}, in the Q_B Binding Site
Biochemistry **31** :855-866

Takamiya, K., Iba, K. and Okamura, K. (1987)
Reaction Center Complex From An Aerobic Photosynthetic Bacterium, *Erythrobacter* Species Och-114
Biochim. Biophys. Acta **890** (2) :127-133

Takemoto, J. & Huang Kao, M. Y. C. (1977)
Effects of Incident Light Levels on Photosynthetic Membrane Polypeptide Composition and Assembly in *Rhodopseudomonas sphaeroides*
J. Bacteriol. **129** (2) :1102-1109

Taraseviciene, L., Miczak, A. and Apirion, D. (1991)
The gene specifying RNase E and a gene affecting mRNA stability (ams) are the same gene
Mol. Microbiol. **5** :851-855

Taylor, D.P., Cohen, S.N., Clark, W.G. and Marrs, B.L. (1983)
Alignment of genetic and restriction maps of the photosynthesis region of the *Rhodopseudmonas capsulata* chromosme by conjugation mediated marker rescue technique
J. Bacteriol. **154** :580-590

Thornber, J.P. (1986)
"Biochemical characterization and structure of pigment-proteins of photosynthetic organism", section 3, In "Photosynthesis III", Eds. Staehelin, C. A., Arntzen, C.
p98

Thornber, J.P. (1970)
Photochemical reactions of purple bacteria as revealed by studies of 3 spectrally different carotenobacteriochlorophyll-protein complexes isolated from *Chromatium* strain D
Biochemistry **9** :2688-2698

Thornber, J.P., Cogdell, R.J., Pierson, B.K., and Seftor, R.E.B. (1983)
Pigment-protein complexes of purple photosynthetic bacteria: an overview
J. Cell. Biochem. **23** :159-169

Thornber, J.P., Trosper, T.L. and Strouse, C.E. (1978)
"Bacteriochlorophyll in vivo: relation of spectral forms to specific membrane components", In "The Photosynthetic Bacteria", Eds. Clayton, R. K. and Sistrom, W. R., Plenum Press, New York.

Tichy, H.V., Albien, K-U., Gad'on, N. and Drews, G. (1991)
Analysis of the *Rhodobacter capsulatus* puc Operon : the pucC Gene Plays a Central Role in the Regulation of LHII (B800-850 complex) Expression.
EMBO J. **10** (10) :2949-2955

Tichy, H.V., Oberle, B., Stehle, E., Schiltz, E. and Drews, G. (1989)
Genes downstream from pucA are essential for formation of the B800-850 complex of *Rhodobacter capsulatus*.
J. Bacteriol. **171** :4914-4922

Tiede, D. M., Budil, D.E., Tang, J., El-Kabbani, O., Norris, J. R. Chang, C.-H., and Schiffer, (1988) M.
"Symmetry breaking structures involved in the docking of cytochrome c and primary electron transfer in reaction centers of *Rhodobacter sphaeroides*", In "The photosynthetic bacterial reaction centre", Eds. Breton, J. and Vermeglio, A., NATO AS1 Series A: vol 149 Life Sciences, Plenum Press, New York.
p13-20

Toussaint, B., Bose, C., Richaud, P., Colbeau, A. and Vignais, P.M. (1991)
A mutation in a *Rhodobacter capsulatus* gene encoding an integration host factor-like protein impairs *in vivo* hydrogenase expression
Proc. Natl. Acad. Sci. USA **88** :10749-10753

Toussaint, B., Bose, C., Richaud, P., Colbeau, A. and Vignais, P.M. (1993)
Purification of the integration host factor homolog of *Rhodobacter capsulatus:* cloning and sequencing of the hip gene, which encodes the beta subunit
J. Bacteriol. **175** :6499-6504

Toussaint, B., David, L., de Sury d'Aspremont, R., and Vignais, P.M. (1994)
The IHF proteins of *Rhodobacter capsulatus* and *Pseudomonas aeruginosa*
Biochimie **76** :951-957

Trumpower, B.L. (1990)
Cytochrome *bc*1 complexes of microorganisms
Microbiol. Rev. **54** 101-129

van Grondelle, R. (1985)
Excitation energy transfer, trapping and annilation in photosynthetic systems
Biochim. Biophys. Acta **811** :147-195

van Grondelle, R., Dekker, J.P., Gillbro, T. and Sundstrom, V. (1994)
Energy transfer and trapping in photosynthesis
Biochim. Biophys. Acta **1187** :1-65

Varga, A.R. and Kaplan, S. (1993)
Synthesis and stability of reaction center polypeptides and implications for reaction center assembly in *Rhodobacter sphaeroides*
Journal of Biological Chemistry **268** (26) :19842-19850

Wagner-Huber, R., Brunisholz, R.A., Bissig, I., Frank, G., Suter, F. & Zuber, H. (1992)
The Primary Structure of the Antenna Polypeptides of *Ectothiorhodospira halochloris* and *Ectothiorhodospira halophila* : Four Core-type Antenna Polypeptides in *E. halochloris* and *E. halophila*.
Eur. J. Biochem. **205** :917-925

Watanabe, Y., Feick, R.G. and Shiozawa, J.A. (1995)
Cloning and Sequencing of the Genes Encoding the Light-harvesting B806-866 Polypeptides and Initial Studies on the Transcriptional Organization of puf2b, puf2a and puf2c in *Chloroflexus-aurantiacus*
Arch. Microbiol. **163** :124-130

Wechsler, T.D., Brunisholz, R.A., Frank, G., Suter, F. and Zuber, H. (1987)
The complete amino acid sequence of the antenna polypeptide B806-866 b from the cytoplasmic membrane of the green bacterium *Chloroflexus aurantiacus*
FEBS Lett. **210** :189-194

Wellington, C.L. and Beatty, J.T. (1989)
Promotor mapping and nucleotide sequence of the bchC bacteriochlorophyll biosynthesis gene from *Rhodobacter capsulatus*
Gene **83** :251-261

Wellington, C.L. and Beatty, J.T. (1991)
Overlapping mRNA Transcripts of Photosynthesis Gene Operons in *Rhodobacter capsulatus*
J. Bacteriol. **173** (4) :1432-1443

Wellington, C.L., Bauer, C.E. and Beatty, J.T. (1992)
Photosynthesis gene superoperons in purple nonsulfur bacteria: the tip of the iceberg?
Can. J. Microbiol. **38** :20-27

Wellington, C.L., Taggart, A.K.P. and Beatty, J.T. (1991)
Functional significance of overlapping transcripts of crtEF, bchCA and puf photosynthesis gene operons in *Rhodobacter capsulatus*
J. Bacteriol. **173** :2954-2961

Westerhuis, W.H.J., Farchaus, J.W. and Niederman, R.A. (1993)
Altered spectral properties of the B875 light-harvesting pigment-protein complex in a *Rhodobacter sphaeroides* mutant lacking pufX
Photochem. Photobiol. **58** :460-463

Wiessner, C., Dunger, I. and Michel, H. (1990)
Structure and transcription of the genes encoding the B1015 light-harvesting complex beta and alpha subunits and the photosynthetic reaction center L, M, and cytochrome c subunits from *Rhodopseudomonas viridis*
J. Bacteriol. **172** (6) :2877-2887

Willems, A., Gillis, M. & De Ley, J. (1991)
Transfer of *Rhodocyclus gelatinosus* to *Rubrivivax gelatinosus* gen. nov., comb. nov., and phylogenetic relationships with *Leptothrix*, *sphaerotilus natans*, *Pseudomonas saccharophila*, and *Alcaligenes latus*
Int. J. Syst. Bacteriol. **41** (1) :65-73

Williams, J.C., Steiner, L.A. and Feher, G. (1986)
Primary structure of the reaction center from *Rhodopseudomonas sphaeroides*
Proteins: Struct. Funct. Genet. **1** :312-325

Williams, J.C., Steiner, L.A., Feher, G. and Simon, M.I. (1984)
Primary structure of the L subunit of the reaction center from *Rhodopseudomonas sphaeroides*
Proc. Natl. Acad. Sci. USA **81** :7303-7307

Williams, J.C., Steiner, L.A., Ogden, R.C., Simon, M.I. and Feher G. (1983)
Primary structure of the M subunit of the reaction center from *Rhodopseudomonas sphaeroides*
Proc. Natl. Acad. Sci. USA **80** :6505-6509

Woese, C. R. (1987)
Bacterial Evolution
Microbiol. Rev. **51** (2) :221-271

Woese, C. R. & Fox, G. E. (1977)
Phylogenetic structure of the prokaryotic domain: The primary kingdoms
Proc. Natl. Acad. Sci. USA **74** (11) :5088-5090

Woese, C. R., Gibson, J. & Fox, E. (1980)
Do genealogical patterns in purple photosynthetic bacteria reflect interspecific gene transfer?
Nature **283** (January) :212-214

Woese, C. R., Stackebrandt, E., Weisburg, W. G., Paster, B. J., Madigan, M. T., Fowler, V. J., (1984)a Hahn, C. M., Blanz, P., Gupta, R., Nealson, K. H. & Fox, G. E.
The phylogeny of purple bacteria: the alpha subdivision
Syst.Appl. Microbiol. **5** :315-326

Woese, C. R., Weisburg, W. G., Paster, B. J., Hahn, C. M., Tanner, R. S., Krieg, N. R., (1984)b Koops, H. P., Harms, H. & Stackebrandt, E.
The phylogeny of purple bacteria: the beta subdivision
Syst.Appl. Microbiol. **5** :327-336

Wong, H.C. and Chang, S. (1986)
Identification of a positive retroregulator that stabilises mRNA's in bacteria
Proc. Natl. Acad. Sci. USA **83** :3233-3237

Wu, Y.Q., Macgreggor, B.J., Donohue, T.J., Kaplan, S. and Yen, B. (1991)
Genetic and physical mapping of the *Rhodobacter sphaeroides* photosynthetic gene cluster from R-prime pWS2.
Plasmid **25** :163-176

Yang, Z., and Bauer, C.E. (1990)
Rhodobacter capsulatus genes involved in early steps of the bacteriochlorophyll biosynthetic pathway
J. Bacteriol. **172** :5001-5010

Yeates, T. O., Komiya, H., Chirino, A., Rees, D. C., Allen, J. P. and Feher, G. (1988)
Structure of the Reaction Center form *Rb. sphaeroides* R26 and 2.4.1.: Protein cofactor (bacteriochlorophyll, bacteriopheophytin and carotenoid) interactions.
Proc. Natl. Acad. Sci. USA **85** :7993-7997

Yeates, T. O., Komiya, H., Rees, D. C., Allen, J. P. and Feher, G. (1987)
Structure of the Reaction Center form *Rb. sphaeroides* R26: Membrane-Protein Interactons
Proc. Natl. Acad. Sci. USA **84** :6438-6442

Yildiz, F.H., Gest, H., Bauer, C.E. (1992)
Conservation of the photosynthesis gene cluster in *Rhodospirillum centenum*
Mol. Microbiol. **6** :2683-2691

Young, D.A., Bauer, C.E., Williams, J.C., and Marrs, B.L. (1989)
Genetic evidence for superoperonal organization of genes for photosynthetic pigments and pigment-binding proteins in *Rhodobacter capsulatus*.
Mol. Gen. Genet. **218** :1-12

Youvan, D.C. and Ismail, S. (1985)
Light harvesting II (B800-850 complex) structural genes from *Rhodopseudomonas capsulata*
Procad. Natl. Acad. Sci. USA. **82** :58-62

Youvan, D.C., Bylina, E.J., Alberti, M., Begusch, H. and Hearst, J.E. (1984)
Nucleotide and deduced polypeptide sequence of the photosynthetic reaction center, B870 antenna, and flanking polypeptides from *Rhodopseudomonas capsulata*
Cell. **37** :949-957

Youvan, D.C., Ismail, S. and Bylina, E.J. (1985)
Chromosonal deletion and plasmid complementation of the photosynthetic reaction center and light harvesting genes from *Rhodopseudomonas capsulata*
Gene. **33** :19-30

Zablen, L.B., Kissel, K.S., Woese, C.R. & Buetow, D.E. (1975)
Phylogenetic origin of the chloroplast and prokaryotic nature of its ribosomal RNA
Proc. Natl. Acad. Sci. USA **72** :2418-2422

Zhu, S.Y. and Kaplan, S. (1985)
Effects of light, oxygen, and substrates on steady state levels of mRNA coding for ribulose-1,5-bisphosphate carboxylase and light-harvesting and reaction center polypeptides in *Rhodopseudomonas sphaeroides*
J. Bacteriol **162** :925-932

Zhu, Y.S. and Hearst, J.E. (1986)
Regulation of Expression of Genes for Light-harvesting Antenna Proteins LH-I and LH-II; Reaction Centre Polypeptides RC-L, RC-M, and RC-H; and Enzymes of Bacteriochlorophyll and Carotenoid Biosynthesis in *Rhodobacter capsulatus* by Light and Oyxgen.
Proc. Natl. Acad. Sci. USA **83** :7613-7617

Zhu, Y.S. and Hearst, J.E. (1988)
Transcription of oxygen-regulated photosynthetic genes requires DNA gyrase in *Rhodobacter capsulatus*
Proc. Natl. Acad. Sci. USA **85** :4209-4213

Zhu, Y.S., Kiley, P.J., Donohue, T.J. and Kaplan, S. (1986)
Origin of the mRNA Stoichiometry of the *puf* Operon in *Rhodobacter sphaeroides*.
J. Biol. Chem. **261** :10366-10374

Zinth, W., Knapp, E.W., Fischer, S.F., Kaiser, W., Deisenhofer, J. and Michel, H. (1985)
Correlation of structural and spectroscopic properties of a photosynthetic reaction center
Chem. Phys. Letts. **119** :1-4

Zsebo, K.M. and Hearst, J.E. (1984)
Genetic-Physical Mapping of a Photosynthetic Gene Cluster from *R. capsulata*
Cell **37** :937-947

Zuber, H. (1990)
"Considerations on the Structural Principles of the Antenna Complexes of Phototrophic Bacteria." In "Molecular Biology of Membrane-bound Complexes in Phototrophic Bacteria." Eds. Drews, G. and Dawes, EA., Plenum Press, New York.
p161-180

Zuber,H. (1986)
"Primary structure and function of the Light-Harvesting polypeptides from cyanobacteria, Red algae, and purple photosynthetic bacteria", section 6.1, In "Photosynthesis III", Eds. Staehelin, C. A., Arntzen, C.
p238-251

Zuber, H. (1985)
Structure and function of light-harvesting complexes and their polypeptides
Photochem. Photobiol. **42** :821-

Zuber, H. & Brunisholz, R. A. (1991)
"Structure and Function of Antenna Polypeptides and Chlorophyll-Protein Complexes: Principles and Variability." In "Chlorophylls" sect. 3.6, Ed. Scheer, H., CRC Press, London.
p627-703

Zuber, H., Brunisholz, R. and Sidler, W. (1987)
"Structure and function of light-harvesting pigment-protein complexes", In "Photosynthesis", Ed. Amesz, J., Elsevier Science, Amsterdam
p233

Zucconi, A.P. and Beatty, J.T. (1988)
Posttranscriptional regulation by light of steady-state levels of mature B800-850 light-harvesting complex in *Rhodobacter capsulatus*
J. Bacteriol. **170** :877-882

Zuerner, R.L., Herrmann, J.L. and Saint Girons, I. (1993)
Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity
J. Bacteriol. **175** :5445-5451