This is to certify that the

dissertation entitled

GENETIC ANALYSIS OF THE PATHWAY FOR THE BIOSYNTHESIS OF THE PLANT SULFOLIPID IN THE PURPLE BACTERIUM *RHODOCATER SPHAEROIDES*

presented by

CHRISTOPH BENNING

has been accepted towards fulfillment of the requirements for

PH.D. degree in GENETICS

Major professor

Date MARCH 14, 1991

DATE DUE	DATE DUE	DATE DUE
MAR 2 9 2001 0411201		AUG 2 2 2004
FEB 1 6 2009		

GENETIC ANALYSIS OF THE PATHWAY FOR THE BIOSYNTHESIS OF THE PLANT SULFOLIPID IN THE PURPLE BACTERIUM *RHODOBACTER SPHAEROIDES*

By

Christoph Benning

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Genetics Interdepartmental Doctoral Program
MSU-DOE Plant Research Laboratory

1991

ABSTRACT

GENETIC ANALYSIS OF THE PATHWAY FOR THE BIOSYNTHESIS OF THE PLANT SULFOLIPID IN THE PURPLE BACTERIUM *RHODOBACTER SPHAEROIDES*

By

Christoph Benning

Approximately thirty years ago, Benson and collaborators discovered the plant sulfolipid and determined its structure to be 6-sulfo-α-D-quinovosyl diacylglycerol. Since then, relatively little progress has been made towards the elucidation of the biosynthetic pathway and the function of sulfolipid.

To solve this classical problem, the purple non-sulfur bacterium *Rhodobacter sphaeroides* was used as a model organism. The presence of authentic sulfolipid in this bacterium was confirmed by comparative mass spectrometry of bacterial and spinach sulfolipid. Mutants of *R. sphaeroides* were isolated that were deficient in sulfolipid biosynthesis. The analysis was focused on three mutants carrying defects in the *sqdA*, *sqdB*, and *sqdC* genes, respectively. The *sqdA* and *sqdC* mutants were found to contain ^{35}S-sulfate labelled, novel compounds, possibly precursors of sulfolipid biosynthesis.

The *sqdA, sqdB,* and *sqdC* genes were isolated using a complementation assay. The *sqdA* mutant was complemented by a different cosmid than the *sqdB* and *sqdC* mutants. Small subclones of the cosmids complementing the *sqdA* and *sqdB* mutants were sequenced, and the

putative coding sequences for the two genes were identified. The *sqdA* gene encodes a 33.6-kD protein, not homologous to any known protein. The *sqdB* gene codes for a 46-kD protein. Its N-terminal amino acid sequence shows homology to the N-terminus of UDP-glucose epimerase from bacteria and fungi, suggesting that the *sqdB* gene product is involved in sulfolipid biosynthesis on the level of UDP-sulfoquinovose or some other sugar nucleotide. The *sqdC* gene was found to be located directly downstream of the *sqdB* gene, suggesting that *sqdB* and *sqdC* are organized in an operon.

To analyze putative precursors of sulfolipid biosynthesis, an *in vitro* sulfolipid synthesis system was developed. The incorporation of ^{35}S-PAPS into authentic sulfolipid by extracts of *R. sphaeroides* wild type cells was observed.

Based on the data provided in this work, definitive experiments can be designed, which will provide a solution to the problem of sulfolipid biosynthesis and function in the near future.

To

My Parents

ACKNOWLEDGMENTS

I thank Chris Somerville for his support throughout my work. He provided me with excellent advice, and taught me how to approach difficult, scientific problems. I am grateful for the support and advice I received from the members of my advisory committee: Hans Kende, John Ohlrogge, Peter Wolk, and Lee McIntosh. I would also like to thank all the members of the lab and all my friends in the PRL for their advice and their helpful discussions. I thank Doug Gage and Dr. Huang for analyzing samples by mass-spectrometry, and for their help concerning the interpretation of the spectra.

I thank my parents, who gave me the freedom to find my own way, and who supported all my decisions regardless of the hardships these brought upon them. I would also like to thank Susi and Urs for their love and patience they gave during the preparation of this work.

I am grateful for a stipend awarded to me by the German Scholarship Foundation, which made it possible for me to get started in E. Lansing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

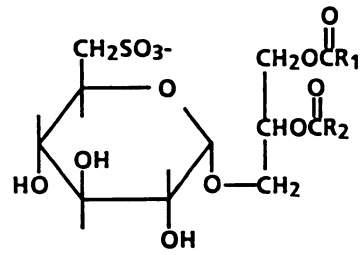
ADP	Adenosine diphosphate
APS	Adenosine phospho-sulfate
ATP	Adenosine triphosphate
CAD	Collision activated dissociation
CDP	Cytosine diphosphate
DGD	Digalactosyl diacylglycerol
EDTA	(Ethylenedinitrilo)tetraacetic aid
FAB	Fast atom bombardement
FAM	Fatty acid methylester
GDP	Guanosine diphosphate
MS	Mass spectrometry
M-H	Molecular ion minus one proton
PAPS	Phosphoadenosine phosphosulfate
pfu	plaque-forming unit
TLC	Thin layer chromatography
TLE	Thin layer electrophoresis

CHAPTER 1

INTRODUCTION

DISCOVERY AND STRUCTURE OF THE PLANT SULFOLIPID

Approximately thirty years ago, Benson and collaborators discovered that $^{35}SO_4$-labelled extracts from various higher plants, algae, and a purple non-sulfur bacterium harbored large amounts of a sulfur containing lipid (2). They determined the structure to be 1'-O-(6'-deoxy-α-D-glucosyl-6'-sulfonic acid)-3-O- diacylglycerol (3). The 6-deoxy-6-sulfo-glucose is also known as sulfoquinovose (Figure 1-1).



6-Sulpho-α-D-Quinovosyl Diacylglycerol
(Plant Sulpholipid)

Figure 1-1: Structure of the plant sulfolipid, 6-sulfo-α-D-quinovosyl diacyl-glycerol.

Since the elucidation of the structure of the sulfolipid, most studies concerning the function and mechanism of synthesis have been carried out with higher plants. Thus, much of the discussion that follows concerns the plant work since virtually nothing has been done with bacteria. The two major fatty acids found in sulfolipid in higher plants are palmitic acid and linolenic acid (4). In *Rhodobacter sphaeroides*, a purple non-sulfur bacterium, the place of linolenic acid is taken by vaccenic acid, a monounsaturated 18-carbon fatty acid typically found in bacteria (40). Positional analysis of the fatty acid composition of sulfolipid in higher plants reveals an interesting difference in comparison to the galactolipid MGD, the most abundant chloroplast lipid. In spinach, the predominant 16-carbon fatty acid found in MGD is hexadecatrienoic acid, which is confined to the sn-2 position of the diacylglycerol moiety, while the most abundant 16-carbon fatty acid in spinach sulfolipid is palmitic acid, which can be found in the sn-2 as well as in the sn-1 position of the glycerol backbone. This unique fatty acid composition of sulfolipid raises a question about the origin of its diacylglycerol moiety. It is generally accepted that in some higher plants, including spinach, the diacylglycerol moiety of the galactolipids is derived from two distinctive pools as a result of the eucaryotic pathway, located in the endoplasmatic reticulum, and the procaryotic pathway, operating inside the chloroplast (33). Due to different substrate specificities of the acyltransferases belonging to one or the other pathway, diacylglycerol derived from the chloroplast pool contains 16-carbon fatty acids predominantly in the sn-2 position, while those molecular

species derived from the endoplasmatic reticulum pool contain 18-carbon fatty acids in the sn-2 position. Higher plants, such as cucumber, which have a reduced procaryotic pathway, virtually lack galactolipids with 16-carbon fatty acids in the sn-2 position of the glycerol backbone. Therefore, by analysis of the fatty acids found in the sn-2 position of the chloroplast lipids, the presence of the two pathways in a particular plant species, as well as their contribution to the synthesis of each lipid, can be estimated. In analogy to the galactolipids, attempts have been made to determine the origin of the diacylglycerol found in sulfolipid, isolated from various higher plants, by positional analysis of the fatty acid composition (4). Based on these data, it was concluded that in spinach the majority of the diacylglycerol in sulfolipid is derived from the chloroplast pool. However, the molecular mechanisms that determine the unique fatty acid composition of each chloroplast lipid are not understood.

OCCURRENCE AND DISTRIBUTION OF THE PLANT SULFOLIPID

Sulfoquinovosyl diacylglycerol occurs in all higher plants, ferns and mosses. Although it seems that the sulfolipid is mainly associated with photosynthetically active membranes, it can be found in small amounts (presumably in plastid membranes) in nonphotosynthetic tissue, such as roots (38), potato tubers (7), or apples (8). It is also present in substantial amounts in etiolated tissue (39). In addition it has been found in various algae such as hemoflagellata, phytoflagellata, red algae, brown algae, and green algae. The

sulfolipid has also been reported to be a constituent of cyanobacteria and some species of purple non-sulfur bacteria. A compilation of the distribution of the sulfolipid in plants and bacteria has been presented by Haines (11).

The purple non-sulfur bacteria can be divided into two groups, depending on whether or not they contain sulfoquinovose diacylglycerol. Species known to contain the sulfolipid are *Rhodobacter sphaeroides* (40) and *Rhodomicrobium vannielii* (29). Benson *et al.* also reported the sulfolipid to be present in labelled extracts from *Rhodospirillum rubrum* (2). This finding could not be confirmed by Wood *et al.* (40), who were also not able to measure detectable levels of the sulfolipid in *R. gelatinosus*, *R. capsulatus*, and *R. palustris*. However, Wood *et al.* did not use isotope labelling experiments and might have missed trace amounts of sulfolipid in these bacteria (11). Thus, it is not certain whether all photosynthetic organisms have the sulfolipid. However, no organism with oxygenic photosynthesis has been found to lack it.

Various photosynthetic organisms have been compared with respect to their sulfolipid content in a careful study by Radunz (31). The amounts, measured in percent of ether-extractable lipids, vary from 2.6% in *R. sphaeroides* up to 18.3% in the brown alga *Fucus vesiculosus*. Leaves of higher plants usually contain around 5% sulfolipid. Within photosynthetic tissue sulfolipid is found exclusively in the chloroplast membranes (28). The detailed subchloroplastic localization of the sulfolipid remains unclear. It has been found in association with the granal lamellae of the chloroplast as well as with the envelope of the chloroplast (6). In a study by Radunz and Berzborn (32),

who employed polyclonal antibodies against *Fucus* sulfolipid, it was found that the antibodies, which were also reactive against *Antirrhinum* sulfolipid, were able to agglutinate the lamellar membranes isolated from *Antirrhinum* chloroplast only indirectly, while mechanically disrupted lamellae could be directly agglutinated. However, the conclusion drawn by the authors, based on these data, that the sulfolipid is predominantly located on the inner surface of the thylakoids needs to be supported by more data. The intracellular localization of sulfoquinovosyl diacylglycerol in the bacterium *R. sphaeroides* has not been investigated. It would be interesting to know whether the occurrence of the sulfolipid is confined to the intracellular membranes that carry the photosynthetic apparatus, or whether it can also be found in the plasma membrane or even in the outer membrane.

In a photosynthetically active leaf of a higher plant, approximately the same amounts of sulfur are bound in sulfolipid as in proteins (14). This implies that a major proportion of organic sulfur on earth exists as sulfolipid, which therefore plays an important role in the global sulfur cycle. In view of the abundance of this compound in plant membranes, it is a quantitatively important component of human and animal diets. Various aspects of degradation of plant sulfolipid have been reviewed by Harwood (15).

PROPOSED BIOSYNTHETIC PATHWAYS FOR THE PLANT SULFOLIPID

Although the plant sulfolipid was discovered over 30 years ago, the biosynthetic pathway remains unknown. Studying sulfur metabolism in general

seems to be very complicated, due to the diversity and reactivity of sulfur containing compounds. The biosynthesis of the sulfoquinovose head group brings together two plant-specific processes: first, photosynthetic sugar biosynthesis and, second, photosynthetic sulfur assimilation. At least by now, we have quite a deep understanding of sugar biosynthesis in plants, and many of the ideas about sulfolipid biosynthesis were derived from our knowledge about this process. The elucidation of the pathway for sugar biosynthesis took many years and was favored by the fact that most of the intermediates are highly abundant in photosynthetically active leaves. Intermediates of sulfolipid biosynthesis, on the other hand, are not quite as abundant.

Several possible pathways for the biosynthesis of sulfoquinovosyl diacylglycerol have been proposed on the basis of "biosynthetic parsimony" rather than the availability of any direct evidence (15, 27). A summary of the most popular ideas is shown in Figure 1-2. The experimental evidence on which Fig. 1-2 is based can be summarized as follows:

Shibuya *et al.* (34) discovered various sulfonic acids in extracts from ^{35}S-sulfate-labelled *Chlorella* cells, including sulfoacetic acid, sulfolactic acid, cysteic acid, and a nucleoside diphosphate sulfoquinovose. Later, in an *in vivo* labelling study employing *Euglena gracilis*, Davis *et al.* (5) showed that labelled sulfate and L-cysteic acid were incorporated by this organism into sulfolipid. However, no methods were employed in these studies to prevent incorporation from degradation products of cysteic acid. In addition,

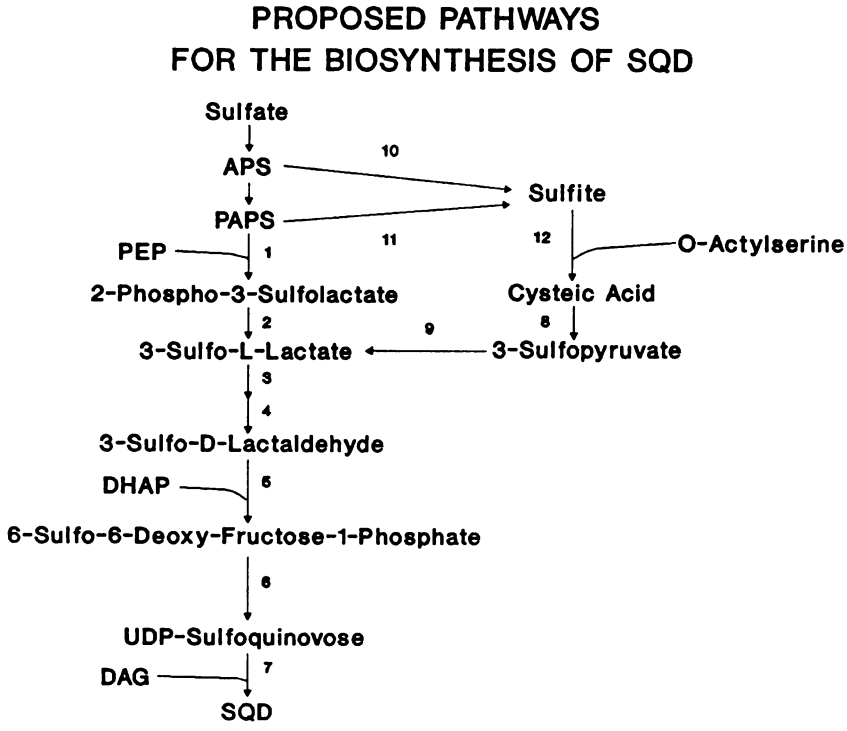


Figure 1-2: Proposed pathways for the biosynthesis of the plant sulfolipid (adopted from 15). The numbers refer to reactions mentioned in the text.

molybdate, which is known to inhibit ATP-sulfurylase, reduced the incorporation of labelled sulfate into sulfolipid, suggesting that adenosyl-1-phosphosulfate (APS) or 3-phosphoadenosyl-1-phosphosulfate (PAPS) might be precursors of sulfolipid biosynthesis. Based on these data, Davis *et al.* (5) outlined a pathway (Fig.1-2, steps 1-7) starting with the reaction of phospho-enolpyruvate and PAPS leading to the formation of 3-L-sulfolactic acid, which might be converted to 3-sulfo-D-lactaldehyde via a proposed epimerase and a dehydrogenase. Sulfolactaldehyde and dihydroxyacetone phosphate could then condense via an aldolase reaction, forming 6-sulfo-6-deoxy-D-fructose-1-phosphate, which might be a direct precursor of 6-sulfoquinovose. It was first suggested by Benson (3) that activated 6-sulfoquinovose in the form of a nucleoside diphosphate sulfoquinovose and diacylglycerol might react to give sulfoquinovose diacylglycerol (Figure 1-2, step 7). The pathway outlined above has some similarity with glycolysis, and therefore the term "sulfoglycolytic pathway" was coined by Benson (3), implying that the sulfonic acid group is equivalent to a phosphate ester. This could be misleading, since the sulfonic acid group might be more similar to a carboxyl group, as pointed out by Mudd *et al.* (27).

Davies *et al.* (5) assumed that the incorporation of cysteic acid into sulfolipid by *Euglena gracilis* was due to a side reaction. Haines proposed that cysteic acid feeds into the pathway on the level of 3-sulfo-L-lactic acid (steps 8 and 9, Figure 1-2) via a transaminase reaction (11). He suggested some alternative, initial reactions that would give cysteic acid a more central role in

the pathway (Figure 1-2, steps 10,11, and 12). In that case, PAPS or APS would be the precursor of sulfite, which could displace O-acetyl from O-acetyl serine analogous to the reactions of thiols with O-acetyl serine (18). The proposal that cysteic acid is a direct precursor of sulfolipid biosynthesis has been controversial. Harwood reported that labeled cysteic acid can be incorporated into sulfolipid by intact alfalfa plants (13), while Mudd *et al.* could not find evidence for incorporation of labeled cysteic acid into sulfolipid by spinach seedlings (26).

Isolated chloroplasts from spinach are capable of incorporating ^{35}S-sulfate into sulfoquinovosyl diacylglycerol, as was first shown by Haas *et al.* (10). Rates of 1.7 pmol/mg chlorophyll per hour were observed in the light. Later, Kleppinger-Sparace *et al.* demonstrated rates of up to 700 pmol/ mg chlorophyll per hour for the incorporation of ^{35}S-sulfate into sulfolipid by isolated spinach chloroplast in the light, which approached the rates observed for phosphatidyl glycerol synthesis in the same system (21). The incorporation of ^{35}S-sulfate into sulfolipid was observed in the dark, when the isolated spinach chloroplast was supplied with ATP either directly, or indirectly via the triose phosphate shuttle, which provided dihydroxyacetone phosphate as the substrate for ATP biosynthesis inside the chloroplast (22). The general conclusion drawn from these results is that all the enzymes required for the biosynthesis of the sulfoquinovose head group are located inside the chloroplast. Kleppinger-Sparace *et al.* (21) measured the effect of various compounds on the incorporation of ^{35}S-sulfate into sulfolipid by intact spinach

chloroplasts in the light. ATP slightly increased the incorporation, while the addition of cysteic acid, cysteine, and other sulfhydryl compounds had no effect. In labelling experiments, using ^{14}C-acetate as a substrate for sulfolipid biosynthesis in isolated spinach chloroplasts, Joyard *et al.* (20) showed that these chloroplasts were capable of synthesizing the diacylglycerol moiety as well as the sulfoquinovose head group *de novo*. Positional analysis of the fatty acid composition of the labeled sulfolipid revealed that palmitic acid was preferentially located in the sn-2 position of the glycerol backbone, typical for sulfolipid synthesized by the procaryotic pathway inside the chloroplast (see above; and reference 4). During the same study it was found that UDP-galactose, the precursor of galactolipid biosynthesis, strongly inhibited the incorporation of labelled sulfate into the sulfolipid. This has been interpreted as competition by UDP-galactose with a nucleoside diphosphate sulfoquinovose for diacylglycerol. Labelling of isolated chloroplasts with ^{14}C-acetate in the presence of UDP-galactose also resulted in a slower rate of incorporation of label into sulfolipid as compared to the controls without UDP-galactose, while at the same time the biosynthesis of galactolipids was stimulated (20).

Direct evidence for the involvement of a nucleoside diphosphate as the last precursor of sulfolipid biosynthesis, has recently been obtained by Heinz *et al.*, who synthesized 6-sulfoquinovosyl-1-phosphate and UDP-sulfoquinovose (16, 17). Chloroplast envelopes were prelabeled with ^{14}C-glycerol to form radioactive diacylglycerol. Addition of UDP-sulfoquinovose

stimulated the incorporation of label into sulfolipid. The GDP derivative were only half as active as UDP-sulfoquinovose, and the ADP and CDP derivatives were inactive.

Unfortunately, chloroplasts have several limitations as a system for investigating the earlier steps in sulfolipid biosynthesis. First, intact chloroplasts are quite selective in their uptake of small molecules, making precursor and inhibitor studies quite difficult or impossible; second, it has not been possible to observe the biosynthesis of sulfolipid from labeled sulfate or PAPS in preparations of lysed chloroplasts (B. Mudd, personal communication). However, Hoppe and Schwenn (19) reported the incorporation of labeled sulfate, sulfite, and PAPS into sulfoquinovosyl diacylglycerol by cell-free extracts from *Chlamydomonas reinhardii*. Incorporation of labelled sulfate required ATP and magnesium, and was stimulated by the addition of glutathione, while addition of cysteic acid did not affect the incorporation. PAPS behaved as a legitimate substrate, since the reaction followed a saturation curve with increasing amounts of PAPS. On the other hand, the incorporation of sulfite into sulfolipid could not be saturated, even with very high amounts of sulfite, suggesting a non-enzymatic artifact.

In conclusion, the biosynthesis of sulfoquinovose diacylglycerol probably starts with activated sulfate, either APS or PAPS. It seems to be fairly clear that the last step involves the reaction between UDP-sulfoquinovose and diacylglycerol to yield the sulfolipid. Every other step in between (see Figure 1-2) is not sufficiently supported by experimental evidence to allow any

conclusions to be drawn. It will be very interesting to see at what level the sulfonate group is introduced into the molecule. It is entirely possible that the pathways outlined in Figure 1-2 are wrong.

PROPOSED FUNCTION OF THE PLANT SULFOLIPID

The exclusive occurrence of sulfoquinovosyl diacylglycerol in photosynthetic organisms and the unique structure of the sulfonic acid head group, suggests that this compound might play a critical role in photosynthesis (for reviews see ref. 1, 15, 27). This idea is supported by a study in which correlations were observed between the amounts of sulfolipid and the levels of chlorophyll in *Chlorella protothecoides* during cycles of bleaching and greening, induced by changes in the growth conditions (35). Cells grown in the light, in medium with high urea and low glucose content, were green and contained high amounts of chlorophyll and sulfolipid , while cells grown in medium rich in glucose and poor in urea were bleached, independent of the illumination, resulting in low chlorophyll and sulfolipid contents. Similarly, Sinensky was able to lower the levels of sulfolipid in *Chlorella pyrenoidosa* by growing the cells on cysteine as the sole sulfur source (37). The cells retained the ability to evolve oxygen, but were unable to fix carbon dioxide, and triacylglycerols accumulated to high levels. However, a specific function for the plant sulfolipid cannot be inferred from these experiments, because it is impossible to evaluate the importance of the non-specific side effects of the growth conditions. A correlation between the

amount of sulfolipid and photosynthetic activity was also observed in a study employing the different developmental stages of plastids along maize leaves. The levels of chlorophyll and sulfolipid were found to increase in parallel with increased maturation of the plastids (24). Correlations have also been observed in higher plants between increased levels of sulfolipid and cold hardiness (23). Unfortunately, such correlative approaches seem suggestive at best, since many cellular constituents may differ as a result of the contrasting treatments.

Reconstitution experiments employing protein complexes common to the thylakoid membrane have been used to examine the function of the sulfolipid. It has been found in very tight association with isolated photosystem II apoproteins (9, 36). Sulfolipid was also present in stoichiometric amounts in highly purified preparations of the ATPase CF_0-CF_1 complex (30). Removal of the tightly bound sulfolipid from the ATPase complex resulted in loss of activity. Enzymatic activity was restored by readdition of sulfoquinovosyl diacylglycerol and other chloroplast lipids in a fixed ratio. It was concluded that the sulfolipid is either required as boundary lipid, or that it might even participate directly in the enzymatic reaction (1).

Haines proposed a model in which the charged, anionic head groups of phospholipids, or the sulfated lipids of *Ochromonas*, could act as a proton conducting pathway in proton pumping membranes (12). This model implies that sulfolipid located in the inner leaflet of the thylakoid membrane might play an essential and direct role in photosynthetic electron transport. It has been

shown that antibodies, raised against sulfolipid, can inhibit photosynthetic electron transport (25). The inhibition was observed for a photosystem I mediated reaction between the electron donor 2,6-dichlorophenol indophenol and the electron acceptor methylviologen. However, the pH-dependence for the inhibitory reaction correlated with the pH-dependence for the interaction of the sulfolipid with polypeptides *in vitro*, and it was concluded that the antibodies indirectly affected the function of proteins of the electron transport chain, to which sulfolipid was tightly bound.

In summary, although several experiments suggest a role for the plant sulfolipid in photosynthesis, the evidence is indirect and not sufficient to propose a specific model for its function.

ABOUT THE PRESENTED WORK

Although sulfolipid is one of the most abundant sulfur-containing molecules in the biosphere, the biosynthetic pathway is not known. Because it has not been possible to identify intermediates in the pathway between activated sulfate in the form of APS or PAPS and sulfoquinovose diacylglycerol by conventional biochemical approaches, a genetic approach to this thirty-year-old problem has been applied. The first step of this approach is to isolate mutants deficient in sulfolipid biosynthesis in the purple non-sulfur bacterium *Rhodobacter sphaeroides*. Specific blocks in the pathway in different mutants provide the tools to dissect the sequence of reactions in the formation of the sulfolipid. Bacterial mutants allow the isolation of the respective genes by

complementation. Mutants may serve as a source of precursors, which build up due to a specific metabolic block in the pathway. In addition, these mutants may help to determine the function of the plant sulfolipid.

In Chapter 2 the identity of purified sulfolipid from spinach and *R. sphaeroides* will be tested by mass spectrometry, since the occurrence of authentic plant sulfolipid in *R. sphaeroides* is critical for the described work and its presumptive relevance to higher plants. In Chapter 3, the isolation of mutants deficient in plant sulfolipid as well as their basic characteristics will be described. Chapters 4 and 5 provide information about the genetic complementation of three sulfolipid mutants and a detailed analysis of two of the genes directly involved in the biosynthesis of the plant sulfolipid in *R. sphaeroides*. The versatility of this purple bacterium for the study of sulfolipid biosynthesis is further demonstrated in Chapter 6, where an *in vitro* system for the biosynthesis of sulfolipid in extracts of *R. sphaeroides* is described. Future prospects for using the genetic and biochemical tools provided in the presented work, to solve the questions about the biosynthesis and function of the plant sulfolipid, will be discussed in Chapter 7.

REFERENCES

1. Barber J., Gounaris K. (1986). What role does sulpholipid play within thylakoid membranes? Photosynthesis Res. 9:239-249

2. Benson A.A., Daniel H., Wiser R. (1959). A sulfolipid in plants. Proc. Natl. Acad. Sci. USA 45:1582-1587

3. Benson A.A. (1963). The plant sulfolipid. Adv. Lipid Res. 1:387-394

4. Bishop D.G., Sparace S.A., Mudd B.J. (1985). Biosynthesis of sulfoquinovosyldiacylglycerol in higher plants: The origin of the diacylglycerol moiety. Arch. Biochem. Biophys. 240:851-858

5. Davies W.H., Mercer E.I., Goodwin T.W. (1966). Some observations on the biosynthesis of the plant sulfolipid by *Euglena gracilis*. Biochem J. 98:369-373

6. Douce R., Holtz R.B., Benson A.A. (1973). Isolation and properties of the envelope of spinach chloroplast. J. Biol. Chem. 248:7215-7222

7. Galliard T. (1968). Aspects of lipid metabolism in higher plants-1: Identification and quantitative determination of the lipids in potato tubers. Phytochem. 7:1907-1914

8. Galliard T. (1968). Aspects of lipid metabolism in higher plants-2: The identification and quantitative analysis of lipids from the pulp of pre- and post-climacteric apples. Phytochem. 7:1915-1922

9. Gounaris K., Barber J. (1985) Isolation and characterization of a photosystem II reaction center lipoprotein complex. FEBS Letter 188:68-72

10. Haas R., Siebertz H.P., Wrage K., Heinz E. (1980). Localization of sulfolipid labeling within cells and chloroplast. Planta 148:238-244

11. Haines T.H. (1973). Sulfolipids and halosulfolipids. In Lipids and Biomembranes of Eukaryotic Microorganisms, Ed Erwin J.A., Academic Press, NY, pp 197-232

12. Haines T.H. (1982). Anionic lipid head groups as a proton conducting-pathway along the surface of membranes: a hypothesis. Proc. Natl. Acad. Sci. USA 80:160-164

13. Harwood J.L. (1975). Synthesis of sulfoquinovosyl diacylglycerol by higher plants. Biochim. Biophys. Acta 398:224-230

14. Harwood J.L., Nicholls R.G. (1979). The Plant sulfolipid - a major component of the sulphur cycle. Biochem. Soc. Trans. 7:440-447

15. Harwood J.L. (1980). Sulfolipids. In The Biochemistry of Plants, Vol 4, Ed Stumpf P.K., Academic Press, NY, pp 301-320

16. Heinz E., Schmidt H., Hoch M., Jung K.H, Binder H., Schmidt R.R. (1989). Synthesis of different nucleoside 5'-diphospho-sulfoquinovoses and their use for studies on sulfolipid biosynthesis in chloroplast. Eur. J. Biochem. 184:445-453

17. Hoch M., Heinz E., Schmidt R.R. (1989). Synthesis of 6-deoxy-6-sulfo-α-D-glucopyranosyl phosphate. Carbohyd. Res. 191:21-28

18. Hodson R.C., Schiff J.A., Mather J.P. (1971). Studies on the sulfate utilization by algae. Plant Physiol. 47:306-311

19. Hoppe W., Schwenn J.D. (1981). *In vitro* biosynthesis of the plant sulpholipid: on the origin of the sulphonate head group. Z. Naturforsch. 36c:820-826

20. Joyard J., Blee E., Douce R. (1986). Sulfolipid synthesis from $^{35}SO^{2-}_4$ and [1-^{14}C]acetate in isolated spinach chloroplast. Biochim. Biophys. Acta 879:78-87

21. Kleppinger-Sparace K.F., Mudd B.J., Bishop D.G. (1985). Biosynthesis of sulfoquinovosyl diacylglycerol in higher plants: the incorporation of $^{35}SO_4$ by intact chloroplast. Arch. Biochem. Biophys. 240:859-865

22. Kleppinger-Sparace K.F., Mudd J.B. (1987). Biosynthesis of sulfoquinovosyldiacylglycerol in higher plants. The incorporation of $^{35}SO_4$ by intact chloroplast in darkness. Plant Physiol. 84:682-687

23. Kuiper P.J.C. (1970). Lipids in *Alfalfa* leaves in relation to cold hardiness. Plant Physiol. 45:684-686

24. Leech R.M., Rumsby M.G., Thomson W.W. (1973). Plastid differentiation, acyl lipid, and fatty acid changes in developing green maize leaves. Plant Physiol. 52:240-245

25. Menke W., Radunz A., Schmid G.H., Koenig F., Hirtz R.D. (1976). Intermolecular interactions of polypeptides and lipids in the thylakoid membrane. Z. Naturforsch. 31c:436-444

26. Mudd J.B., Dezacks R., Smith J. (1980). Studies on the biosynthesis of sulfoquinovosyldiacylglycerol in higher plants. In Biogenesis and Function of Plant Lipids, Eds Mazliak P., Beneviste P., Costes C., Douce R., Elsevier, Amsterdam pp 57-66

27. Mudd B.J., Kleppinger-Sparace K.F. (1987). Sulfolipids. In The Biochemistry of Plants, Vol 9, Eds Stumpf P.K and Conn E.E., Academic Press, NY, pp 159-173

28. Ongun A., Thomson W.W., Mudd J.B. (1968). Lipid composition of chloroplast isolated by aqueous and nonaqueous techniques. J. Lipid Res. 9:409-415

29. Park C., Berger L.R., (1967). Complex lipids of *Rhodomicrobium vannielii.* J. Bacteriol. 93:221-229

30. Pick U., Weiss M., Gounaris K., Barber J. (1987). The role of different thylakoid glycolipids in the function of reconstituted chloroplast ATP synthase. Biochim. Biophys. Acta 891:28-39

31. Radunz A. (1969). Ueber das Sulfochinovosyl-diacylglycerin aus hoeheren Pflanzen, Algen und Purpurbakterien. Hoppe-Seylers Z. Physiol. Chem. 350:411-417

32. Radunz A., Berzborn R. (1970). Antibodies against sulfoquinovosyl-diacylglycerol and their reactions with chloroplast. Z. Naturforsch. 25b:412-419

33. Roughan G., Slack, R. (1984). Glycerolipid synthesis in leaves. Trends Biochem. Sci. 9:383-386

34. Shibuya I., Yagi T., Benson A.A. (1963). Sulfonic acids in Algae. In Microalgae and Photosynthetic Bacteria, Ed Jpn. Soc. Plant Physiol., Univ of Tokyo Press, Tokyo, pp 627-636

35. Shibuya I., Hase E. (1965). Degradation and formation of sulfolipid occurring concurrently with de- and re-generation of chloroplast in the cells of *Chlorella protothecoides* Plant Cell Physiol. 6: 267-283

36. Sigrist M., Zwillenberg C., Giroud C.H., Eichenberger W., Boschetti A. (1988). Sulfolipid associated with the light harvesting complex associated with photosystem II apoproteins of *Chlamydomonas reinhardii*). Plant Science 58:15-23

37. Sinensky M. (1977). Specific deficit in the synthesis of 6-sulfoquinovosyl diglyceride in *Chlorella pyrenoidosa.* J. Bacteriol. 129:516-524.

38. Stuiver C.E.E., Kuiper P.J.C., Marschner H. (1978). Lipids from barley and sugar beet in relation to salt resistance. Physiol. Plant. 42:124-128

39. Tevini M. (1977). Light, function and lipids during plastid development. In Lipids and Lipid Polymers in Higher Plants, Eds Tevini M. and Lichtenthaler H.K., Springer Verlag, Berlin pp 121-145

40. Wood B.J.B., Nichols B.W., James A.T. (1965). The lipids and fatty acid metabolism of photosynthetic bacteria. Biochim. Biophys. Acta 106:261-273

CHAPTER 2

COMPARISON OF SULFOLIPID FROM SPINACH AND *R. SPHAEROIDES* BY MASS SPECTROMETRY

ABSTRACT

To test the identity of sulfolipid isolated from the purple non-sulfur bacterium *R. sphaeroides* with the plant sulfolipid sulfoquinovosyl diacylglycerol, the structures of sulfolipid from spinach and *R. sphaeroides* were compared by negative mode FAB-MS and FAB-CAD-MS/MS tandem mass spectrometry. Utilizing these techniques, it was demonstrated that the sulfur containing head group from the plant and bacterial lipids produced identical fragmentation patterns in agreement with sulfoquinovose. However, it should be mentionend that the employed method did not allow to distinguish between various stereo isomeres. The overall differences in the samples were based on different fatty acid composition of the sulfolipid isolated from the two organisms. Fragmentation pathways were postulated, to interpret head group specific fragments found at m/z 299, 283, 241, 225, 161 and 80 in the two samples. The two predominant molecular species of spinach sulfolipid contained either palmitic and linolenic acids (m/z 815) or two molecules of linolenic acid (m/z 837), while the two predominant species of bacterial sulfolipid contained palmitic and vaccenic acids (m/z 819) or stearic and vaccenic acids (m/z 847).

INTRODUCTION

The plant sulfolipid was discovered by Benson *et al.* (1), who also provided the first structural data about it. By chromatography of the hydrolysis products of uniformly ^{14}C-labeled sulfolipid, it was established that the molecule contained glycerol, two fatty acids and a 6-carbon compound, which could also be labeled with ^{35}S-sulfate. The sulfur containing, 6-carbon compound was recognized as a sulfonic acid by its complete resistance to acid hydrolysis. The presence of hydroxyl groups, as well as its reaction with 2,4-dinitrophenyl- hydrazine suggested a aldohexose structure. By analogy to the galactolipids, and based on the fact that the glycosidic linkage was not cleaved by α-galactosidase from yeast, a structure was initially proposed, that featured a basic diacylglycerol structure, with a 6-sulfo-6-deoxy-galactose head group linked in a 1-ß-glycosidic linkage to the glycerol moiety (1). However, further experiments, which included elemental analysis, proton magnetic resonance spectrometry, infrared spectrometry, optical rotary dispersion spectrometry, and X-ray crystallography of the sulfolipid from *Chlorella* and alfalfa, required modifications of the originally proposed structure. The sulfolipid was thus correctly recognized as 6-sulfo-α-D-quinovosyl diacylglycerol (Figure 1-1). Benson summarized the data on which the final structural elucidation of the plant sulfolipid was based in a review (2).

The occurrence of the plant sulfolipid in the purple bacterium *R. sphaeroides* was originally demonstrated by Wood *et al.* (8), and later by Radunz (7). Both groups identified the sulfolipid by cochromatography with

standards. This finding offered the opportunity to study the biosynthesis of the plant sulfolipid in a bacterial system. However, since it was crucial to the study that the bacterial and the plant sulfolipid were both sulfoquinovosyl diacylglycerol, sulfolipids isolated from spinach and *R. sphaeroides* were compared by mass spectrometry to confirm their identity. This technique is very sensitive, requiring minimal amounts of substance (1 nmol, equivalent to 0.1-1.0 μg depending on the molecular weight) to be analyzed. Fast Atom Bombardment (FAB) employing a high energy (6 keV) Xenon atom beam is used to desorb and ionize nonvolatile analytes from a matrix. The ions produced are then accelerated in an electric field (10 kV), passed through an electric sector (E), which acts to narrow the energy spread of the accelerated ions, and then mass analyzed by their momentum in a magnetic field (B). To acquire a spectrum the magnetic field strength is scanned so that ions of different momenta are sequentially allowed to reach the detector. In the case of sulfolipids the instrument is run in the negative ion mode, since these compounds readily form anions. The FAB ionization method is very "soft", resulting in a relatively low amount of fragmentation and a high percentage of molecular ions ($[M-H]^-$), although sulfolipids do fragment to some extent. The negative mass spectrum of the sulfolipids displays prominent peaks, representing the different molecular species present in the sample.

For structural information other than molecular weight, diagnostic fragmentation patterns are observed in a FAB-CAD-MS/MS experiment, which is based on an extension of the technique described above. This allows

"daughter ions" of one particular "parent ion" to be selectively observed. To generate "daughter ions", a helium gas filled chamber (0.1 Torr) is placed in the first field-free region of the instrument resulting in the collision of the accelerated ions with helium gas atoms. Through "Collision Activated Dissociation" (CAD), "daughter ions" are generated. Fragments or "daughter ions" created in this process are formed with an energy that is proportional to their mass in a linear fashion. The value of the energy/mass ratio is dependent upon the mass of the "parent ion". To detect "daughter ions" of a particular "parent ion", the instrument is scanned at a constant ratio of B/E (magnetic field strength/electrical field strength in the electrical sector), typical for the respective parent ion, permitting the observation of only those "daughter ions" generated from that particular "parent ion" without interference from matrix related ions or ions from other compounds in the sample (for more detailed physical theory, see reference 4).

In this chapter, a diagnostic fragmentation pattern for sulfoquinovosyl diacylglycerol will be postulated, and the identity of the plant and bacterial sulfolipid will be confirmed by the MS/MS techniques described above.

MATERIALS AND METHODS

Plant and bacterial material. Spinach leaves were obtained from a local grocery store. The wild type *R. sphaeroides* strain 2.4.1 used, was kindly provided by Dr. Uffen's lab at Michigan State University. Bacterial 4-I cultures were grown photoheterotrophically in Sistrom's medium, as will be described

in detail in the methods section of chapter 3.

Large scale preparation of sulfolipid. Spinach leaves (500 g) were blended in 1.2 l of chloroform-methanol-formic acid (10:10:1, v/v). Solids were separated by filtration and reextracted with 440 ml chloroform-methanol-water (5:5:1, v/v). The two combined filtrates were extracted with 600 ml 1 M KCl, 0.2 M phosphoric acid. The dried chloroform phase yielded about 4.5 g of lipids. This material was purified by column chromatography according to O'Brien and Benson (6). Similarly, sulfolipid was extracted and purified from 20 g of *R. sphaeroides* (2.4.1) cells, except that volumes were cut in half. As a final purification step, TLC on ammonium sulfate plates was employed to remove contaminating phosphatidyl glycerol. (The TLC system and its characterization will be described in detail in chapter 3). The sulfolipid band was scraped off the plate after staining with iodine vapor, and the material was eluted using chloroform-methanol (1:1, v/v). In addition, this TLC system was used to monitor the purity of the sulfolipid throughout the purification procedure.

Mass spectrometry. Approximately 10 μg of the sulfolipid sample dissolved in chloroform/methanol (1:1, v/v) were mixed with 2 μl of triethanolamine matrix on the FAB probe tip. Ions were produced by bombardment with a beam of Xe atoms (6 keV) or Cs^+ ions (8 keV) in a JEOL HX-110 double focusing mass spectrometer (EB) configuration operating in the negative ion mode; no differences were observed in spectra acquired by neutral beam (Xe^o) or charged cesium beam (Cs^+) bombardment. The

accelerating voltage was 10 kV and the resolution was set at 1000. For FAB-CAD-MS/MS, helium was used as the collision gas in a cell located in the first field-free region. The helium pressure was adjusted to reduce the abundance of the parent ion by 50%. A JEOL DA-5000 data system generated linked scans at a constant ratio of magnetic to electrical fields (B/E). The instrument was scanned in 2 minutes from m/z 0-1500. Data presented were acquired in a single scan and were found to be reproducible.

RESULTS AND DISCUSSION

Comparison of molecular species of sulfolipid found in samples of spinach and *R. sphaeroides*. Negative FAB mass scans made it possible to detect the most abundant molecular species contained in the two sulfolipid samples obtained from spinach and *R. sphaeroides*. Two predominant molecular species were found in the spinach sample at m/z 815 and 837, which corresponded to sulfoquinovosyl diacylglycerol ($[M-H]^-$) containing palmitic and linolenic acids, or two linolenic acids respectively. This result was in agreement with data obtained by conventional analysis of the fatty acid composition of spinach sulfolipid (3). Loss of one fatty acid in the form of a ketene (R=C=O) resulted in ions of m/z 555 and 577, depending on which fatty acid was lost from which molecular ion (Figure 2-1, top). The equivalent negative FAB mass scan for bacterial sulfolipid showed two predominant species of sulfoquinovosyl diacylglycerol at m/z 847 and 819 (Figure 2-1, bottom), which corresponded to two molecular species containing either

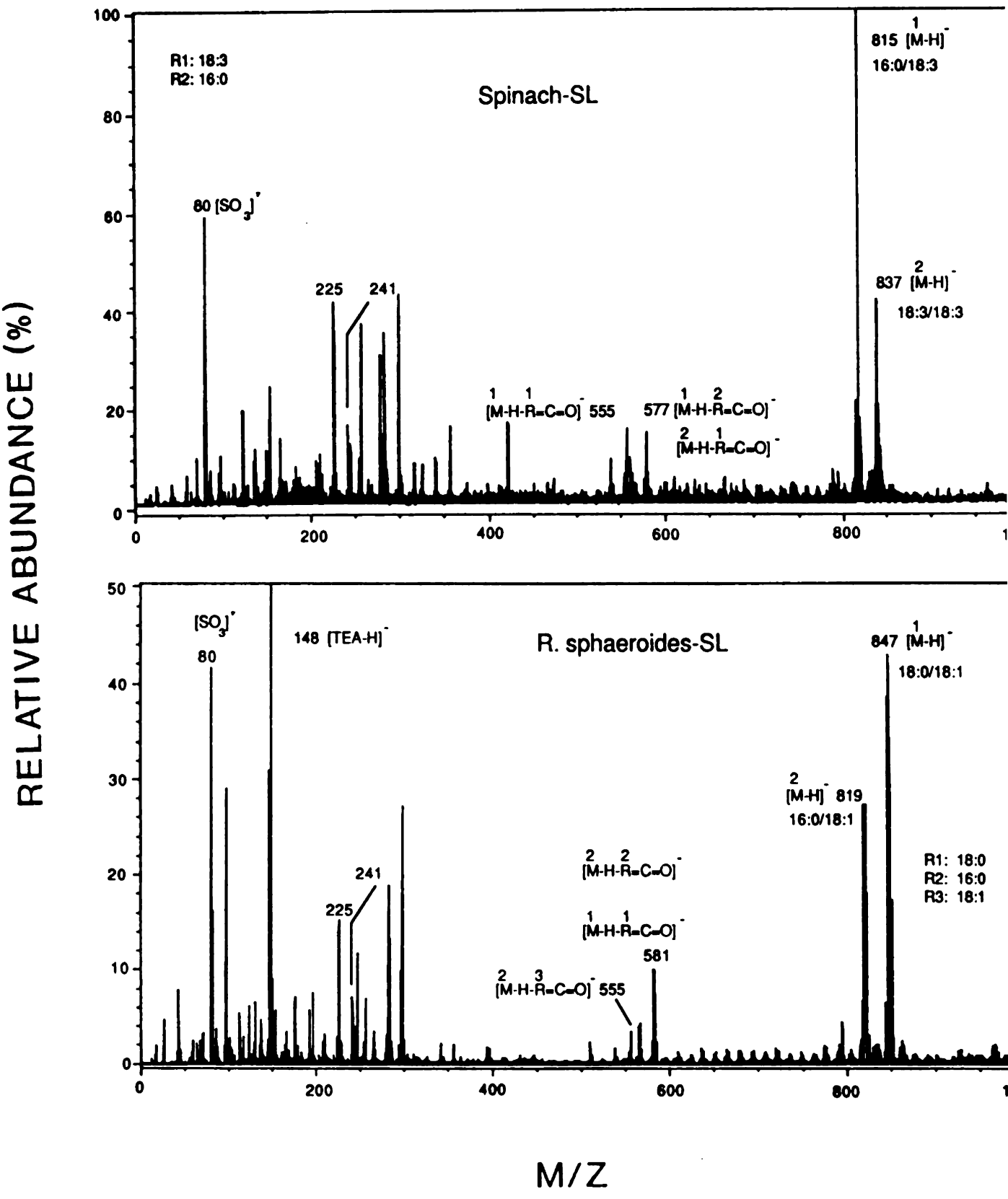


Figure 2-1: Negative FAB mass spectra of sulfolipid samples from spinach (top) and *R. sphaeroides* (bottom). R, residues characteristic for various fa acids. $[M-H]^-$, molecular ion following the loss of a proton. $[TEA-H]^-$, trietha amine matrix peak. The mass of characteristic fragments is indicated.

stearic and palmitic acids, or palmitic and vaccenic acids respectively. The most unsaturated fatty acid commonly found in bacteria was vaccenic acid, a *cis*-Δ11 mono-unsaturated 18-carbon fatty acid. The observed fatty acid composition was in agreement with data obtained by Radunz (7). Loss of one fatty acid moiety (as the ketene analog) resulted in ions of m/z 581 and 555, depending on the fatty acid and the molecular species. Both spectra from spinach and bacterial sulfolipid showed peaks at m/z 241, 225, and 80, which were suspected to be characteristic for the head group.

B/E linked scans of molecular species of sulfolipid. Two B/E linked scans focusing on the major species found in the spinach sample (m/z 815) and the bacterial sample (m/z 847) are shown in Figure 2-2. Due to collision activated dissociation as described above, there was a high abundance of fragments at m/z 600-800 at intervals of about 14 mass units. These corresponded to fragmentation along the fatty acid chains, generating fragments with increments of one methylene (CH_2). In the low mass range fragments characteristic for the lipid head group were found. Both spectra showed a comparable pattern of fragments with peaks at m/z 283, 241, 225, and 80. (It must be emphasized that the relative abundance of peaks in mass spectra generated by the use of FAB and CAD is not always reproducible. The relative abundance of "daughter ions" seems to be influenced by sample concentration and purity.) Similar, low mass daughter ions were detected when B/E linked scans were performed on the parent ions of m/z 555 for spinach and of m/z 581 for the bacterial sample (Figure 2-3), which represent

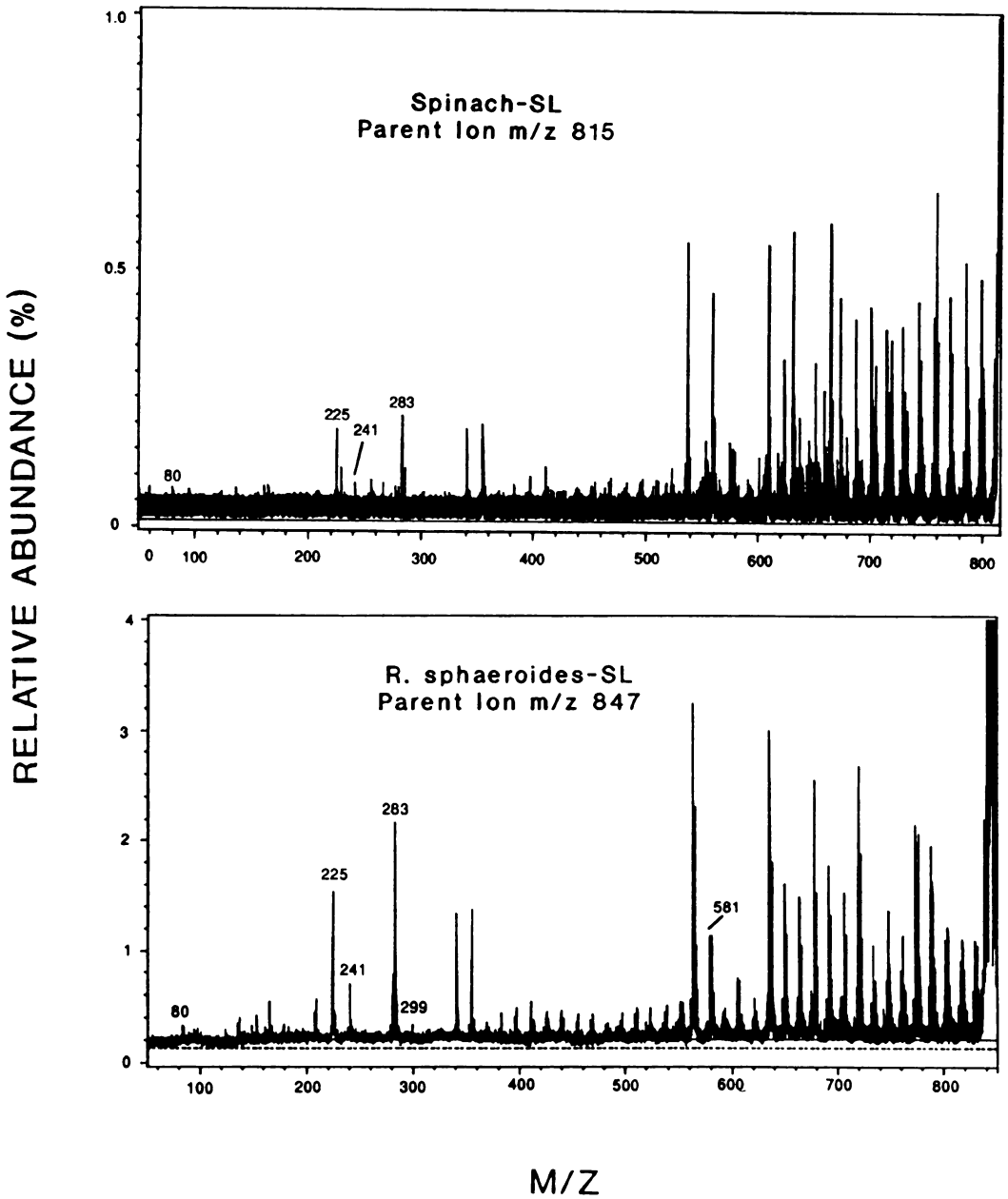


Figure 2-2: B/E linked FAB-CAD-MS/MS scans focusing on the most abundant molecular ion of the spinach (top) and the bacterial sample (bottom). The mass of characteristic head group fragments is indicated. In the high mass region the abundant peaks are due to fragments formed from cleavage along the fatty acyl chains.

the respective molecular species following the loss of one fatty acid (compare Figure 2-1).

B/E linked scans of sulfolipid head group specific fragments. The sulfoquinovose sugar moiety has a molecular weight of 243. Consistently in all the scans discussed, an ion was observed at m/z 241 and seemed to be a key ion. It was proposed that this ion represents a stable oxidation product of the ion at m/z 243, which has lost two hydrogens. The same phenomenon has been observed for sulfated sugar head groups from animal lipids (5). B/E linked scans focusing on the ion at m/z 241 from the spinach and the bacterial samples again showed a comparable pattern (Figure 2-4). The "daughter ion" at m/z 225 was proposed to be the epoxy derivative of the sulfoquinovose following the loss of water, while the ion at m/z 80 corresponded to the sulfonic acid group SO_3^- (Figure 2-5). The most abundant ion at m/z 161 resulted from the loss of the sulfonic acid group. Based on the limited available information, it was not attempted to propose a precise structure for this fragment.

B/E linked scans focusing on the 299 m/z key ion and its derivatives. The B/E linked scan for the"parent ion" at m/z 299 from the spinach sample showed fragments at m/z 225, 165, and 80, but non at m/z 241 (Figure 2-6, top). This suggested that the m/z 241 key ion was preferentially derived from the molecular ion, but that its formation from subfragments was not favored. Therefore, an alternative fragmentation pathway was proposed, which started with the formation of the key ion at m/z 299 (Figure 2-8). The ion at m/z 225

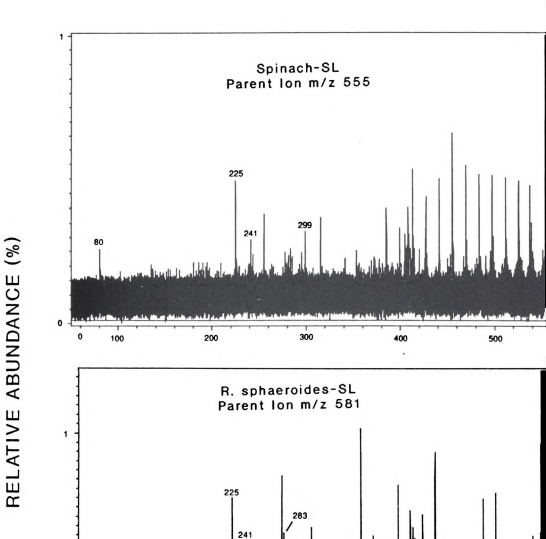


Figure 2-3: B/E linked FAB-CAD-MS/MS scans focusing on monoacyl spe
of sulfolipid from spinach (top) and from *R. sphaeroides* (bottom). The ma
of important fragments is indicated.

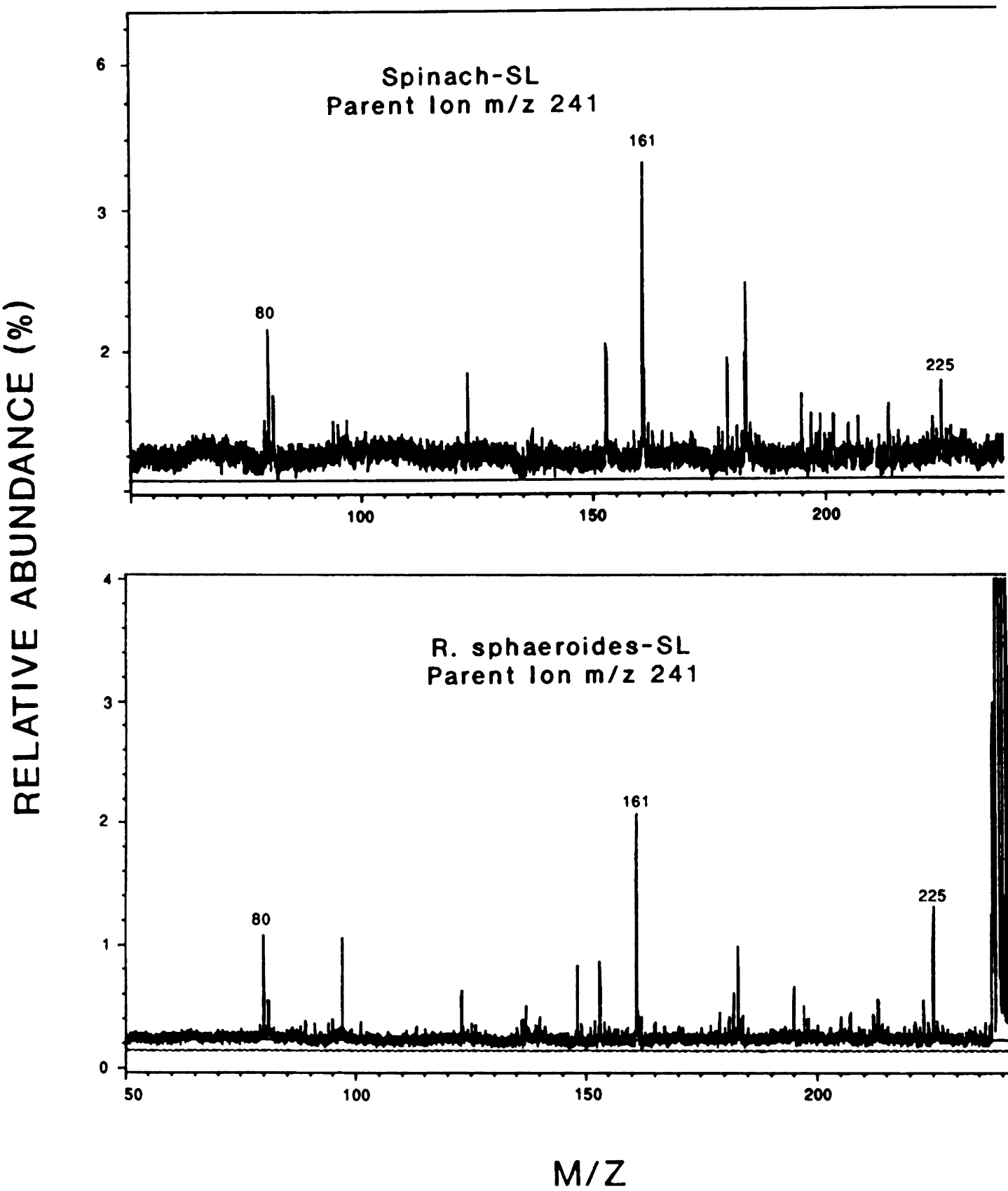


Figure 2-4: B/E linked FAB-CAD-MS/MS scans focusing on the head grou
key ion at m/z 241 from spinach (top) and bacterium (bottom). The mass
important fragments is indicated.

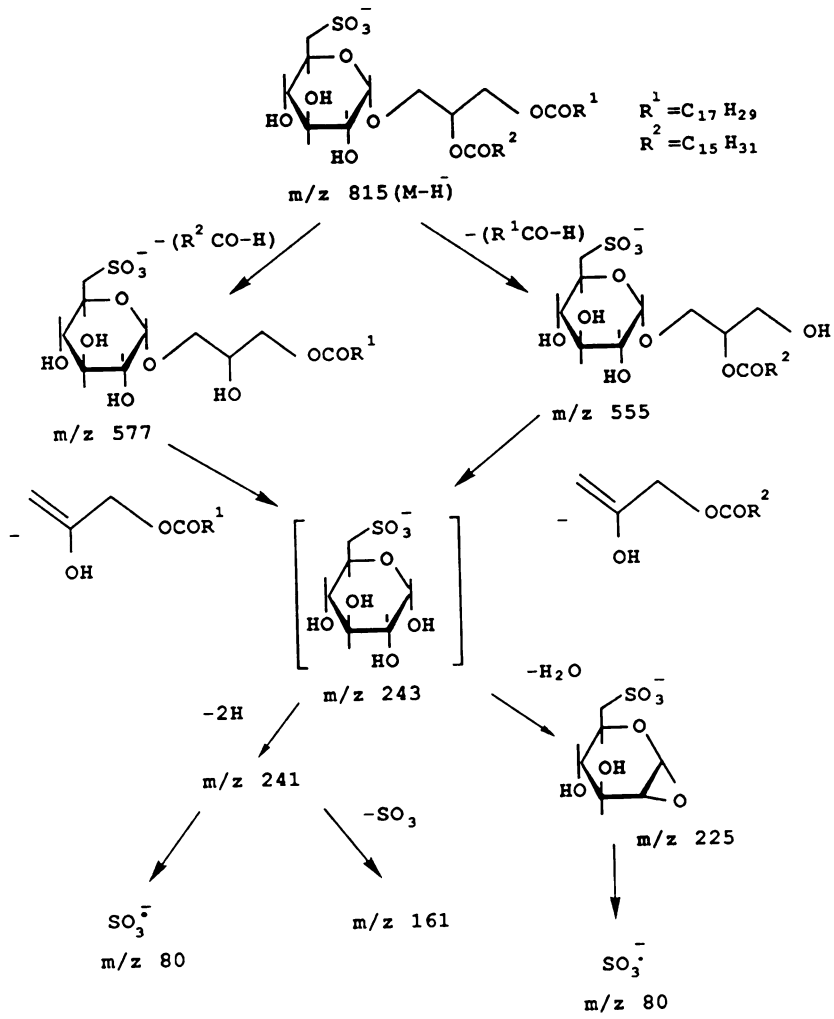


Figure 2-5: Proposed fragmentation pathway showing the decomposition of the key ion at m/z 241.

was interpreted as epoxy sulfoquinovose, while the SO_3^- group was represented by a peak at m/z 80. The structure of the ion at m/z 165 could not be proposed on the basis of the available data. This ion was not observed in the B/E linked scan focusing on the ion at m/z 241 (Figure 2-4). The B/E linked scan for the ion m/z 283 was very similar to the B/E linked scan of the ion at m/z 299 (Figure 2-6). The proposed structure for this ion is shown in Figure 2-8. The Carboxylate ion of stearic acid $[RCOO]^-$ has a mass of 283. Since there were considerable amounts of stearic acid in the bacterial sample, B/E linked scans focusing on the m/z 283 ion of the bacterial sample could not be interpreted.

To complete the series of B/E linked scans the fragmentation patterns for the ions m/z 225 and m/z 80 from the spinach sample were obtained (Figure 2-7). The m/z 225 fragment, as part of the sulfoquinovose head group, produced the SO_3^- fragment at m/z 80. In addition, several other fragments, including that at m/z 165, were generated, for which no structure could be proposed. The ion at m/z 80 gave rise to the SO_2 ion with a mass of 64, confirming that this ion represents the sulfonate group (Figure 2-7, bottom). B/E linked scans for the ions at m/z 299, 283, 225, and 80 for the bacterial sample were not run and could therefore not be compared with the equivalent ions of the spinach sample.

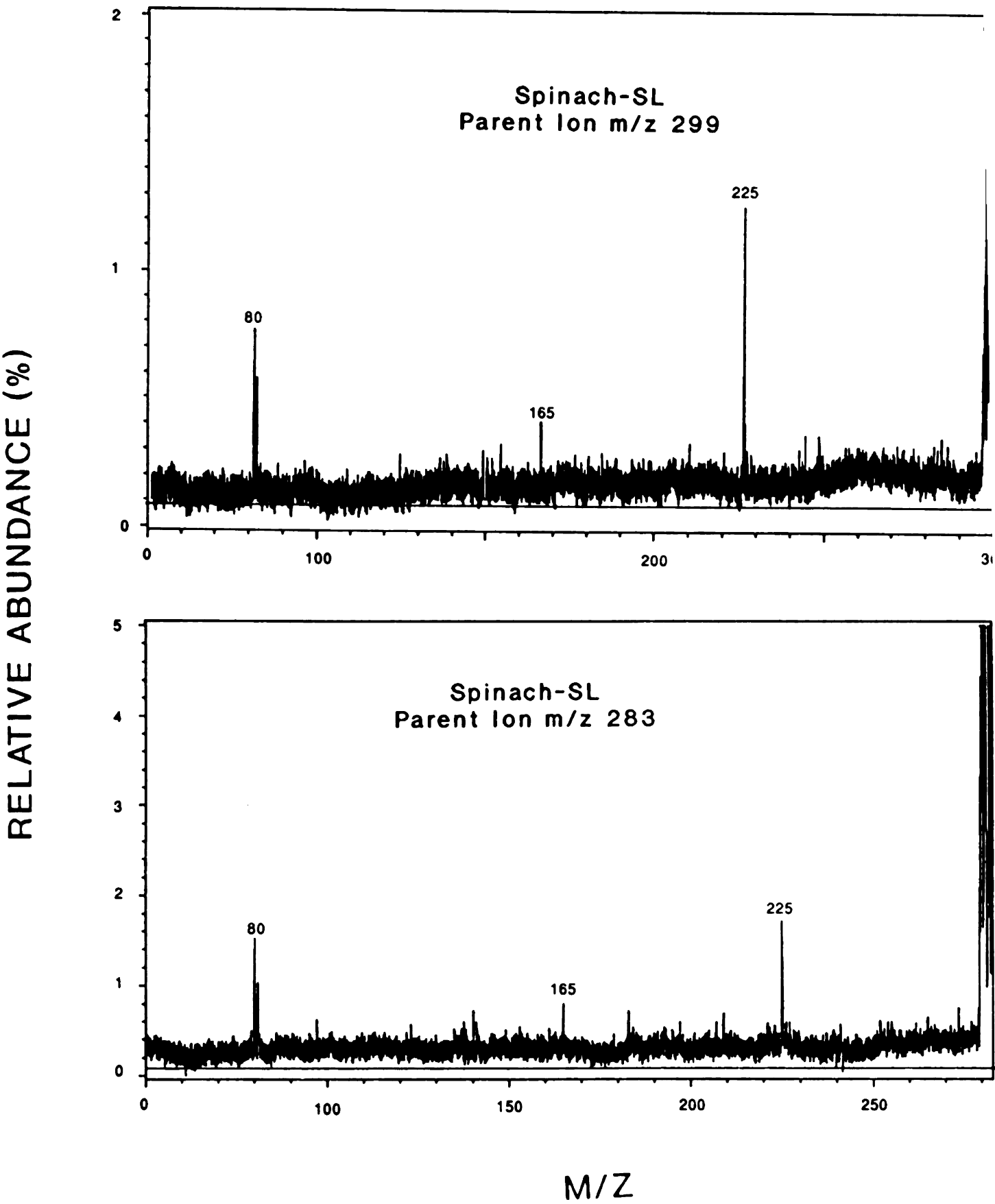


Figure 2-6: B/E linked FAB-CAD-MS/MS scans of the spinach sulfolipid, focusing on the "daughter ions" with the mass 299 (top), and 283 (bottom) The mass of important ions is indicated.

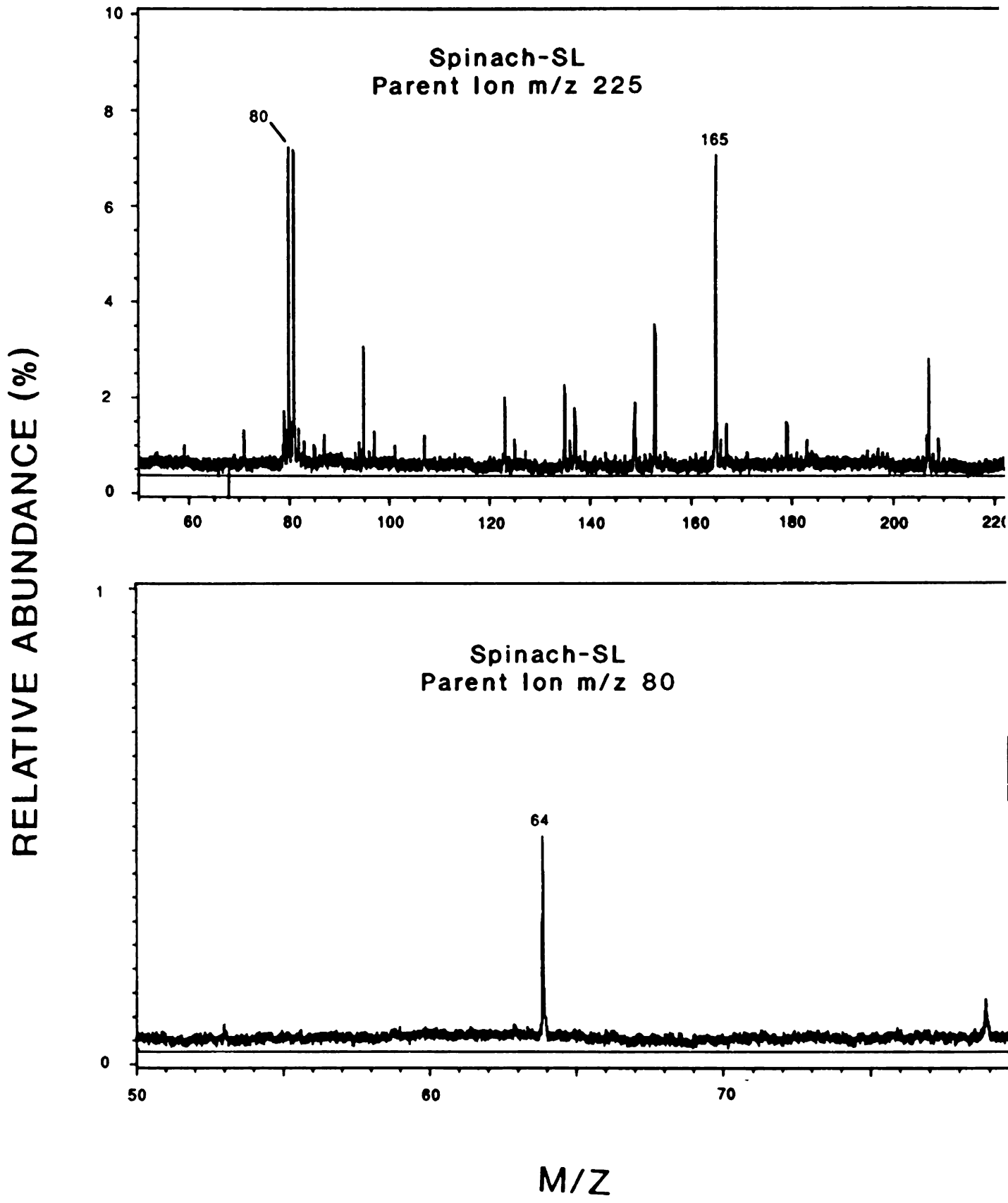


Figure 2-7: B/E linked scans of the spinach sulfolipid, focusing on the "daughter ions" with the mass 225 and 80. The mass of important ions is i

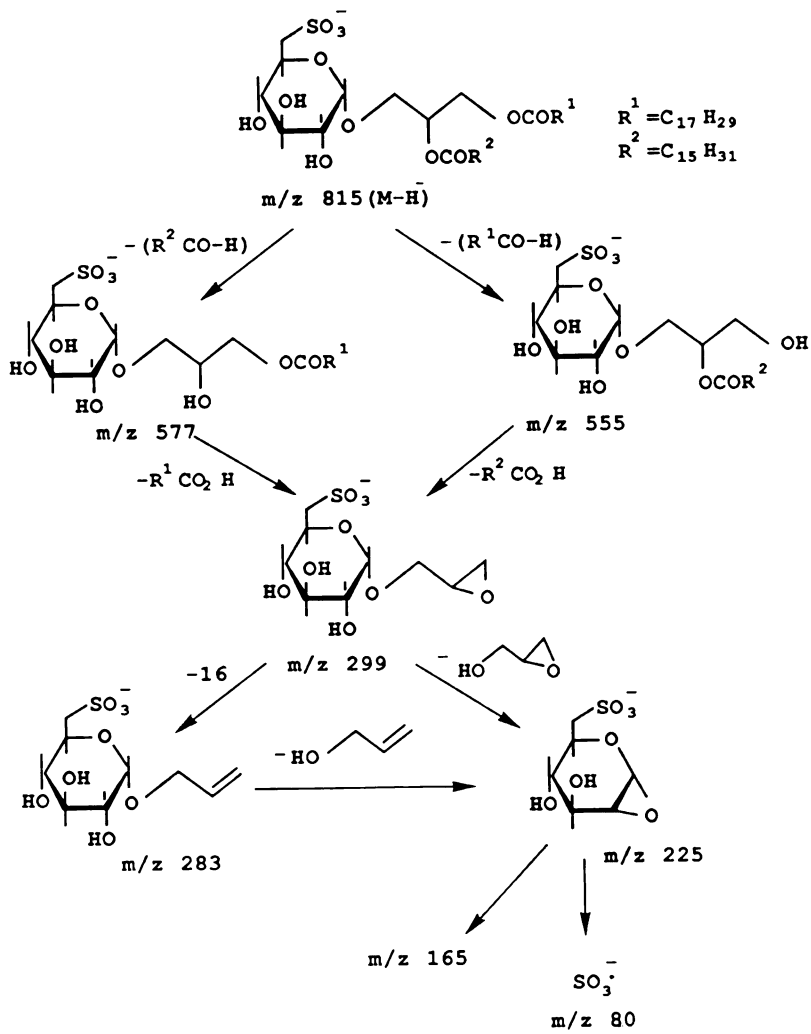


Figure 2-8: Alternative fragmentation pathway showing the decomposition of the key ion at m/z 299.

CONCLUSIONS

Based on the observed data, two diagnostic "daughter ions" of sulfolipid were postulated: first, the sulfoquinovose head group specific ion with a mass of 241, and second, the sulfoquinovose glycerol specific deacylation product with a mass of 299. Since the "daughter ion" at m/z 299 did not give rise to the ion at m/z 241, and since the B/E linked scans focusing on these two ions revealed distinct fragmentation patterns, it was postulated that both ions were produced in simultaneously occurring reactions from the molecular ion, and that further fragmentation occurred in a pattern specific for each of the two "daughter ions". Two alternative fragmentation pathways were postulated featuring the m/z 241 and the m/z 299 ions as key ions (Figures 2-5 and 2-8).

Comparison of linked B/E scans of the high mass ions and in particular of the head group specific ion at m/z 241, revealed similar fragmentation patterns for the spinach and bacterial sulfolipid, strongly suggesting that the sulfolipid found in both organisms is sulfoquinovosyl diacylglycerol. The data were supported and extended earlier findings obtained by Wood *et al.* (8) and Radunz (7), who identified the sulfolipid in *R. sphaeroides* by cochromatography with standards. However, it should be pointed out that the mass-spectrometry technique use did not allow to distinguish between the sulfoquinovose sugar and its stereo isomers. Although the data were consistent with the presence of the sulfoqinovose in the bacterial lipid, it could not be excluded that the sugar head group is a stereo isomere of sulfoquinovose.

REFERENCES

1. Benson A.A., Daniel H., Wiser R. (1959). A sulfolipid in plants. Proc. Natl. Acad. Sci. USA 45:1582-1587

2. Benson A.A. (1963). The plant sulfolipid. Adv. Lipid Res. 1:387-394

3. Bishop D.G., Sparace S.A., Mudd B.J. (1985). Biosynthesis of sulfoquinovosyldiacylglycerol in higher plants: The origin of the diacylglycerol moiety. Arch. Biochem. Biophys. 240:851-858

4. Jennings K. R., Mason R. S. (1983). Tandem mass spectrometry utilizing linked scanning of double focusing instruments. In: Tandem mass spectrometry; Ed. McLafferty; John Wiley & Sons, New York pp. 197-222

5. Kushi Y., Handa S., Ishizuka I. (1985). Secondary ion mass spectrometry for sulfoglycolipids: Application of negative ion detection. J. Biochem. 97:419-428

6. O'Brien J.S., Benson A.A. (1964). Isolation and fatty acid composition of the plant sulfolipid and galactolipids J. Lipid Res. 5:432-436

7. Radunz A. (1969). Ueber das Sulfochinovosyl-diacylglycerin aus hoeheren Pflanzen, Algen und Purpurbakterien. Hoppe-Seylers Z. Physiol. Chem. 350:411-417

8. Wood B.J.B., Nichols B.W., James A.T. (1965). The lipids and fatty acid metabolism of photosynthetic bacteria. Biochim. Biophys. Acta 106:261-273

CHAPTER 3

R. SPHAEROIDES AS A MODEL SYSTEM FOR THE BIOSYNTHESIS OF THE PLANT SULFOLIPID

ABSTRACT

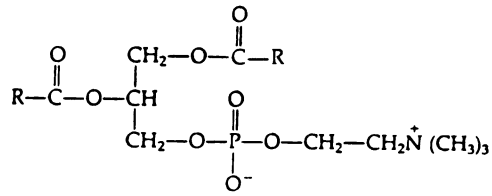
A TLC screening method for the isolation of chemically induced mutants of *R. sphaeroides* deficient in sulfolipid biosynthesis has been developed. Initially, 15 putative mutants were isolated from 1512 mutagenized cell lines examined. Following the preliminary analysis of all 15 putative mutants, further analysis was confined to three strains. Two of these, CHB16 and CHB17, were found to contain about 10% of the amount of sulfolipid as compared to the wild type, while in the mutant strain CHB18 the sulfolipid levels were reduced by about 50% with respect to the wild type. All mutants showed greatly reduced rates of incorporation of labelled sulfate into sulfolipid, while proteins were labelled at a normal rate compared to the wild type, indicating that the general sulfate assimilation pathway was not affected. Mutants CHB17 and CHB18 accumulated ^{35}S-labelled water-soluble compounds, which are potential intermediates in the pathway of sulfolipid biosynthesis.

INTRODUCTION

The four major polar lipids found in *R. sphaeroides* are phosphatidyl glycerol, phosphatidyl ethanolamine, phosphatidyl choline, and sulfoquinovosyl diacylglycerol (17). In addition, the presence of cardiolipin (17) as well as the occurrence of two unusual lipids, the ornithine lipid (4) and N-acylphosphatidylserine (2) has been reported. The structures of the *R. sphaeroides* lipids are given in Figure 3-1.

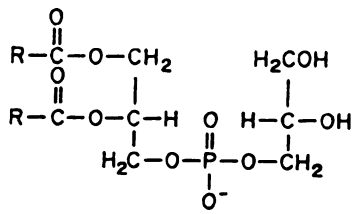
The occurrence of the plant sulfolipid in the purple bacterium *R. sphaeroides* was originally demonstrated by Wood *et al.* (17) and was later confirmed by Radunz (11). Both groups identified the sulfolipid by cochromatography with standards. The structural identity of the plant and the bacterial sulfolipid was further confirmed by mass spectrometry, as was described in the previous chapter. This finding offered the opportunity to study the biosynthesis of sulfoquinovosyl diacylglycerol in a bacterial system, which has several advantages as a genetic model system as compared with a higher plant system such as *Arabidopsis thaliana*. There are two main reasons for the choice of *R. sphaeroides* as a model system for the study of sulfolipid biosynthesis.

First, there is circumstantial evidence that the plant sulfolipid, which is predominantly found in association with photosynthetic membranes, might play an essential role in photosynthesis. The chemical properties of sulfoquinovosyl diacylglycerol, such as its low pK, make this lipid unique amongst the other



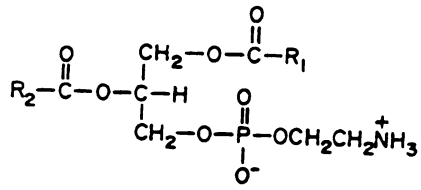
PC

(Phosphatidyl Choline)



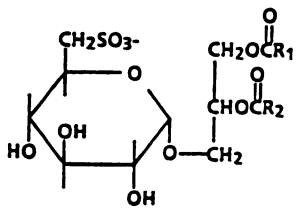
PG

(Phosphatidyl Glycerol)



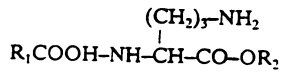
PE

(Phosphatidyl Ethanolamine)



SL

(6-Sulphoquinovosyl Diacylglycerol)



OL

(Ornithine Lipid)

Figure 3-1: Structures of the most abundant polar lipids found in *R. sphaeroides.*

membrane lipids in the chloroplasts of higher plants. In addition, the sulfolipid is found in tight association with intrinsic protein complexes of the thylakoid membrane, leading to the suggestion that it might be required for the catalytic activity of these protein complexes (1; see also discussion about the function of sulfolipid in chapter 1). There are sufficient reasons to assume that the loss of sulfolipid by an obligate photoautotrophic organism such as a higher plant may not be tolerated.

Purple non-sulfur bacteria, including *R. sphaeroides*, can grow under a wide range of growth conditions. The photosynthetic activity of the membranes is dispensable when the bacterium is grown chemoheterotrophically under aerobic conditions. The bacterium ceases production of the photosynthetic apparatus in certain growth regimes (e.g. in a high oxygen atmosphere in the presence of organic substrate). If the atmospheric pressure of oxygen is lowered below a certain point, *de novo* synthesis of the photosynthetic membranes is induced. *R. sphaeroides* is therefore ideally suited for the dissection of photosynthetic membrane biogenesis by classical and molecular genetic studies (8). Thus, if the plant sulfolipid is required for photosynthetic growth, it will still be possible to identify mutations that inactivate sulfolipid biosynthesis under conditions in which the bacterium can grow chemoheterotrophically.

A second reason for the use of *R. sphaeroides* is that all the genetic techniques required for modern genetic analysis of the pathway for sulfolipid biosynthesis have previously been applied to *R. sphaeroides*. Mutants

affecting complex biochemical pathways have been generated by chemical mutagenesis (e.g. reference 10); genes have been cloned by complementation of mutants using cosmid libraries, prepared from wild type DNA (e.g. reference 16); and gene replacement techniques, commonly used in other bacteria, have been used to construct specific mutations (e.g. reference 3). In addition, a physical map of the whole genome of *R. sphaeroides* has been recently constructed (15). This map might become useful for the linkage analysis of the various genes involved in sulfolipid biosynthesis.

Because of the genetic tools and the rapid life cycle, it is much more straightforward to isolate a gene by genetic complementation in the bacterial system, than in a higher plant.

It is expected that the information gained by studying sulfolipid biosynthesis in *R. sphaeroides* will be applicable to higher plants. Any hypothesis about sulfolipid biosynthesis in *R. sphaeroides* can be tested in higher plants using the tools developed during the course of this work. However, there is the possibility that the biosynthesis of sulfolipid proceeds along different routes in the bacteria compared to higher plants.

In this chapter, the feasibility of using mutants of *R. sphaeroides* to study the biosynthesis of the plant sulfolipid is reported. A TLC screening procedure for the isolation of polar lipid mutants and the primary phenotypes of three mutants deficient in sulfoquinovosyl diacylglycerol are described.

MATERIALS AND METHODS

Bacterial strain. The wild type strain *Rhodobacter sphaeroides* 2.4.1. used throughout this study was kindly provided by Dr. Uffen's lab at MSU.

Growth conditions for *R. sphaeroides*. Cell cultures were initially grown in the malate-basal-salt medium described by Ormerod et al. (9). The ingredients of this medium are listed in Table 3-1. For aerobic growth, stock solutions (I, II, III, IV, V, VI) were diluted into water according to the dilution factor given in Table 3-1. The pH was adjusted to 7.0 with NaOH following addition of 3 g D,L-malate per liter. Sterile solutions V and VI were added after autoclaving. For photoheterotrophic growth, sterile solutions VII and VIII were also added. In labelling experiments using ^{32}P-phosphate, solution I was replaced by solution I-P, and in labelling experiments involving ^{35}S-sulfate solutions I and II were replaced by I-S and II-S respectively.

In later experiments, Ormerod's medium was replaced by Sistrom's succinate-basal-salt medium A (13, 14), which was specifically adapted to the growth requirements of *R. sphaeroides*. The stock solutions required for Sistrom's medium are listed in Tables 3-2, 3-3, and 3-4. The 10X stocks were kept at -20°C. Following dilution of the stock solution, the pH was adjusted to 7.0 with KOH and the solution was autoclaved. Agar plates (1.5% agar) were either incubated in a dark incubator at 30°C or in a plexiglass box, which was flushed at a flow rate of 1 ml/min with a gas mixture containing 5% CO_2 and 95% N_2. Light and heating were provided by a 60 Watt incandescent light bulb placed in such a distant from the box to sustain a temperature of about

Table 3-1: Stock solutions for Ormerod's medium.

Stock#	(x)	Ingredients	g/l
I	(10):	KH_2PO_4	0.6
		K_2HPO_4	0.9
		$(NH_4)_2SO_4$	1.0
I-P	(10):	KCl	1.0
		KH_2PO_4	0.06
		K_2HPO_4	0.09
		$(NH_4)_2SO_4$	1.0
I-S	(10):	KH_2PO_4	0.6
		K_2HPO_4	0.9
		NH_4Cl	0.75
II	(10):	$MgSO_4{\cdot}7H_2O$	0.2
II-S	(10):	$MgCl_2{\cdot}6H_2O$	0.14
		$MgSO_4{\cdot}7H_2O$	0.02
III	(10):	Fe-EDTA	0.03
IV	(1000):	H_3BO_3	4.3
		$MnCl_2$	2.7
		$CuSO_4$	0.125
		$ZnSO_4$	0.288
		$(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$	0.037
		NaCl	0.584
		$CoCl_2$	0.0024
V	(1000):	Biotin	0.15
		Thiamine	1.0
		Nicotinic acid	1.0
VI	(100):	$CaCl_2$	8.0
VII	(50):	$NaHCO_3$	84.0
VIII	(100):	Ascorbic acid	50.0

Table 3-2: Ingredients of Sistrom's 10X medium.

Components	g/l
1. K_2HPO_4	34.8
2. $(NH_4)_2SO_4$ $(NH_4Cl)^*$	5.0 (1.95)*
3. Succinic acid	40.0
4. Glutamic acid	1.0
5. Aspartic acid	0.4
6. NaCl	5.0
7. Nitrilo acetic acid	2.0
8. $MgSO_4 \cdot 7H_2O$ $(MgCl_2 \cdot 6H_2O)^*$	3.0 (2.44)*
9. $CaCl_2 \cdot 2H_2O$	0.334
10. $FeSO_4 \cdot 7H_2O$	0.02
11. $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	0.002
12. Trace element solution (ml)	(1.0)
13. Vitamin solution (ml)	(1.0)

(* chloride salts to replace sulfate salts for ^{35}S-sulfate label experiments)

Table 3-3: Trace element solution for Sistrom's medium.

Components	g/100 ml
$Na_2EDTA \cdot 2H_2O$	1.41
$ZnSO_4 \cdot 7H_2O$	10.95
$FeSO_4 \cdot 7H_2O$	5.0
$MnSO_4 \cdot H_2O$	1.54
$CuSO_4 \cdot 5H_2O$	0.392
$Co(NO_3)_2 \cdot 6H_2O$	0.248
H_3BO_3	0.114

Table 3-4: Vitamin solution for Sistrom's medium.

Component	g/100 ml
Nicotinic acid	1.0
Thiamine·HCl	0.5
Biotin	0.010

30-35°C and a light irradiance of approximately 100 μE/s per m^2 at the level of the plates. Liquid cultures were inoculated with a single colony grown on an agar plate containing Sistrom's medium. Aerobic chemoheterotrophic, liquid cultures were incubated at 30°C under shaking in Erlenmeyer flasks closed with a foam stopper. Cells grown in these aerobic shaker cultures had a tendency to clump. Anaerobic photoheterotrophic cultures were grown in tightly closed, filled 200 ml bottles, which were kept under a 60 Watt incandescent lamp. This lamp was placed at such a distance as to provide a temperature of 30°C inside the bottle and a light irradiance of approximately 100 μE/s per m^2 at the surface of the bottle. The bottles were mixed once or twice a day by shaking. The generation time for photoheterotrophic cultures was about 4 h. For the large scale preparations, photoheterotrophic cultures were grown in 4-l Erlenmeyer flasks, with two 60 Watt bulbs placed on the outside (approximately 100 μE/s per m^2 at the surface of the flask). A 200-ml preculture was used for inoculation. Adequate mixing was provided by placing a magnetic stir bar inside the flask and the flask was placed in a water bath

kept at 30°C.

For growth of mutagenized cells on complete medium, Z-broth (14) was used. One liter of aqueous solution contained 10 g tryptone, 5 g yeast extract, 10 g NaCl, 1 g glucose, 0.4 g $CaCl_2{\cdot}2H_2O$ (pH 6.8-6.9). Plates, containing solidified medium were incubated in a dark incubator at about 35°C, liquid cultures were grown in a dark shaker at 35°C.

Mutagenesis of *R. sphaeroides*. A 50 ml culture (Z-broth) was grown from a single colony to a density of OD_{720}=0.5 (mid-log phase). To 20 ml of this culture 2 mg of nitrosoguanidine crystals were added (0.1 mg/ml), and the culture was incubated at room temperature in the fume hood with vigorous shaking. At various times, aliquots of 3 ml were removed, and the cells were immediately spun down and the supernatant was removed. The cells were washed once with 3 ml of Z-broth. Small aliquots were plated directly onto agar-solidified Z-broth to establish a killing curve, while the bulk of the culture was incubated for 17 h with shaking at 35°C to allow the cells to replicate two to three times. Aliquots of these cultures were plated onto Z-broth plates containing 100 μg/ml rifampicin or 100 μg/ml streptomycin to test for the degree of mutagenesis, by scoring the frequency of antibiotic resistant mutants. Further aliquots were plated onto plain Z-broth plates to determine the total amount of cells in the culture, and to initiate the screening for mutants. To the remainder of the cultures DMSO was added to a final concentration of 20% (v/v), to allow storage of the mutagenized cells at -70°C.

Screening for polar lipid mutants. Randomly chosen colonies from a mutated population of *R. sphaeroides* cells were picked to patches on a fresh Z-broth plate (24 per plate). Lipids were isolated from these patches by taking cells onto the wide end of a flat tooth pick and swirling the material in 75 µl chloroform/methanol (1:1, v/v) contained in polypropylene microfuge tubes. Following the addition of 25 µl 1 N KCl, 0.2 M H_3PO_4 the tubes were vortexed and centrifuged to separate the organic and the aqueous phases. Using a glass capillary micropipet, 10 µl from the lipid-containing lower phase were withdrawn and directly spotted onto an activated ammonium sulfate plate (12 samples per plate). The plates were developed and the lipids were visualized by exposure to iodine vapor as described below. Using this procedure, it was possible to examine about 144 samples per day.

Preparation of lipids labelled with various isotopes. A 500 µl sample of a stationary, photoheterotrophic *R. sphaeroides* culture was centrifuged, and the cells were resuspended in 5 ml of the respective Ormerod's labelling medium for ^{32}P-phosphate or ^{35}S-sulfate labelling (see above). For ^{14}C-acetate labelling experiments, regular Ormerod medium was used. The cultures, contained in closed 8-ml screw cap tubes, were incubated for 6 h under a 60 Watt light bulb (irradiance approximately 100 µE/s per m^2) prior to addition of 50 µCi carrier- free ^{32}P-phosphate, 50 µCi carrier-free ^{35}S-sulfate, or 1 µCi ^{14}C-acetate with a specific activity of 56.6 mCi/mmole. Following further incubation for 18 h under the same conditions, the cells were spun down in a clinical centrifuge and resuspended in 0.5 ml water. The cells were extracted by vortexing with 4

ml chloroform/methanol (1:1, v/v). Addition of 1.3 ml 1 M KCl, 0.2 M H_3PO_4, vortexing and centrifugation resulted in phase partitioning of the lipids into the lower chloroform phase. This phase was removed and concentrated down to 0.2 ml by evaporation under a stream of N_2. The lipid extracts were analyzed by TLC as described below.

Time course for the incorporation of ^{35}S-sulfate into proteins and sulfolipid by the wild type and mutants of *R. sphaeroides*. Cells (grown photoheterotrophically in Ormerod's medium) from a 2 ml late log phase culture were spun down and resuspended in 8 x 2 ml low sulfate Ormerod's medium containing 10 μCi ^{35}S-sulfate each (specific activity adjusted to 0.5 μCi/nmole). These cultures, contained in microfuge tubes (2 ml), were grown in the dark at 30°C. One of each of the 8 parallel cultures was analyzed at 3, 6, 9, 12, 15, 18, 24, and 30 hours after addition of the label. Following centrifugation in a microfuge, the cells were extracted with 0.5 ml methanol/chloroform/formic acid/water (12:5:1:2, v/v). The extracts were stored at -20°C. Proteins and cell debris were pelleted by centrifugation. The supernatant was transferred to a new tube and the pellet was reextracted with 0.25 ml methanol/formic acid/water (40:1:59, v/v). Following centrifugation, the supernatants from the two extractions were combined and phase partitioned by addition of 0.15 ml chloroform and 0.05 ml water. The chloroform phase contained the labelled sulfolipid, which was further analyzed by TLC on ammonium sulfate plates (see below). The sulfolipid band was scraped from the plate and the radioactivity was determined by scintillation counting. The

pellets, containing most of the protein and cell debris, were resuspended in 0.05 ml 0.1 N NaOH and incubated at 37°C for 30 min. The samples were diluted with 0.45 ml H_2O and 0.1 ml were used for protein determination using the Pierce BCA protein assay according to the manufactor's instructions. The radioactivity in 0.01 ml samples was measured by scintillation counting.

Preparation of ^{35}S-labelled, water-soluble compounds and sulfolipid from wild type and mutant strains. Cells grown aerobically with shaking in Ormerod's medium in the dark at 30°C were spun down, and resuspended in 20 ml of low- sulfate labelling medium containing 0.5 mCi ^{35}S-sulfate (specific activity 2.5 μCi/nmole). The cultures were incubated for 17 h under the same conditions as the precultures. The cells were harvested by centrifugation and extracted with 5 ml methanol/chloroform/formic acid/H_2O (12:5:1:2, v/v). Following centrifugation in a clinical centrifuge, the supernatant was saved and the pellet was reextracted with 2.5 ml methanol/formic acid/H_2O (40:1:59, v/v). The sample was centrifuged again. To the combined supernatants from the two extractions, 1.5 ml chloroform and 0.5 ml H_2O were added, resulting in the partitioning of water soluble compounds into the aqueous phase and lipids into the chloroform phase. The labelled lipids were analyzed by TLC (see below), while the aqueous phase was evaporated at 45°C under a stream of N_2 and the residue was resuspended in 0.3 ml water. Insoluble material was removed by centrifugation. To separate basic and acidic compounds, the solution was applied to a 0.25 ml Dowex $50H^+$ column which was prepared in a 1 ml plastic pipet tip. The column was eluted with 0.5 ml water followed with

0.5 ml 2 N NH_4OH resulting in the acid and basic fractions respectively. This material was dried under a stream of N_2 at 45°C prior to further analysis.

Quantitative lipid analysis of mutants and wild type by gas chromatography of fatty acid methyl esters. For each strain three 50 ml cultures inoculated with single colonies were grown aerobically with shaking at 32°C using Sistrom's medium. Lipid extracts were prepared from pelleted cells as described for the preparation of labelled lipids (see above) using the same volumes of solutions. After concentration of the chloroform phase, the material was spotted onto two activated silica TLC plates (Si250, Baker). The plates were developed in two dimensions using the system described below. Lipids were visualized with iodine vapor and the spots were individually scraped into 8 ml screw cap tubes. To the samples 5 μg myristic acid methyl ester in 0.1 ml hexane was added as an internal standard. Fatty acid methyl esters were prepared by addition of 1 ml of 1 N methanolic HCl followed by incubation at 80°C for 1 h. Following the addition of 1 ml 0.95% (w/v) KCl, the fatty acid methylesters were extracted into 1 ml hexane. The hexane phase was dried down to 0.1 ml. Samples of 2 μl were injected into the gas chromatograph (Varian 2000), which was equipped with a 2.4 m column (2 mm inner diameter), packed with 3% SP-2310/2% SP-2300 on 100/120 Chromosorb W AW (Supelco). The carrier gas (N_2) flow rate was adjusted to 20 ml/min, and the column temperature was set for 2 min at 180°C, in 10 min to 200°C, and 4 min at 200°C. The fatty acid methyl esters were detected by a flame ionization detector, and the data were integrated by a Spectra Physics integrator. To

calculate the relative amounts of the eight polar lipids included in the analysis, the amount of fatty acids contained in each lipid was calculated. The validity of calculation was based on the assumption that each of the lipids, including the unidentified lipids, contained two fatty acids per molecule, and that the different lipids had a similar fatty acid composition. The amount (μg) of fatty acids in each lipid of a particular sample was calculated from the chromatograms as following:

[(total area under all peaks - standard peak area)/standard peak area] X 5 μg. The relative amount for each lipid in the sample was expressed as percent of all lipids analyzed. It should be pointed out that the data were not corrected for the slightly different response of the different fatty acids in the flame ionization detector. However, the resulting error was considered to be very small compared to the experimental error occurring during sample preparation and handling. A total of 6 samples (2 for each of the three cultures) per strain was analyzed, and means and standard deviations were calculated. The data were subjected to an analysis of variance, to identify significant differences between the lipid composition of the different strains at the 0.05% confidence level.

Quantitative analysis of fatty acid methylesters was also used to estimate the total amount of lipids in ^{35}S-labelled lipid samples.

Analysis of lipids by TLC. One dimensional separation of polar lipids from *R. sphaeroides* was accomplished by using ammonium sulfate treated silica plates (modified according to reference 7). Baker Si250 silica plates with

preadsorbent layer were soaked in 0.15 M ammonium sulfate for 30 s and air dried completely for a minimum of 2 days. Prior to use, the plates were activated for 2.5 h at 120°C. This process drives off the ammonia and leaves behind sulfuric acid. The amount of acidification of the plate depends on the activation time and is crucial for successful separation of phosphatidyl glycerol and sulfoquinovosyl diacylglycerol. The plates were developed in acetone/benzene/water (91:30:8, v/v). This qualitative system was employed for screening of sulfolipid mutants.

For quantitative lipid analysis, a two dimensional TLC-system was used (6). Lipid samples were spotted in the lower right corner of Baker Si250 silica plates (without preadsorbent layer) activated for 30 min at 120°C. The plates were air-dried for 10 min following development of the first dimension with chloroform/methanol/water (65:25:4, v/v). The second dimension was developed with chloroform/acetone/methanol/acetic acid/water (50:20:10:10:5, v/v). Lipids were nondestructively visualized by exposing the plates to iodine vapor, or by spraying the plates with 50% sulfuric acid followed by heating at 160°C for 10-15 min to char the lipids (destructive, permanent). Radioactive labelled lipids were visualized by autoradiography. The ornithine lipid and phosphatidyl ethanolamine were specifically detected due to their free amino group by spraying the plate with ninhydrin solution (0.2% w/v in *sec*-butanol). Following heating at 120°C for 10-15 min the respective spots turned pink.

Analysis of ^{35}S-labelled compounds by 2D TLE/TLC. Water-soluble, ^{35}S-labelled compounds extracted from wild type and mutant strains were

analyzed employing a two dimensional separation system which was originally designed to separate the intermediates of the Calvin cycle (12). For electrophoresis in the first dimension a modified BioRad isoelectric focusing apparatus was used. The apparatus was equipped with two electrode compartments large enough to accommodate paper wicks of 20 cm width. A 0.5 cm thick aluminum plate (20X25 cm) was placed on top of the cooling table, which was flushed with a 20% (v/v) ethylene glycol mixture. The solvent was cycled through a Lauda water bath with the temperature set at -10°C. The aluminum plate was necessary to accommodate a 20 X 20 cm TLC plate. Due to this setup the temperature at the level of the TLC plate was kept at about +10°C throughout the electrophoresis run. Samples (5 µl) were spotted into the lower right corner of a MN300 cellulose TLC plate (Analtech), about 2 cm from each edge. Prior to sample application the plate was sprayed with electrophoresis buffer (955 ml H_2O, 35 ml acetic acid, 10 ml pyridine pH 4.0) and dried in air for about 10 min. A picric acid standard (2.5 µl of 1:10 diluted water saturated solution) was spotted into the lower left corner. The plate was sprayed again with electrophoresis buffer avoiding excessive liquid on the plate to prevent smearing of the sample. The plate was immediately placed onto the precooled aluminum plate, which was covered with plastic wrap to avoid electric shorts. The orientation was such that the spotting area was near the cathode and migration of anions towards the anode could be observed. Two 20 X 7 cm Whatman 3MM paper wicks were soaked in buffer and placed onto the edges of the TLC plate connecting the two electrode compartments.

A sequencing power pack was used to run the plate at a constant voltage of 800 V. The run was finished when the picric acid standard had migrated 5.5 cm towards the anode. Following air drying, the plate was developed in the second dimension by conventional TLC (solvent: 2-butanol/formic acid/H_2O; 60:10:20, v/v). Labelled compounds were detected by autoradiography.

RESULTS AND DISCUSSION

Isolation of mutants deficient in the biosynthesis of the plant sulfolipid in *R. sphaeroides*. To enhance the mutation rate, wild type *R. sphaeroides* cells were treated with the chemical mutagen nitrosoguanidine as described in the methods section of this chapter. This compound causes clustered point mutations, since it acts preferentially at the replication fork (5). The degree of mutagenesis was optimized by incubating the cells for various lengths of time in the presence of 100 µg/ml nitrosoguanidine. Survivor ratios were determined by plating a small aliquot of the cultures immediately following exposure to the mutagen. Since actively growing cells during the exponential phase of growth contain at least two copies of the genome due to continuous replication, the cultures were allowed to grow for 18 h before plating, to ensure segregation of the mutagenized chromosomes and genotypic purity of the resulting colonies. The degree of mutagenesis was monitored by the frequency of mutations for resistance to the antibiotics rifampicin and streptomycin. In addition, the occurrence of pigment mutations (e.g. dark red colonies as opposed to wild type pink colonies) provided a visual indication of

Table 3-5: Effect of duration of exposure to nitrosoguanidine on survival and on induction of mutations.

	Exposure time (min)				
	0	10	20	30	45
Survivors (%)	100.0	10.0	1.3	0.7	0.7
rifR[1]	$2.5X10^{-6}$	$1.0X10^{-5}$	$2.3X10^{-5}$	$5.0X10^{-4}$	$2.5X10^{-4}$
strepR[1]	$5.0X10^{-7}$	$2.5X10^{-6}$	$7.0X10^{-6}$	$4.0X10^{-5}$	$7.0X10^{-5}$

[1]. The frequency of antibiotic resistance is expressed as resistant colonies per total number of colonies.

the extent of the mutagenesis. Table 3-5 shows the results for a typical mutagenesis experiment.

The exact measurements of the frequencies by plating was adversely affected by the fact that cells treated with the mutagen had a tendency to clump. Therefore, the values given in Table 3-5 can only be considered approximate. However, the results suggested that in the most heavily mutagenized population (45 min exposure) the occurrence of antibiotic resistance is increased approximately 100-fold compared to spontaneous frequencies in wild type cultures (0 exposure). The number of survivors fell below 1% for the longest exposures to the mutagen. Upon visual examination

of the mutagenized colonies the extent of the mutagenesis became quite clear, since a high frequency of white, yellow, dark red, or abnormally shaped colonies compared to wild type colonies was observed. A low frequency of colonies showed sectors with altered pigmentation, indicating that even after 18 h of incubation subsequent to the removal of the mutagen, some segregation or mutagenesis of the multiple genomes was still occurring.

The most heavily mutagenized cell population from this experiment (Table 3-5) was used to screen for mutants deficient in the biosynthesis of the plant sulfolipid. Colonies were restreaked on Z-broth plates and lipids were extracted as described in the method section. Crucial to the screening procedure was a TLC system employing ammonium sulfate treated silica plates (see methods). This TLC system allowed for the separation in one dimension of all the major polar lipids found in *R. sphaeroides*. Figure 3-2 shows the separation of ^{35}S-, ^{32}P-, and ^{14}C-labelled lipids on these plates.

The incorporation of the three different labels helped to identify the various lipids. All the lipids which were clearly identified by comigration with standards and through specific staining procedures were labelled. Three additional phospholipids of lower abundance (Figure 3-2, U) were not identified. The three lanes shown in Figure 3-2 were run on different TLC plates. As is obvious from comparison of the absolute mobility of the lipids in the two right lanes, the R_f values for each lipid varied considerably on these type of plates. The R_f of phosphatidyl glycerol was particularly sensitive to the degree of activation of the TLC plate. As described in the methods section,

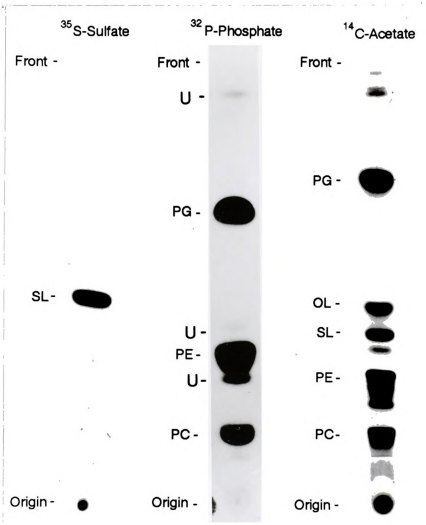


Figure 3-2: Separation of *in vivo* labelled lipids from *R. sphaeroides* on ammonium sulfate plates. The labelling substrate is indicated on top of the lanes. The three autoradiograms shown were derived from three different experiments. OL, ornithine lipid; PC, phosphatidyl choline; PE, phosphatidyl ethanolamine; PG, phosphatidyl glycerol; SL,sulfolipid; U, unidentified phospholipids.

activation of ammonium sulfate-treated plates at 120°C produces sulfuric acid, which protonates phosphatidyl glycerol, making it less polar. Less polar lipids run in this TLC system closer to the front. Sulfolipid on the other hand is not being protonated, and its migration was therefore not affected by the activation process. However, given that the plate was sufficiently activated prior to the run, phosphatidyl glycerol was readily separated away from the sulfolipid, permitting easy scoring of mutants deficient in the biosynthesis of the sulfolipid or of any other lipid.

During the screening procedure, 1512 cell lines, the equivalent of 126 TLC plates, were analyzed for their lipid composition. The plates were visually examined following staining of the lipids with iodine. Figure 3-3 shows a picture of a developed TLC plate as it appeared during the screening procedure. One of the samples on this plate contained no visible sulfolipid band, identifying the cell line from which this extract was derived as a probable mutant. A total number of 15 putative mutants showing very low to moderately low levels of sulfolipid in relation to the other lipids was saved. In addition, 6 putative mutants with dramatic changes in lipid composition other than sulfolipid were isolated. These included two mutants seemingly lacking phosphatidyl choline.

For practical reasons, it was necessary to focus further work on a selected number of mutants, which were most likely to provide some clues about the biosynthetic pathway for sulfolipid. Two criteria were used to select the mutants for further study. First, the mutant must grow on Ormerod's

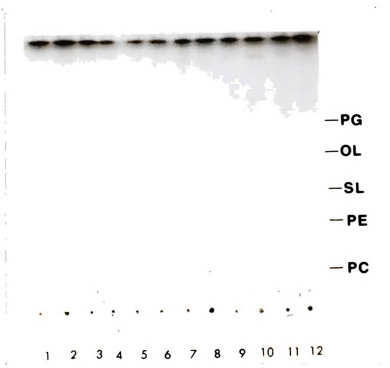


Figure 3-3: Separation of lipid samples prepared from twelve mutagenized cell lines on ammonium sulfate plates. One sample (lane 8) does not show a sulfolipid band indicating a putative sulfolipid mutant. OL, ornithine lipid; PC, phosphatidyl choline; PE, phosphatidyl ethanolamine; PG, phosphatidyl glycerol; SL, sulfolipid.

minimal medium to exclude the possibility that it was not deficient in sulfolipid due to a defect in the general sulfate assimilation pathway. Second, it was decided to focus on mutants that accumulated labelled compounds, which might represent putative precursors of sulfolipid biosynthesis.

To obtain a broader knowledge of the phenotype of the putative mutants, the 15 putative sulfolipid mutants were rescreened using a labelling procedure. Those which grew on Ormerod's minimal medium were labelled with ^{35}S-sulfate. Labelled cells were extracted, and the lipids and water soluble compounds were analyzed. Equal amounts of lipids from each mutant sample, as determined by fatty acid methyl ester quantification by gas chromatography, were loaded onto the TLC plate. Following development, the radioactivity in the sulfolipid band was determined and expressed as μCi/mg fatty acid methyl esters. The water soluble compounds were separated into basic and acidic compounds, and analyzed by 2D TLE/TLC. The patterns observed in mutant samples were compared to that obtained from a wild type sample following autoradiography. The data obtained in this screening experiment are shown in Table 3-6. They have to be considered preliminary, since only one labelling experiment for each putative mutant was performed.

Based on the screening experiment described above, the three mutant lines CHB16, CHB17, and CHB18 were selected for further analysis. The selection was based on the fact that CHB16 and CHB17 showed very small, almost undetectable amounts of labelled sulfolipid, while all three of these

Table 3-6: Preliminary characterization of putative sulfolipid mutants. Growth refers to visible colony formation after 2 weeks of anaerobic, photoheterotrophic growth on Ormerod's minimal medium. The incorporation of ^{35}S-sulfate into sulfolipid (SL) is expressed as radioactivity in the sulfolipid band per total fatty acid methyl esters (FAM) in the lipid sample (μCi/mg FAM). The mean value for three independent wild type labelling experiments was 22.2 ± 6.7 (μCi/mg FAM). The column "New Compounds" indicates the presence of additional ^{35}S-sulfate-labelled, water-soluble compounds as compared to the wild type.

Strain	Growth	SL(μCi/mg FAM)	New Compounds
CHB2	+	23.3	–
CHB5	+	7.9	+
CHB7	–	n.d.	n.d.
CHB8	–	n.d.	n.d.
CHB9	+	8.1	–
CHB10	+	36.2	–
CHB12	+	15.9	–
CHB13	–	n.d.	n.d.
CHB14	+	11.3	–
CHB15	+	*	–
CHB16	+	0.3	(+)
CHB17	+	0.2	+
CHB18	+	1.3	+
CHB19	+	0.2	–
CHB21	+	4.2	–

* The sample was lost prior to scintillation counting. However, examination of the autoradiogram showed no clear reduction of labelled sulfolipid compared to wt in this mutant.

mutants seemed to accumulate water-soluble, ^{35}S-sulfate-labelled, compounds. In particular CHB18 showed three very prominent additional spots on the 2D chromatograms compared to the wild type (see further below). However, the additional spot initially seen upon analysis of the extract from mutant CHB16 could not be reproducibly detected in later experiments. Several of the putative mutants clearly did not rescreen (CHB2, CHB10, CHB12), while others (CHB5, CHB9, CHB14, CHB15), showing slight reductions of labelled sulfolipid compared to the wild type, need to be retested in the future to be certain. Two certain candidates for sulfolipid mutants, on which no further analysis has been performed to date, are CHB19 and CHB21. It will be interesting to know whether these two mutants carry mutations in genes other than the three genes mutated in CHB16, CHB17, and CHB18.

The primary phenotype of three selected sulfolipid mutants. As will be seen in chapters 4 and 5, the three mutants CHB16, CHB17, and CHB18 represent three different genes involved in sulfolipid biosynthesis. These genes were given the designation *sqdA* (CHB17), *sqdB* (CHB16), and *sqdC* (CHB18) and the mutants will be referred to by these designations throughout this text.

To confirm the accumulation of novel ^{35}S-labelled, water-soluble compounds in extracts from the three mutants, three individual labelling experiments for each of the mutants and the wild type were performed. The labelled, water-soluble compounds were analyzed as described in the methods section. The results were found to be reproducible, and the 2D chromatograms from one representative experiment are shown in Figures 3-4

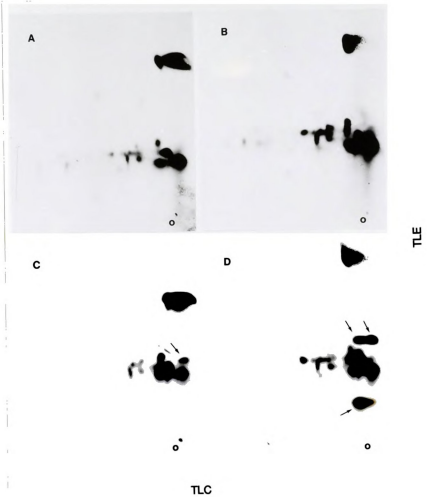


Figure 3-4: Separation of water-soluble, ^{35}S-labelled compounds contained in the acidic fraction of wild type and mutant extracts by 2D TLE/TLC. Samples containing approximately 0.5-1.0 μCi were applied to each plate, and the film was exposed for 3 days. A: wild type; B: *sqdB* mutant; C: *sqdA* mutant; D: *sqdC* mutant. Spots which reproducibly appeared in the mutant samples but not in the wild type are indicated by arrows. The origin is marked with an o.

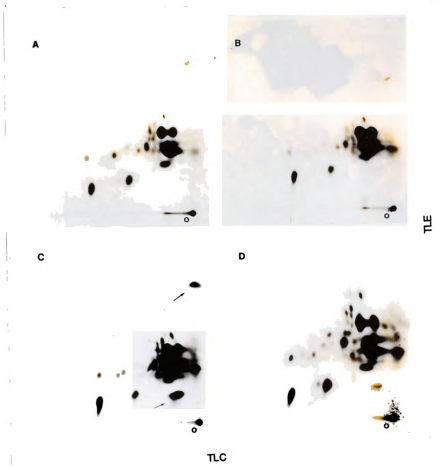


Figure 3-5: Separation of water-soluble, ^{35}S-labelled compounds contained in the basic fraction of wild type and mutant extracts by 2D TLE/TLC. Samples containing approximately 0.5-1.0 μCi were applied to each plate, and the film was exposed for 3 days. A: wild type; B: *sqdB* mutant; C: *sqdA* mutant; D: *sqdC* mutant. Spots which reproducibly appeared in the mutant samples but not in the wild type are indicated by arrows. The origin is marked with an o.

and 3-5. Based on the assumption that some of the precursors might contain sulfonic acid or phosphate ester groups, a separation system was chosen, which is commonly used for the separation of photosynthetic intermediates of the reductive pentose phosphate cycle (12). The electrophoresis was run in such a way as to observe anions, which are charged at the pH (4.0) of the running buffer. These included all the phosphate esters and sulfonic acids harboring a net negative charge at pH (4.0). However, ions which were positively charged at pH (4.0) such as certain amino acids migrating towards the cathode, were lost for the analysis, since they ran off the plate into the paper wick. Aqueous extracts from the various cell lines were first separated into acid and basic compounds using a Dowex $50H^+$ column. The aqueous "flow through" contained the acidic compounds while the ammonium hydroxide eluate made up the basic fraction, containing mostly amino acids. The compounds contained in the acid fraction (Figure 3-4) and those contained in the basic fraction (Figure 3-5) showed characteristic and reproducible patterns. By spotting a sample into the middle of the plate instead of into the lower right corner, it could be estimated that about 5% of the radioactivity in the acidic fractions and about 30-40% of the radioactivity in the basic fractions was lost due to migration towards the cathode into the paper wick.

As Figure 3-4 shows (indicated by arrows), the *sqdC* mutant accumulated 3 compounds visible in the acid fraction, which could not be detected in the wild type sample. The *sqdA* mutant accumulated one

compound compared to wild type, while no reproducible difference between the wild type and the *sqdB* mutant could be detected. In the basic fraction (Figure 3-5) only the *sqdA* mutant showed the accumulation of new compounds compared to the wild type. It was assumed that the appearance of these compounds in the mutants was somehow related to a block in the pathway for the biosynthesis of the plant sulfolipid. Due to this specific block caused by the mutation, precursors of sulfolipid biosynthesis accumulated, which were visualized by the experiment described above. The *sqdB* mutant is clearly deficient in sulfolipid biosynthesis, yet did not show the accumulation of ^{35}S-labelled compounds in the experiment described above. There are three possible explanations: first, such compounds accumulated but were missed because they ran off the plate during electrophoresis; second, the respective precursor was shunted into a different pathway and did not accumulate; or third, the block in the pathway occurred at a very early stage in sulfolipid biosynthesis prior to the incorporation of sulfur into a precursor molecule. Therefore, no ^{35}S-labelled compound would be expected to accumulate.

To test whether the three *sqd* mutants harbor defects in the general sulfur assimilation pathway which, in turn, could affect protein biosynthesis, incorporation rates of ^{35}S-sulfate into proteins and the sulfolipid were compared as described in the method section. As can be seen in Figure 3-6, all three mutants and the wild type incorporated labelled sulfate into protein at a similar rate, while the capability of the mutants to label sulfolipid was greatly reduced. Only in the *sqdC* mutant was the sulfolipid labelled at a detectable

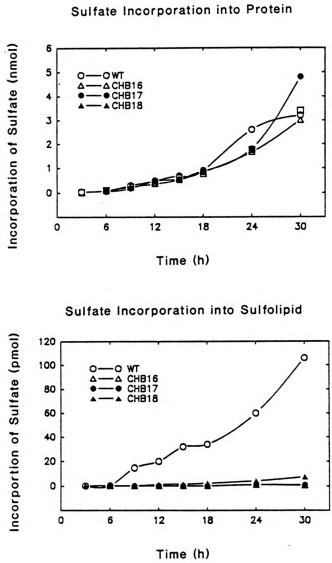


Figure 3-6: Time course for the incorporation of ^{35}S-labelled sulfate into protein (upper panel) and sulfolipid (lower panel) by wild type and mutant strains CHB16 (*sqdB*), CHB17 (*sqdA*), and CHB18 (*sqdC*).

rate. The increase in labelled protein was paralleled by the increase of total protein as determined by a colorimetric protein assay (data not shown). The result was in agreement with the fact that all three mutants can grow on Ormerod's malate-basal-salt medium, which does not contain any reduced sulfur compounds such as cysteine. Therefore, it was assumed that the defects in the three mutants must be specific for sulfolipid biosynthesis, since the biosynthesis of cysteine and its methionine were not affected in the three mutants.

In the experiments described above, the capability of the mutants to incorporate ^{35}S-sulfate into sulfolipid was used as a measure of their ability to synthesize sulfolipid *de novo*. This did not necessarily reflect the steady state levels of sulfolipid in the mutants compared to the wild type. To determine the lipid composition of the wild type and the mutants in a quantitative way, lipid extracts from the wild type and the *sqd* mutants were separated in two dimensions on silica TLC plates. In Figure 3-7 the typical lipid pattern for the wild type and the three *sqd* mutants is shown. Lipids were identified by isolation of the spots from the plates, followed by cochromatography of the lipids with standards on ammonium sulfate plates. No sulfolipid spot could be detected by charring of the separated lipids from the *sqdA* and *sqdB* mutant samples, while a weak sulfolipid spot was visible in the *sqdC* mutant sample. To quantify the lipids, the iodine stained spots were isolated from the plates and fatty acid methyl esters were prepared and quantified by gas chromatography using an internal standard. Myristyl methyl ester was used as

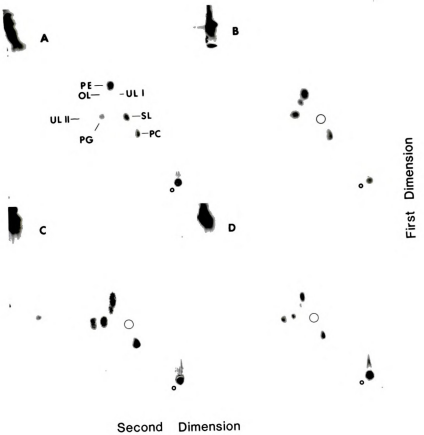


Figure 3-7: Separation of lipids isolated from wild type (A) and three *sqd* mutants by 2D TLC. A, wild type; B, *sqdB*; C, *sqdA*; D, *sqdC*. Ol, ornithine lipid; PC, phosphatidyl choline; PG, phosphatidyl glycerol; PE, phosphatidyl ethanolamine; SL, sulfolipid; ULI and ULII, two unidentified lipids. The position for the sulfolipid in panels B,C, and D are circled. The origin is indicated with o. The lipids on the plates were visualized by charring.

Table 3-7: Polar lipid composition (weight %) of wild type and *sqd* mutant strains of *R. sphaeroides* grown under chemoheterotrophic conditions. OI, ornithine lipid; PC, phosphatidyl choline; PG, phosphatidyl glycerol; PE, phosphatidyl ethanolamine; SL, sulfolipid; ULI and ULII, two unidentified lipids; *WT*, wild type. For each strain two samples were analyzed from each of three independent cultures. These data were subjected to analysis of variance. Means with identical superscripts are not significantly different at the 0.05% level.

Lipid	*WT*	*sqdA*	*sqdB*	*sqdC*	pooled SD
SL	$5.4^{a} \pm 0.4$	$0.7^{b} \pm 0.2$	$0.4^{b} \pm 0.2$	$2.6^{c} \pm 1.0$	0.6
PC	$15.4^{a} \pm 0.8$	$20.0^{b} \pm 1.5$	$17.2^{a} \pm 0.7$	$21.2^{b} \pm 1.9$	1.3
PG	$18.7^{a} \pm 0.9$	$24.2^{c} \pm 1.4$	$29.6^{b} \pm 1.6$	$23.1^{c} \pm 2.8$	1.8
PE	$44.8^{a} \pm 0.9$	$38.4^{b} \pm 2.1$	$42.7^{a} \pm 1.4$	$35.4^{b} \pm 2.6$	1.8
OL	$1.8^{a} \pm 0.5$	$3.2^{ab} \pm 1.4$	$2.4^{ab} \pm 0.4$	$3.5^{b} \pm 0.5$	0.8
ULI	$1.7^{a} \pm 0.3$	$2.0^{ab} \pm 0.4$	$2.1^{ab} \pm 0.5$	$2.7^{b} \pm 0.4$	0.4
ULII	$12.0^{a} \pm 1.1$	$11.6^{a} \pm 1.3$	$5.3^{b} \pm 1.0$	$11.4^{a} \pm 1.0$	1.2

standard, since only negligible amounts of endogenous myristic acid were found in the bacterial lipids. The lipids included in the analysis were phosphatidyl choline, phosphatidyl ethanolamine, phosphatidyl glycerol, ornithine lipid, sulfolipid, and two unidentified lipids (Figure 3-7, panel A). These lipids include more than 95% of the polar lipids. The large spot in the upper left corner represents neutral lipids, in particular carotenoids. Two aliquots, each from three independent aerobic cultures grown in Sistrom's medium, for each of the three *sqd* mutants and the wild type were analyzed, and the data were submitted to an analysis of variance. From the result shown in Table 3-7, it became clear that the sulfolipid content in the *sqdA* and *sqdB* mutants was reduced to about 10% of the wild type level.

Based on the labelling experiments described above, in which very small amounts of labelled sulfolipid were detected (Table 3-6), as well as based on the result of the lipid analysis, it was concluded that the *sqdA* and *sqdB* mutants are slightly leaky, and still retained about 10% of their capacity to synthesize sulfolipid. The *sqdC* mutant shows a reduction of the steady state levels of sulfolipid down to little less than 50% of the wild type levels (Table 3-7). This value seemed to be surprisingly high since incorporation of labelled sulfate into sulfolipid occurred at much lower rates than 50% of the wild type rate (Figure 3-6, and Table 3-6). As mentioned above, incorporation rates of labelled sulfate into sulfolipid may not necessary reflect steady state levels. The pool size of the various intermediates, the uptake of the labelled substrate, as well as the turnover rates may influence the level of labelled

sulfolipid detected in a labelling experiment. Unless the system was completely equilibrated with the labelled precursor, the degree of label found in the sulfolipid may not necessary reflect the steady state level of sulfolipid as determined by quantification of the total mass of the lipid. It was most likely that the system had not been in equilibrium during the labelling time course described in Figure 3-6, explaining the differences between the label experiment and the lipid analysis experiment. Therefore, it is suggested that the values presented in Table 3-7 reflect more realistically the steady state levels for each lipid than those observed in the labelling experiments.

It should be mentioned that the *sqdB* mutant showed significantly higher levels of phosphatidyl glycerol, and lower levels of the unidentified lipid II (ULII; see Table 3-7) than the wild type. The *sqdA* and *sqdC* mutants on the other hand, showed significantly higher levels of phosphatidyl choline and phosphatidyl glycerol than the wild type. It was not clear, whether this observation was related to the defect in sulfolipid biosynthesis, or whether mutations in the background resulted in these changes.

CONCLUSIONS

The experiments described in this chapter were designed to establish the purple bacterium *R. sphaeroides* as a model system for the study of the biosynthetic pathway for the plant sulfolipid. This goal was clearly achieved. A TLC screening procedure was developed, facilitating the isolation of a series of sulfolipid deficient mutants from a chemically mutagenized population of

R. sphaeroides. Some of the mutants accumulate ^{35}S-labelled, water-soluble compounds as detected by 2D-TLE/TLC. These compounds may be precursors which accumulate due to specific blocks in the pathway for sulfolipid biosynthesis.

The sulfolipid mutants described in this work are the first isolated from any organism. Because it was not efficient to perform a simultaneous characterization of all mutants, the following work was focused on three *sqd* mutants: *sqdA* (CHB17), *sqdB* (CHB16), and *sqdC* (CHB18). These three mutants were selected, because they showed a pronounced deficiency in their sulfolipid levels, and because they revealed three distinctive patterns for the accumulation of ^{35}S-labelled compounds. Based on this result, it was assumed that these three mutants might mark three different genes involved in sulfolipid biosynthesis. It should be noted that the mutant collection described above (Table 3-6) contains at least two more mutants with significant sulfolipid deficiencies (CHB19, CHB21). Once the genes for the three *sqd* mutants have been isolated, a complementation analysis can be performed to test whether these mutants carry mutations in different genes involved in sulfolipid biosynthesis, or whether they represent alleles of the three *sqd* mutants.

Based on the results described above, it was concluded that the next logical step for the characterization of these mutants would be the isolation of the wild type genes complementing these mutations. First of all, any characterization of the phenotype of the mutants, could be misleading due to background mutations. Since the cells were quite heavily mutagenized, it is

certain that the described strains contain more than one mutation. Thus, in principle, the accumulation of labelled compounds seen in the *sqdA* and *sqdC* mutants may be due to a second mutation in the background and therefore unrelated to the sulfolipid phenotype. To avoid this problem, isogenic lines have to be constructed. In order to accomplish this, a series of experiments was undertaken to clone the genes, which complement the various *sqd* mutants. The next two chapters of this work deal with the isolation and characterization of the genes complementing the *sqdA*, *sqdB*,and *sqdC* mutants.

REFERENCES

1. Barber J., Gounaris K. (1986). What role does sulpholipid play within thylakoid membranes? Photosynthesis Res. 9:239-249

2. Donahue T.J., Cain B.D., Kaplan S. (1982). Purification and characterization of an N-acylphosphatidylserine from *R. sphaeroides*. Biochemistry 21:2765-2773

3. Falcone D.L., Quivey R.G. JR, Tabita F.R. (1988). Transposon mutagenesis and physiological analysis of strains containing inactivated form I and form II ribulose bisphosphate carboxylase/oxygenase genes in *R. sphaeroides*. J. Bacteriol. 170:5-11

4. Gorchein A. (1967). The separation and identification of the lipids of *R. sphaeroides*. Proc. Roy. Soc. N. 170:279-297

5. Guerola N., Ingraham J.L., Cereda-Olmedo E. (1971). Induction of closely linked multiple mutations by nitrosoguanidine. Nature New Biology 230:122-125

6. Joyard J., Blee E., Douce R. (1986). Sulfolipid synthesis from $^{35}SO^{2-}_{4}$ and [1-^{14}C]acetate in isolated spinach chloroplasts. Biochim. Biophys. Acta 879:78-87

7. Kahn M.-U., Willams J. P. (1977). Improved thin-layer chromatographic method for the separation of major phospholipids and glycolipids from plant lipid extracts and phosphatidyl glycerol and bis(monoacylglyceryl) phosphate from animal lipid extracts. J. Chromato. 140:179-185

8. Kiley P.J., Kaplan S. (1988). Molecular genetics of photosynthetic membrane biosynthesis in *R. sphaeroides*. Microbiol. Rev. 52:50-69

9. Ormerod J.G., Ormerod K.S., Gest H. (1961). Light-dependent utilization of organic compounds and photoreduction of molecular hydrogen by photosynthetic bacteria; relationships with nitrogen metabolism. Arch. Biochem. Biophys. 94:449-463

10. Pemberton J.M., Bowen A.R.S.G.(1981). High-frequency chromosome transfer in *R. sphaeroides* promoted by broad-host range plasmid RP1 carrying mercury transposon Tn501. J. Bacteriol. 147:110-117

11. Radunz A. (1969). Ueber das Sulfochinovosyl-diacylglycerin aus hoeheren Pflanzen, Algen und Purpurbakterien. Hoppe-Seylers Z. Physiol. Chem. 350:411-417

12. Schurmann P. (1969). Separation of phosphate esters and algal extracts by thin-layer electrophoresis and chromatography. J. Chromatog. 39:507-5-9

13. Sistrom W.R. (1960). A requirement of sodium in the growth of *R. sphaeroides*. J. Gen. Microbiol. 22:778-785

14. Sistrom W.R. (1962). The kinetics of the synthesis of photopigments in *R. sphaeroides*. J. Gen. Microbiol. 28:607-616

15. Suwanto A., Kaplan S. (1989). Physical and genetic mapping of the *R. sphaeroides* 2.4.1. genome: presence of two unique circular chromosomes. J. Bacteriol. 171:5850-5859

16. Weaver K.E., Tabita F.R. (1985). Complementation of a *R. sphaeroides* ribulose bisphosphate carboxylase-oxygenase regulatory mutant from a genomic library. J. Bacteriol. 164:147-154

17. Wood B.J.B., Nichols B.W., James A.T. (1965). The lipids and fatty acid metabolism of photosynthetic bacteria. Biochim. Biophys. Acta 106:261-273

CHAPTER 4

ISOLATION AND CHARACTERIZATION OF A CLONE COMPLEMENTING THE SULFOLIPID-DEFICIENT *sqdA* MUTANT OF *R. SPHAEROIDES*

ABSTRACT

A cosmid library was prepared from *R. sphaeroides* wild type DNA. Individual cosmids were mated into the sulfolipid *sqdA* mutant, and lipid extracts from the exconjugants were analyzed by TLC for the presence of the sulfolipid. Two cosmid clones were isolated which were able to restore the biosynthesis of the plant sulfolipid in the *sqdA* mutant. A physical map for one of the cosmids was constructed, and two 5-6 kb subfragments derived from this cosmid were found to complement the mutant. A region of 2.1 kb common to both of these subclones was sequenced and found to contain an open reading frame of 295 amino acids preceded by a ribosome binding site. Deletion analysis in combination with cloning of the open reading frame into a novel expression vector, constructed during this study, suggested that the protein encoded by the open reading frame can restore the wild type phenotype of the *sqdA* mutant in a complementation assay. By searching two protein banks, no known proteins with striking homology to the putative protein encoded by the *sqdA* gene were found, indicating that this gene codes for a novel protein involved in plant sulfolipid biosynthesis.

INTRODUCTION

In the previous chapter the isolation and characterization of three sulfolipid-deficient mutants was described. It was noted that further biochemical or physiological analysis is hampered by the fact that the mutants contain secondary mutations in their genetic background, making it difficult to link the observed aspects of the phenotype directly to the mutation causing the sulfolipid deficiency. Isogenic lines can be created by cloning the gene, which complements the mutation, and using it, either to construct a null mutation in the wild type background, by gene replacement, or by introducing the wild type gene carried on a single copy plasmid into the mutant, restoring the wild type phenotype. Such a line, which would be wild type with respect to sulfolipid biosynthesis, could be compared with the mutant carrying a control plasmid lacking the complementing gene. Because both lines carry the same background mutations, any differences observed between the lines can be directly attributed to the mutation.

Analysis of the DNA sequence of a gene which complements can also provide clues about the nature of the protein product of the gene. Given the sequence, it is possible to determine by hydropathy analysis according to the method of Kyte and Doolittle (16) whether a protein is membrane bound or soluble (16). Also, protein-sequence data bases can be searched to identify similarities with other proteins, which may provide clues about the type of enzymatic reaction catalyzed by the protein.

By expression of the cloned gene in *E. coli*, it is frequently possible to

purify large amounts of the protein. The purified protein can either be studied directly, or it can be used to raise specific antibodies. These antibodies can serve as tools to study the protein and may be useful in identifying a similar protein involved in sulfolipid biosynthesis in other photosynthetic organisms.

For the reasons described above, the next logical step in the establishment of a model system for the study of the biosynthesis of the plant sulfolipid was the isolation of the genes complementing the mutations in the *sqdA*, *sqdB*, and *sqdC* mutants. In this chapter, the isolation and characterization of the gene complementing the *sqdA* mutant is described.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains used in the experiments described in this chapter and in chapter 5 are shown in Table 4-1. The source of the plasmids used in this chapter, and the construction of novel plasmids, is presented in Table 4-2.

Growth media. *R. sphaeroides* strains were grown in Sistrom's medium as described in the methods section of chapter 3. When required, 0.8 μg/ml tetracycline were added to agar plates containing Sistrom's medium.

E. coli strains were grown on LB medium containing 10 g bacto-tryptone, 5 g bacto-yeast extract, and 10 g NaCl per liter (pH 7.5). For the preparation of single stranded DNA for sequencing purposes, the cells were grown in 2X YT medium containing 16 g bacto-tryptone, 10 g yeast extract, and 10 g NaCl per liter (pH 7.4). When required, 10 μg/ml tetracycline,

Table 4-1: Bacterial strains.

Bacterial Strains	Relevant Characteristics	Source or Reference
R. sphaeroides:		
2.4.1	wild type	R. L. Uffen MSU
CHB17	sulfolipid deficient, carrying a mutation in *sqdA*	this study, chapter 3
CHB16	sulfolipid deficient, carrying a mutation in *sqdB*	this study, chapter 3
CHB18	sulfolipid deficient, carrying a mutation in *sqdC*	this study, chapter 3
E. coli:		
HB101	F- proA2 recA13 mcrB	(17)
DH5αF'	F'/endA1 recA1 Δ(lacZYA-argF)U169 (φ80dlacΔ(lacZ)M15)	(20)
MM294	F- endA1 hsdR17 thi-1	(2)
S17-1	RP4-2-Tc::Mu-Km::Tn7 (integrated) pro^-, res^- mod^+	A. Puehler (22)

Table 4-2: Plasmid sources and derivation.

Plasmid	Description or Construction	Source or Reference
pBS-KS$^+$, pBS-KS$^-$	Bluescript Vector	Stratagene
pRK2013	Km^r self-transmissible RK2 derivative containing ColE1 replicon and transfer functions to mobilize RK2 derivatives	(10)
pLA2917	Cosmid Vector (Tc^r) derived from RK2	R.S. Hanson (1)
pCHB500	Expression vector for *R. sphaeroides*	Appendix B
pCHB1701	pLA2917 derived Cosmid clone complementing the *sqdA* mutant	Figure 4-2 this study
pCHB17011, pCHB17012	2 subclones of pCHB1701 in pLA2917 complementing the *sqdA* mutant	Figure 4-3 this study
pCHB17H3B	10.5 kb HindIII fragment of pCHB1701 in pBS-KS$^+$	Figure 4-2 this study
pCHB17B2B	10.5 kb BglII fragment of pCHB1701 in pBS-KS$^+$	Figure 4-2 this study
pCHB17B20-D$^+$, pCHB17B20-A$^-$,	2.1 kb BamHI fragment of pCHB17012 in BamHI of pBS-KS$^+$ and -KS$^-$ (same orientation)	Figure 4-3 this study
pCHB17B20-A$^+$	(as above, but opposite orientation)	

Table 4-2 (cont'd)

Plasmid	Description or Construction	Source or Reference
pCHB1741, pCHB1742	2.1 kb XbaI, EcoRI fragment of pCHB17B20-D$^+$ and -A$^+$ in XbaI, EcoRI of pCHB500	Figure 4-9 this study
pCHB1751, pCHB1752	1.2 kb BamHI fragment of pCHB17B20-A$^+$ in both orientations in BamHI of pCHB500	Figure 4-9 this study
pCHB17NS-I, pCHB17NS-II	1.74 kb StuI, NaeI (partial) fragment of pCHB17B20 in both orientations in EcoRV of pBS-KS$^+$	this study
pCHB1761, pCHB1762	1.8 kb KpnI, EcoRI fragment of pCHB17NS-I and -II in KpnI, EcoRI of pCHB500	Figure 4-9 this study

50 μg/ml kanamycin, or 100 μg/ml ampicillin were added to the medium. In order to score the lacZ phenotype of clones carrying pBluescript derivatives, 80 μl of a solution of 2% (w/v) X-gal in dimethyl formamide were spread onto the agar plates prior to plating of the cells.

Construction of the cosmid library. High molecular weight DNA from *R. sphaeroides* wild type cells grown photoheterotrophically in Sistrom's medium (chapter 3), was prepared using a modified procedure originally

designed for yeast chromosomes (7). Pelleted cells from a 200 ml culture were resuspended in 13 ml 8% (w/v) sucrose, 50 mM Tris-Cl, 50 mM EDTA (pH 8.0) buffer. Following the addition of 3.2 ml of the same buffer containing 10 mg/ml lysozyme (final concentration 2 mg/ml), the suspension was incubated at 30°C. Lysis was monitored by mixing 100 µl of the solution with 100 µl of 1% SDS. Sufficient lysis capability was generally observed after about 30 min. Four ml aliquots of the sphaeroplast solution were lysed by pipeting the solution very slowly under gentle shaking into a 6 ml solution of 3% (w/v) Na Lauryl Sarcosine, 0.5 M Tris-Cl, 0.2 M EDTA (pH 9.0). The solution was heated for 15 min at 65°C and subsequently cooled on ice. Aliquots (10 ml) of this DNA solution were loaded onto a sucrose step gradient (bottom to top: 10 ml 50% w/v sucrose cushion, 10 ml 20% w/v sucrose. and 10 ml, 10% w/v sucrose), and the samples were centrifuged at 27,000 rpm for 3 h at 15°C in a SW27 rotor (Beckman). Brown, flocculent genomic DNA floating just inside the 50% sucrose cushion, was removed from the tubes and dialyzed overnight at 4°C in 1.5 l 10 mM Tris-Cl, 1 mM EDTA (pH 8.0) using dialysis tubing (10,000 MW cut- off). Aliquots of 0.5 ml were distributed into microfuge tubes using a sterilized pair of scissors. Following soaking of the DNA for 10 min in 0.5 ml buffer-equilibrated phenol and 0.5 ml chloroform, the DNA was dialyzed again for two days with 4 changes of buffer in 1 l of 10 mM Tris-Cl, 1 mM EDTA (pH 8.O) at 4°C. The DNA, which appeared quite green, was stored at 4°C.

Aliquots of 40 µl (approximately 5-10 µg) from this DNA preparation were partially digested for 1 h at 37°C with Sau3Al in a total volume of 80 µl of

the recommended enzyme buffer using between 10 and 0.001 units of the enzyme. The solution was stirred periodically using a closed plastic pipet tip. The reaction was stopped by addition of 20 µl of 50 mM EDTA solution. The DNA was separated on a 0.4% (w/v) agarose gel as described further below. Intact lambda phage (48 kb), phage cut with SalI or with HindIII were used as size markers. Despite its "dirty" appearance, the DNA was readily digested. Fragments in the size range from 20 to 35 kb were isolated by electroelution from the respective gel slice and purified using standard DNA procedures (17). About 150 ng of genomic DNA fragments were ligated into the BglII site of the cosmid vector (500 ng of prepared pLA2917). The vector DNA was purified twice on a cesium chloride gradient, overdigested with BglII, and treated with calf intestinal phosphatase using standard procedures (17). Ligation tests using the vector prepared in this way, revealed only very small amounts of self-ligation. The resulting cosmids were packaged *in vitro* using commercially available (Promega) lambda phage packaging extracts, and the resulting phage were used to infect *E. coli* (HB101 or S17-1). Cell lines containing the cosmids were selected on LB plates containing 10 µg/ml tetracycline. Individual colonies were transferred to 96-well plates (200 µl LB tetracycline medium per well). Following the addition of 20% (v/v) DMSO, the plates were kept as frozen stocks at -70°C.

Triparental mating and complementation assay. Individual cosmids carried by HB101 host cells were introduced into the *R. sphaeroides sqdA* mutant using a triparental mating procedure. The *sqdA* mutant cells were

grown photoheterotrophically in 200 ml Sistrom's medium as described in chapter 3, until the culture reached mid to late log phase. An aliquot of 5 ml from this culture was diluted with 20 ml Sistrom's medium, and 100 µl of a concentrated suspension of MM294 cells containing the helper plasmid pRK2013 were added. The helper strain was grown in 50 ml LB kanamycin medium to midlog phase, and was resuspended in 1 ml LB following centrifugation. This mixture was poured over a sterile cellulose acetate filter (MSI, 9 cm diameter) under suction. Following the complete removal of the liquid, the underside of the filter was blotted on sterile Whatman #2 filter paper and placed onto a LB agar plate. Using a sterilized pronging device matching 48 wells of a 96-well microtiter plate, frozen material from the -70°C cosmid library stocks was directly transferred onto the filter containing the lawn of *sqdA* mutant and helper cells. Mating efficiency increased slightly when freshly grown cultures inoculated from the stock cultures were used as donors. However, since there was a concern that any additional growth of the cosmid clones increases the chance for deletions, direct transfer from frozen stocks was preferred. The plates, containing the filters carrying the three parent cell lines, were incubated for 6 h at 30°C in the dark prior to transfer of the filters onto agar plates containing Sistrom's medium with 0.8 µg/ml tetracycline. These plates were incubated in a plexiglass box, which was illuminated using a 60 Watt incandescent bulb with an irradiance of 50-100 µE/s per m^2 at the surface of the plates. The box was continuously flushed with a gas mixture of 95% N_2 and 5% CO_2 (flow rate approximately 20 ml/min). The selective growth

conditions permitted only the growth of exconjugants (*sqdA* mutant cells containing cosmids), which became visible after about three to four days of incubation at the positions where the pronging device had touched the filter. These colonies were restreaked in patches onto Sistrom's media agar plates containing 0.8 μg/ml tetracycline. The sulfolipid phenotype of the exconjugants was tested by extraction of the lipids, followed by TLC analysis on ammonium sulfate-impregnated plates as described for the screening procedure used for the isolation of the sulfolipid mutants in chapter 3.

Isolation of small, complementing subclones from cosmids. Purified DNA from the complementing cosmid was partially digested with Sau3AI, and the fragments were separated by gel electrophoresis as described above for the preparation of the cosmid library. Fragments of the 2-3 kb, 3-4 kb, and 4-6 kb range were isolated and cloned separately into the BglII site of pLA2917, which was prepared as described above. Following ligation, the resulting plasmids (23-27 kb) were used to transform *E. coli* HB101. The cells were plated onto LB/tetracycline plates, and the resulting colonies were restreaked onto LB/tetracycline and LB/kanamycin plates to test for the presence of inserts (loss of kanamycin resistance). Clones containing inserts were transferred to microtiter plates, which were kept as frozen sublibrary stocks as described above for the cosmid library. Complementing clones were identified using the complementation assay described in the previous section.

Preparation of cosmid and plasmid DNA. Cosmid DNA was isolated on a large scale using the alkaline lysis method followed by purification on a

vertical CsCl gradient (standard protocol in reference 17). Preparations starting from 1 l cultures yielded about 200-400 μg DNA for low copy plasmids (pLA2917 and pCHB500 derivatives), and about 1-2 mg for multicopy plasmids (Bluescript derivatives) as determined by UV absorption.

Plasmid and cosmid minipreps were made using a rapid alkaline lysis method (26) with the exception that the samples were never vortexed, and that the DNA suspension was extracted once with 1 volume chloroform prior to precipitation of the DNA. Preparations of single copy plasmids were started from 3 ml overnight cultures, while preparations of multicopy plasmids were started from 1.5 ml overnight cultures. The DNA was resuspended in 30 μl 10 mM Tris-Cl, 1 mM EDTA (pH 8.0). For cosmids, the complete sample was digested with restriction enzymes and analyzed by gel electrophoresis. For smaller single-copy plasmids 1/3, and for multi-copy plasmids 1/10 of the DNA was analyzed in one reaction.

Restriction analysis and agarose gel electrophoresis. For detailed restriction maps, CsCl gradient-purified DNA was used. The DNA (1 μg) was digested in 20 μl of the recommended buffer for 1-2 h at the recommended temperature. Following the addition of loading dye (dye #3, reference 17) the samples were directly loaded into a well of an 0.8% (w/v) agarose gel.

DNA from minipreps was generally digested in 100 μl of the recommended buffer. Following digestion, the DNA was precipitated by adding 1/10 volume of 10 M LiCl and 2 volumes of cold ethanol. To remove salts, the samples were washed once with 200 μl 80% (v/v) ethanol. The dried

DNA was resuspended in an appropriate amount of 10 mM Tris-Cl, 1 mM EDTA (pH 8.0) and loading buffer prior to loading of the gel.

Gel electrophoresis was generally performed using 0.8% (w/v) agarose gels in combination with Tris-Borate-EDTA buffer as described in (17).

Southern analysis. DNA fragments separated on agarose gels were blotted onto Hybond-N^+ (Amersham) nylon membranes using the alkaline transfer method. Following soaking of the gel in 0.25 N HCl for 15 min and rinsing with water, the DNA was denatured in 1.5 M NaCl, O.5 N NaOH for 30 min and was then allowed to equilibrate in the transfer buffer (1.5 M NaCl, 0.25 N NaOH) for 15 min. The gel was placed into a plastic tray equipped with a snap lid, and the bottom of the tray was covered with transfer buffer. The filter, soaked in transfer buffer, was placed onto the gel and was covered with 3 layers of whatman 3MM paper. Paper towels, cut to size, were stacked onto the 3MM paper to a height such that upon closing of the lid a slight pressure was exerted onto the gel on the bottom of the tray. Following the transfer of the DNA for about 5 hours, the filter was prehybridized in 50 ml of 1 mM EDTA, 0.5 M NaH_2PO_4, 7% (w/v) SDS (pH 7.2) at 65°C for 15 min, prior to addition of the labeled probe in a small volume, just enough to cover the filter, of the same buffer inside tightly closed plastic boxes. The filters were incubated in the presence of the probe overnight at 65°C with shaking. To remove background label, it was usually sufficient to wash the filters 2X for 10 min at 65°C in 50-100 ml of 1 mM EDTA, 40 mM NaH_2PO_4, 5% (w/v) SDS (pH 7.2). On some occasions the filters were reprobed after removal of the label

by repeated boiling in water.

Probes were prepared by labelling DNA fragments excised from low-melting agarose gels using the random priming method (9).

Cloning of DNA fragments. DNA restriction fragments were usually isolated from agarose gels using NA45 paper strips (Schleicher & Schüll). The DNA fragment was electrophoresed into a small piece of NA45 paper placed into an incision in the gel in front of the band. The paper was rinsed with 10 mM Tris-Cl, 1 mM EDTA (pH 8.0) and incubated in 400 µl of 1.5 M NaCl at 70°C for 1 h to elute the DNA. Following the precipitation of the DNA by addition of 1 ml of cold ethanol, subsequent washing with 80% ethanol, and drying, the fragment was directly ligated without further purification to the prepared vector using a standard protocol (17). This method was efficient enough to yield the desired clones, following transformation of competent cells ($CaCl_2$ method, reference 17) and analysis of the transformants. Clones containing inserts were identified as colonies lacking ß-galactosidase activity on plates containing X-gal, prepared as described above.

Preparation of single stranded DNA for sequencing. *E. coli* DH5αF'cells containing derivatives of the bluescript plasmid were grown on LB/ampicillin plates at 37°C overnight. Three to five colonies were used to inoculate 25 ml of prewarmed 2X YT medium containing 100 µg/ml ampicillin (see above) in a 250 ml flask. The cells were grown for about 2 h at 37°C under vigorous shaking until a OD_{600} of about 0.1 was reached. At this stage, 20 µl of VCS phage (Stratagene; 5 x 10^{11} pfu/ml) were added, and incubation was

continued for about 5-6 h. Cells were removed by two consecutive centrifugations at 12,000 rpm for 10 min in a Sorvall SS34 rotor (Dupont). To 20 ml of the supernatant 5 ml of 20% (w/v) PEG^{800} , 2.5 M NaCl were added, and the samples were incubated for a minimum of 30 min at 4°C. To pellet the phage, the samples were centrifuged at 15,000 rpm in an SS34 rotor for 20 min. The supernatant was completely removed under suction followed by wiping the tube with a cotton tipped applicator. Following the resuspension of the phage pellet in 200 µl 10 mM Tris-Cl, 1 mM EDTA, 300 mM Na acetate (pH 8.0), the suspension was extracted twice with 1 volume of buffered phenol and 1 volume of chloroform. The DNA was precipitated by addition of 3 volumes of cold ethanol, and washed and dried as usual. The DNA yield was approximately 1-2 µg as estimated by agarose gel electrophoresis. This was sufficient for about 5-10 sequencing reactions.

Sequencing. DNA sequences were determined by employing the chain termination reaction method (21). In reaction mixtures, single stranded templates (about 200 ng per reaction) were sequenced using [α-^{35}S]-dATP as labelled substrate (6). To reduce sequencing artifacts due to the high GC content of *R. sphaeroides* DNA, GTP was replaced by 7-deaza-dGTP (3) provided in a sequencing kit purchased from United States Biochemical Corporation. Modified T7 DNA polymerase (Sequenase, reference 23) obtained along with the sequencing kit from the same supplier, was used as the sequencing enzyme. The reactions were done according to the instructions provided by the company. The labelled DNA fragments were

separated on 6% polyacrylamide, 8 M urea slab gels. Gels were double loaded and run according to reference (19).

Preparation of synthetic oligonucleotides. Oligonucleotides used as sequencing primers were synthesized on an Applied Biosystems 380A DNA synthesizer using the phosphoramidite method (4) according to the manufacturer's instructions. Following "trityl-on" synthesis, the oligonucleotides were cleaved from the support with concentrated NH_4OH and deprotected by incubation for 8 h at 55°C. The material was diluted with one third volume of H_2O and purified using a SepPac C_{18} cartridge (Waters) attached to a 5 ml syringe. Prior to loading of the sample, the cartridge was flushed with 5 ml acetonitrile followed by 5 ml of 2 M triethylamine acetate (pH 7.0). The flow rate was always kept at 1-2 drops per second. The sample was pushed twice through the cartridge, and the resulting flow through, still containing oligonucleotides, was stored at -20°C. The cartridge was flushed 3 X with 5 ml 10% ammonium hydroxide and 2 X with 5 ml H_2O. The oligonucleotide was detritylated with 2 X 1 min treatments of 5 ml 2% trifluoroacetic acid. Following flushing of the cartridge with 2 X 5 ml H_2O, the oligonucleotides were eluted with 2 ml 20% acetonitrile in water. The sample was evaporated to dryness, and the DNA was resuspended in 50-200 µl of 10 mM Tris-Cl, 1 mM EDTA (pH 8.0). The concentration was determined by measuring the absorption at 260 nm.

RESULTS AND DISCUSSION

Isolation of a cosmid complementing the *sqdA* mutant. A cosmid library from *R. sphaeroides* wild type DNA was constructed using the 21 kb cosmid vector pLA2917 (Figure 4-1). This vector contains the λ phage cos site, which is sufficient for the packaging of approximately 39-51 kb DNA fragments into λ phage particles by *in vitro* packaging extracts. The BglII cloning site is located in between the promotor and the coding region of the neomycin phosphotransferase gene located on pLA2917 (5). Upon insertion of a DNA fragment into this site, expression of the gene is prevented, allowing for the identification of insert-containing clones as kanamycin-sensitive colonies. Two libraries were prepared by infecting two different *E. coli* hosts, S17-1 (1728 clones) and HB101 (1920 clones). It was noted that upon infection with equal amounts of phage under comparable conditions, infection of HB101 consistently yielded about 4 X more colonies than S17-1. It was suspected that S17-1 discriminated against certain sequences as compared to HB101. HB101 is partially deficient in the methyl-cytosine specific restriction system (*mcrBC*), while the genotype of S17-1 with respect to this locus is not known. The presence of the methylation specific restriction system has been known to cause discrimination of methylated DNA during the construction of libraries (18). Therefore, it was concluded that the library in HB101 was more representative for the *R. sphaeroides* genome and should be used for the screening procedure. To test for the presence of inserts in this library, 768 clones were replica-plated onto LB/kanamycin plates, and 113 (14.7%) were found to be

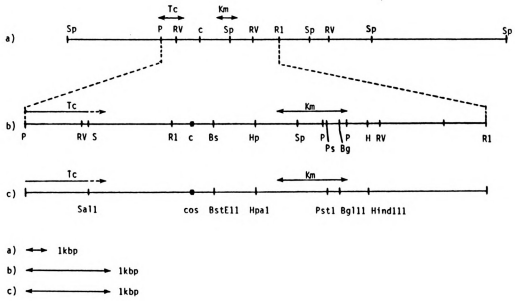


Figure 4-1: Detailed restriction map of pLA2917 (taken from ref. 1). The vector is shown linearized at one of the SphI sites. Abbreviations: Bg, BglII; Bs, BstEII; c, cos site; H, HindIII; Hp, HpaI; Km, kanamycin; Ps, PstI; P, PvuII; R1, EcoRI; RV, EcoRV; S, SalI; Sp, SphI; and Tc, tetracycline. (a) pLA2917; (b) expanded view of the region between the PvuII site in the tetracycline gene and one of the EcoRI sites. (c) same as (b) showing antibiotic resistance determinants, cos site, and unique restriction sites.

resistant, indicating that they might not contain inserts. Given the genome size of *R. sphaeroides* of 5 X 10^6bp (15) and the number of 1920 clones with an average insert size of 25 kb, the HB101 library was expected to cover the genome about four to five times. It was assumed that on average about one positive clone could be found in about 400 clones tested.

As opposed to the S17-1 library, which does not require a third helper strain for the transfer of cosmids into the *sqdA* mutant (22), cosmids contained in HB101 had to be transferred using the helper plasmid pRK2013 in MM294. Triparental matings were performed as described in the methods section. About 60-70 % of the cosmid clones were transferred (formation of exconjugants) in this particular experiment. Testing of about 450 exconjugants for restoration of the sulfolipid deficiency in the *sqdA* mutant by TLC of the lipids yielded two complementing cosmid clones, of which one (pCHB1701) was further analyzed.

Restriction analysis of cosmid pCHB1701. To establish a physical map for the complementing cosmid pCHB1701, cesium chloride gradient-purified DNA from this cosmid was cut with a series of restriction endonucleases. Simple restriction patterns were obtained for the restriction enzymes HindIII, EcoRV, BglII, EcoRI, and PstI, listed in order of increasing complexity. EcoRV had no site internal to the insert, while the other five enzymes cut the insert at least twice. By double digesting with all possible combinations for the five enzymes and applying Southern analysis to the resulting fragment patterns, it was possible to establish the restriction map for pCHB1701 as shown in

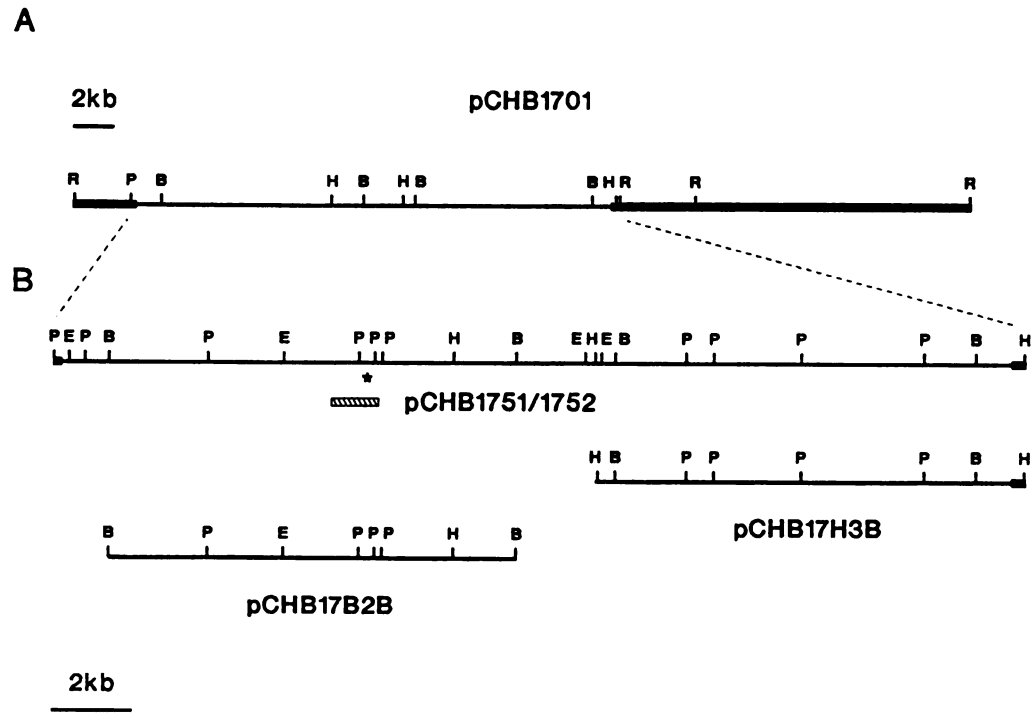


Figure 4-2: Complete restriction map for the cosmid clone pCHB1701. (A) overview showing insert and vector. (B) expanded view of the insert of pCHB1701, and the two subclones pCHB17B2B and pCHB17H3B. The location of the smallest subclones complementing the *sqdA* mutant (pCHB1751, pCHB1752) is indicated (cross hatched bar). The pLA2917 vector part was drawn with thicker line. Restriction sites: H, HindIII; R, EcoRV; E, EcoRI; B, BglII; and P, PstI. (*) location of the 0.4 PstI fragment, crosshybridizing with the comparable fragment from the subclones pCHB17011 and pCHB17012 .

Figure 4-2. One filter was prepared, which was probed with the complete pLA2917 vector, and stripped and reprobed with the internal 3.5 kb HindIII fragment. Use of the vector as a probe allowed the assignment of the vector/insert boarder fragments, as well as the vector internal fragments. Similarly, by using the HindIII fragment as a probe the central region of the cosmid could be mapped. This information was sufficient to establish a complete map for all the enzymes listed above except for PstI. To obtain the map for this enzyme, an approximately 10.0-kb BglII fragment from the left side of the cosmid, and a 10.5-kb HindIII fragment from the right side were subcloned into the Bluescript vector pBS-KS$^+$, giving rise to clones pCHB17B2B and pCHB17H3B respectively (Figure 4-2). Comparative restriction analysis of these two clones and pCHB1701 confirmed the map for HindIII, BglII, and EcoRI. In addition, the PstI map for the left half of the cosmid could be solved, and the number of PstI sites located on the 10.5-kb HindIII fragment from the right side of the cosmid were determined. However, to establish the order of the PstI sites located on the right side of the cosmid, the 10.5-kb HindIII fragment was excised from pCHB17H3B and partially digested with PstI. The separated fragments from the partial digest were blotted onto nitrocellulose filters and probed with the 0.35-kb BglII, HindIII fragment of pLA2917 (Figure 4-1). This fragment contained the right border fragment of the pLA2917 cosmid vector, which was present in the 10.5-kb HindIII fragment of pCHB17H3B. This strategy allowed the identification of the right end of the insert. Therefore, the sequential order of the PstI sites beginning from the right

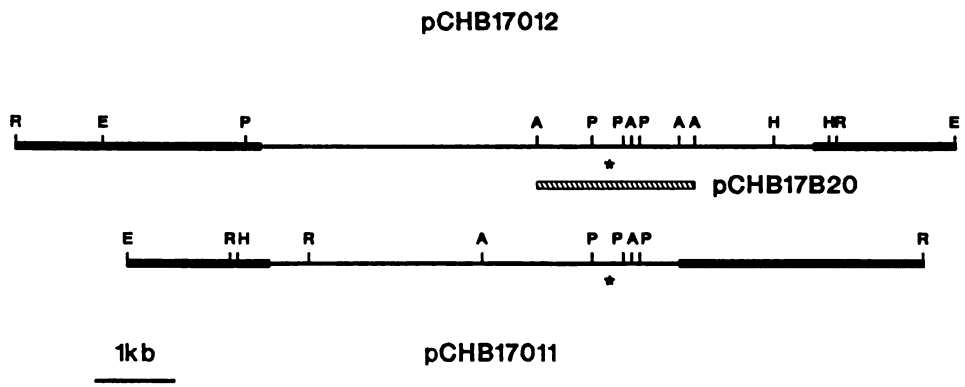


Figure 4-3: Complete restriction map of two subclones, derived from pCHB1701, which complement the *sqdA* mutant. The two clones pCHB17011 and pCHB17012 were aligned to match the common region around the 0.4 kb PstI fragment marked with (*). The location of the pCHB17012 BamHI fragment giving rise to the pCHB17B20 subclone series is indicated (cross hatched box). Vector sequences were drawn with thick line. Restriction sites: R, EcoRV; E, EcoRI; A, BamHI; P, PstI; H, HindIII.

end of the insert could be established. Summing up the size of each individual PstI fragment, the total length of the cosmid clone including the vector was estimated to be about 44.5 kb. Given the low accuracy to which the size of longer fragments could be determined on agarose gels, this value might have been underestimated by 1-2 kb. It was possible that sites closer than 0.2-0.3 kb outside of the sequenced subclone containing the *sqdA* gene (see below) were missed, as well as sites generating fragments of similar size below 0.5 kb.

Isolation and characterization of two subclones from pCHB1701 complementing the *sqdA* mutant. To isolate smaller subclones from the cosmid pCHB1701, three sublibraries with insert sizes of 2-3 kb, 3-4 kb, and 4-6 kb were prepared as described in the method section. Testing of 57 clones from the sublibrary containing the 2-3 kb inserts for complementation of the *sqdA* mutant yielded one positive clone (pCHB17011). In addition, one positive clone (pCHB17012) was found in the sublibrary (87 clones), containing 3-4 kb inserts. Restriction analysis of purified DNA from both clones indicated that pCHB17011 with its insert size of 5.2 kb as well as pCHB17012 (insert size 7 kb) obviously contained two independent inserts (Figure 4-3). They both shared a common region centering around a 0.4-kb PstI fragment. The 0.4-kb PstI fragment isolated from pCHB17012 was used as probe against restricted DNA from pCHB17011 and pCHB17012 (Figure 4-4), as well as against DNA from pCHB1701 (Figure 4-5). The Southern hybridization experiment confirmed that this fragment was common to all three clones, and

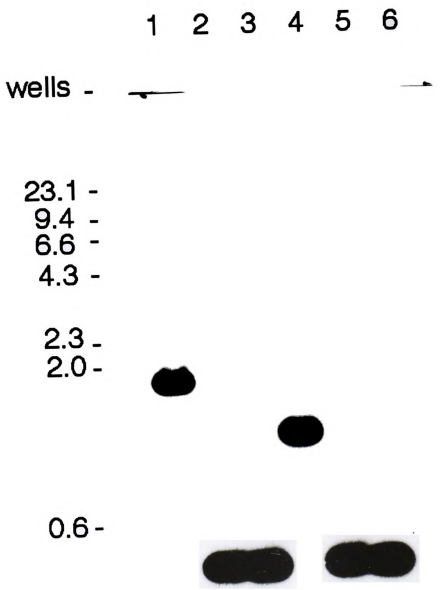


Figure 4-4: Southern blot of pCHB17011 and pCHB17012 probed with the 0.4 kb PstI fragment of pCHB17012. Lanes 1-3, pCHB17011; lanes 4-6, pCHB17012. The DNA was restricted with BamHI (lanes 1 and 4), PstI (lanes 2 and 5) and BamHI/PstI (lanes 3 and 6). The position of size markers (kb) is indicated on the left.

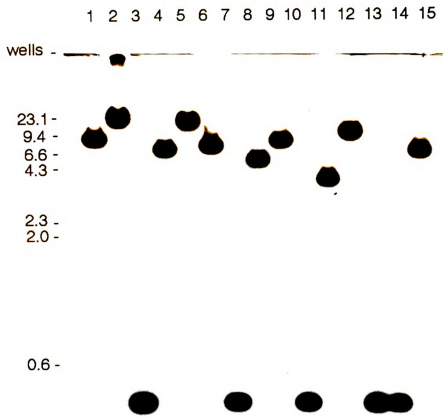


Figure 4-5: Southern blot of pCHB1701 probed with the 0.4 kb PstI fragment from pCHB17012. The DNA was restricted with BglII (1), HindIII (2), PstI (3), EcoRI (4), EcoRV (5), BglII/HindIII (6), BglII/PstI (7), BglII/EcoRI (8), BglII/EcoRV (9), HindIII/PstI (10), HindIII/EcoRI (11), HindIII/EcoRV (12), PstI/EcoRI (13), PstI/EcoRV (14), and EcoRI/EcoRV (15). The position of size markers (kb) is indicated on the left.

established the location of the 0.4-kb PstI fragment on the map of cosmid pCHB1701 (Figure 4-2). The fusion of the two inserts in both clones was assumed to occur between the left BamHI site of pCHB17012 and the neighboring PstI site to the right (Figure 4-3). A 2.1-kb BamHI fragment obtained in a partial digest from pCHB17012 (Figure 4-3) was cloned in both orientations into the BamHI site of pBS-KS$^+$ (pCHB17B20-A$^+$, -D$^+$), and into pBS-KS$^-$ (pCHB17B20-A$^-$, same orientation as pCHB17B20-D$^+$). Based on the information obtained from comparative restriction analysis and Southern hybridization, this fragment was thought to contain the stretch of DNA common to both subclones, and being responsible for the complementation of the *sqdA* mutant.

Sequence analysis of pCHB17B20. The various pCHB17B20 Bluescript clones were used to prepare single stranded templates for each strand. Using the universal M13 primers, initial sequencing runs were done from both ends of the complete clone as well as from a small number of deletion clones created by using internal restriction sites. Based on the sequence information obtained from these runs, oligonucleotides were synthesized and used as sequencing primers (Table 4-3). The bulk of the sequence was obtained by using these synthetic oligonucleotide primers (Figure 4-6).

The sequence (Figure 4-7), confirmed by sequencing of both strands, had a GC-content of 70%, which is typical for *R. sphaeroides* DNA (15). Analyzing the sequence for open reading frames initially revealed three larger frames (Figure 4-6), which could encode proteins of about 33 kD (ORF1),

Table 4-3: List of synthetic oligonucleotide primers used for sequencing of pCHB17B20.

Primer	Sequence
CB24	5' - CGT, GGA, ACA, GAG, TGA, TGT -3'
CB25	5' - GAT, CGA, GAT, GCT, GCG, CGA -3'
CB26	5' - GCC, AAC, ATG, GAA, GGC, GAA -3'
CB28	5' - GGC, CTC, CTG, CTC, TCG, CAT -3'
CB29	5' - TGC, CAC, GGC, GTG, CGG, CGA -3'
CB30	5' - GAT, CCA, TGC, GTG, CGC, ACT -3'
CB31	5' - GGC, CAT, TTT, CCT, CAT, GCT -3'
CB32	5' - CAC, AAG, GTG, CCC, TAC, CAG -3'
CB33	5' - ACA, CGG, CTC, ATG, TCA, TGC -3'
CB34	5' - CTC, CAC, CAC, ATC, CTC, GCG -3'
CB35	5' - AAG, TTC, GAC, CTG, CCT, GTG -3'
CB36	5' - GTG, CAA, GAG, CAT, GAC, GAA -3'
CB38	5' - TGA, AGG, ACG, ATC, CGA, CGC -3'

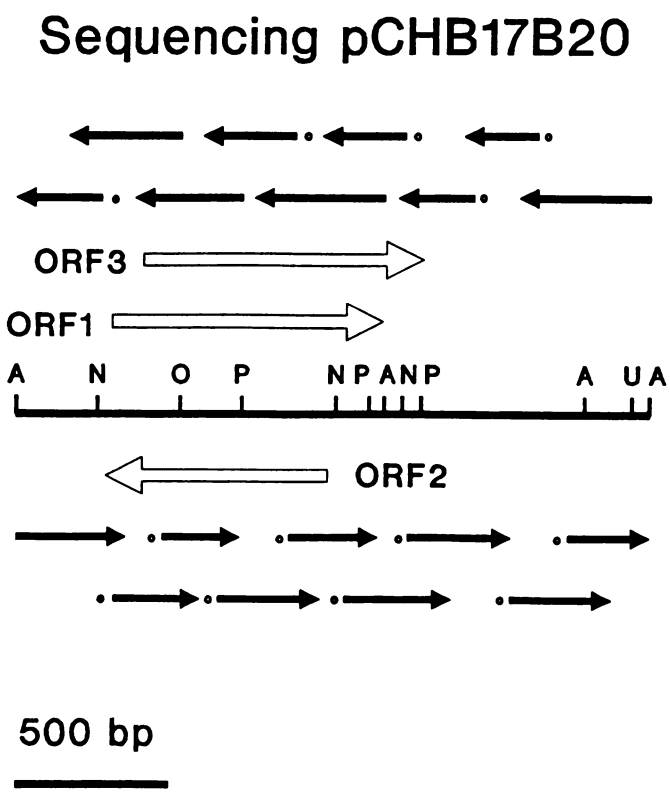


Figure 4-6: Strategy for sequencing pCHB17B20. Only the 2.1 kb BamHI fragment from pCHB17012 is shown. Sequencing runs (closed arrows) primed with synthetic oligonucleotides are preceded by an open circle. The location and orientation of three open reading frames (ORF1, ORF2, and ORF3) are indicated by open arrows. Restriction sites: A, BamHI; N, NaeI; O, XhoI; P, PstI; U, StuI.

Figure 4-7: DNA sequence of pCHB17B20. Only the proposed coding strand for the *sqdA* gene is shown. The protein sequence predicted by translation of the proposed *sqdA* coding sequence, is given below the sequence. The locations of stop codons (S^{1-2}) in frame with the coding sequence, a potential ribosome binding site (SD), 4 possible initiation codons (I^{1-4}), and restriction sites, used to construct deletion clones are marked above the sequence by a line.

```
                                                                  S1
1     GGATCCGGCTCGACCCGCTGAACCCGCAGACGGCGGCGCGCATGTCGACCGCCTTCGAGA

                                                                  S1
61    CCCGGGCCCGCTACGATGCGGACCGGCAGGGCATGATACTGGCCGAACTCGACCGCATCC

                                                                  S1
121   TCGCCACGCCGGGCCTCAGCCGCGATCTGGGCGAGATGGCGGGGCGGATGAGGCACGCAT

                                                                  S1
181   GAGCGGGTGGGAACAGTTGGACGGGAAGCGGCGTTCTGACCGGCGGATGACTGCCGAGCG

                                        NaeI                      I1
241   CAGATCGGTGCAAAGGACGGCGCTTGCCGGCCGGGACAACCGGGGCTCATGTCATGCGAG

          SD          I2          I3 I4
301   AGCAGGAGGCCCCGGGTGTCGCTCATGATGCAGTCTCGAATCGATCCGCTCATCGAAGAG
1                     M  S  L  M  M  Q  S  R  I  D  P  L  I  E  E

361   CGTGCCCCCTGGCTCTTCTCCGGCCGTCCCCATCATGAAGCCTCGCGCAGACTTCTGTGC
16     R  A  P  W  L  F  S  G  R  P  H  H  E  A  S  R  R  L  L  C

421   CGCGTTCTGGGCTACGACCGCACCGTCGAGATCGGAGAGCAGCTGAAGGACGATCCGACG
36     R  V  L  G  Y  D  R  T  V  E  I  G  E  Q  L  K  D  D  P  T

481   CCCGAGATCATGCGCCGCGTGGCGCAGATGATCGCGCGGGACGTAGAGGTCTCGGGTCTC
56     P  E  I  M  R  R  V  A  Q  M  I  A  R  D  V  E  V  S  G  L

541   GAGAACCTGCCGCGCAACGGCCCGGCGCTGGTCGTCTGCAACCATCCGACCGGCATCGCC
76     E  N  L  P  R  N  G  P  A  L  V  V  C  N  H  P  T  G  I  A

601   GACGGGGTGATCCTCCACCACATCCTCGCGCCCCTGCGGCCCGATCTCTTCTTCTACGCC
95     D  G  V  I  L  H  H  I  L  A  P  L  R  P  D  L  F  F  Y  A

661   AACCGCGACATTCTTCGCGTGCTGCCGCAGATGGAGGACCTGATCTGCCCGGTCGAATGG
116    N  R  D  I  L  R  V  L  P  Q  M  E  D  L  I  C  P  V  E  W

721   CGGCAGGAGAAGCGGAGCCTGCAGAAGACGCGCGAGACGATGGACTACACCCGCCGCGCC
136    R  Q  E  K  R  S  L  Q  K  T  R  E  T  M  D  Y  T  R  R  A

781   ATCGAGGCCGGGCGGCTGGGCGTAATCTTCCCCTCCGGGCGGCTGGCCAAGCGGACGGGA
156    I  E  A  G  R  L  G  V  I  F  P  S  G  R  L  A  K  R  T  G

841   CTGAAGCTCATGGAGCGGCCGTGGATGGCCTCCGCTGCGATGATCGCGCGGAAGTTCGAC
176    L  K  L  M  E  R  P  W  M  A  S  A  A  M  I  A  R  K  F  D

901   CTGCCTGTGGCGCCGGTGAACATCCGGGCGCGCAACTCGACGCTGTTCTACCTCTTCGAC
196    L  P  V  A  P  V  N  I  R  A  R  N  S  T  L  F  Y  L  F  D
```

Figure 4-7 (cont'd)

```
961   CTGATCCACCCCACGCTGCGGGACATCACTCTGTTCCACGAGGTGCTGAACAAGCATCGC
216    L  I  H  P  T  L  R  D  I  T  L  F  H  E  V  L  N  K  H  R

1021  CAGCCGTTCCGACTGACGATGGGCCGGCCGATCCCGGCCGAGCGTCTGCCGTCGAAGTCC
236    Q  P  F  R  L  T  M  G  R  P  I  P  A  E  R  L  P  S  K  S

1081  GAAGAGGGGATCGAGATGCTGCGCGAGGCGACGCTCGCGCTCGGCGGGCAATATTCGCCT
256    E  E  G  I  E  M  L  R  E  A  T  L  A  L  G  G  Q  Y  S  P
                                                              S²
1141  GCCGTCAGCATCCTGCAGATGACGAGCCTGCGCGGCGCGCTCCGCTACGCACGCGGGTAG
276    A  V  S  I  L  Q  M  T  S  L  R  G  A  L  R  Y  A  R  G
      BamHI
1201  GGGATCCCGGAGCGCGCGGCTGGCCGGACAGGCGGCCGCGCCTCCTCCCCACCCCTCCGC

1261  GGCCGGCAGGAGCTGTCGCCGCTGCGTTCCGGCAGCCGCGGGCGGCCATTTTCCTCATGC

1321  TCCTGCAGTGACCGGAGGACCCGTCGCCGCCGCGTCCAACGGCAGGCGCAGGACGCCCCC

1381  TTCGTCATGCTCTTGCACGGCGGCGCGGGATGGCGTCAACTCGCGCCATGACCGGACCGC

1441  GCTACACCCCGCTCGTCGAGAGCCTGCCCTCCACCGTTCCCTTCGTCGGCCCCGAGACGC

1501  AGGAGCGTGCGCGTGGCCGCCCGTTCCGCGCCCGCTTGGGCGCGAACGAGAGCCTCTTCG

1561  GCCCCTCGCCGCACGCCGTGGCAGCCATGGCCGCGGCCGCCGCCGAGGTCTCGCTGTATG

1621  GCGATCCAGAGATCCATGCGCTGCGCACTGCCATCGCGGCCCATCACGGCGTGGCGCCGG

1681  ATCAGGTGACGGTGGGCGAGGGCGTCGACGGGCTTCTGGGCAACCTCGTCCGGCTGCTTG

1741  TCGACGAGGGCGCGCCGGTCGTGACCTCGGAGGGCGCCTATCCGACCTTCGCCTTCCATG

1801  TGGCGGGGTTCGGCGGCCGCCTGCACAAGGTGCCCTACCAGGGCGATCACGAGGATCCCG

1861  AGGCGCTGGCGGCCGAGGCCCGGAGGTCGGGGCCCGGATCGTCTATTTCGCCAATCCGGA

1921  CAATCCGATGGGGTCGCACCATCCTGCCGCCTCCGTCCGCGCCCTGATCGAGGCGCTGCC
                                  StuI
1981  CGAGGGCACGCTGCTCGCGCTCGACGAGGCCTATGTGGAACTGGCCCCCGAGGGCACGGC

2041  GCCCGAGATCGCCCCCGAGGATCC
```

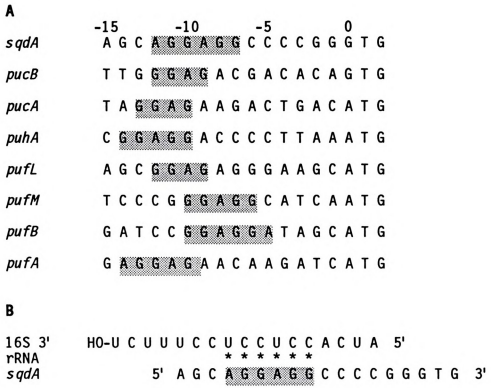


Figure 4-8: Comparison of the putative *sqdA* ribosome binding site to binding sites from various *R. sphaeroides* genes (A) and to the 3'-end of the 16S rRNA from *R. sphaeroides* (B). Bases matching the complementary sequence of the 16S rRNA 3'-end representing the ribosome binding site are shaded. Sequences were taken from: *pucB* and *pucA* (14), *puhA* (8), *pufL* (25), *pufM* (24), *pufB* and *pufA* (13), and *R. sphaeroides* 16S rRNA 3'-end (11).

26 kD (ORF2) and 32 kD (ORF3). However, complementation analysis of deletion clones, as described below, excluded two of the three frames (ORF2 and ORF3). The maximum size of the third reading frame (ORF1) was found to be determined by the position of two in-frame stop codons (Figure 4-7, $S^{1,2}$). Downstream of the first stop codon, four possible in-frame initiation codons were located, of which only the second, a GTG codon (Figure 4-7, I^2), was preceded by a putative ribosome binding site. This putative ribosome binding site showed perfect sequence match complementary to the 3' end of the *R. sphaeroides* 16S rRNA, as well as correct spacing as compared to ribosome binding sites from other *R. sphaeroides* genes (Figure 4-8). It has been shown that the presence of a ribosome binding site consensus sequence, correctly spaced upstream of the initiation codon, can greatly affect the translation of the respective mRNA (12). The presence of a sequence satisfying the requirements for a correctly spaced ribosome binding site only upstream of the second initiation codon, suggested that a GTG codon 7 bp into the 3'-direction was the best candidate for the translational initiation site of the *sqdA* gene. Although the usual initiation codon is ATG, the use of GTG has been observed previously. One example for a GTG initiation codon found in *R. sphaeroides* is the *pucB* gene (Figure 4-8).

Complementation analysis using various DNA fragments of pCHB17B20 cloned into an expression vector. To directly determine, which of the open reading frames in the 2.1-kb BamHI fragment of pCHB17012 represented the *sqdA* gene, the complete insert as well as deletion fragments of pCHB17B20

were cloned in both orientations behind a promotor to create transcriptional fusions, which would result in transcription of the various fragments in *R. sphaeroides*. The clone pCHB17B20 by itself could not be employed for complementation analysis, since Bluescript derivatives can not replicate in *R. sphaeroides*. Instead, the vector used was pCHB500, which was constructed by inserting the promotor region of the cytochrome-c_2 gene (*cycA*) from *Rhodobacter capsulata* into the poly-linker of pRK415 (construction and testing of pCHB500 are described in Appendix B). Insertion of the complete fragment of pCHB17B20, corresponding to the 2.1-kb BamHI fragment of pCHB17012, into the BamHI site of pCHB500 in both orientations, gave rise to clones pCHB1741 and pCHB1742 (Figure 4-9), both of which restored the wild type phenotype when present in the sulfolipid deficient *sqdA* mutant. This result confirmed that the 2.1 kb fragment of pCHB17012 was indeed capable of complementing the *sqdA* mutant. In addition, these data suggested that the fragment contained a promotor, making the complementation independent of the insert orientation and therefore independent of the presence of the external *cycA* promotor provided by pCHB500. There were two possible origins for a promotor present in the 2.1 kb BamHI fragment of pCHB17012: first, the endogenous promotor naturally occurring in front of the gene was located on this fragment; or second, the *sqdA* open reading frame was part of an operon, and the promotor necessary for complementation was provided by the second inserts found in the complementing subclones pCHB17011 and pCHB17012 (Figure 4-3). The fact that the only two positive subclones found, both

contained two inserts, supported the second possibility.

A 1.2 kb BamHI subfragment of pCHB17B20 extending from the left end e03of the insert to the BamHI site directly downstream of the stop codon of the putative *sqdA* gene (Figure 4-7), was cloned in both orientations into the BamHI site of pCHB500 giving rise to subclones pCHB1751 and pCHB1752 (Figure 4-9). Both clones transferred into the *sqdA* mutant were able to restore the wild type phenotype, suggesting that the 1.2 kb BamHI subfragment of pCHB17B20 contained a promotor preceding the *sqdA* coding sequence. These two subclones were the smallest clones capable of complementing the *sqdA* mutant. The location of this fragment on the cosmid clone pCHB1701 is shown in Figure 4-2. The 3'-end of one of the three potential open reading frames ((Figure 4-9, ORF3) was not included in this fragment, making it unlikely to represent the *sqdA* gene. However, there were still two intact reading frames (ORF1, and ORF2; Figure 4-9) located on opposite strands of this fragment.

An experiment was designed to determine the coding strand for the *sqdA* gene, and thereby the correct open reading frame by testing the orientation of the insert with respect to the *cycA* promotor, necessary for complementation. To observe orientation dependent complementation using the vector pCHB500, the *sqdA* coding sequence lacking a promotor had to be cloned in proper orientation behind the *cycA* promotor provided by pCHB500. A NaeI site upstream of the putative initiation codon of ORF1 was used to remove most of the upstream sequence of the putative *sqdA* gene (Figure 4-7). Using this site, a 1.8-kb partial NaeI/StuI fragment from pCHB17B20 was cloned

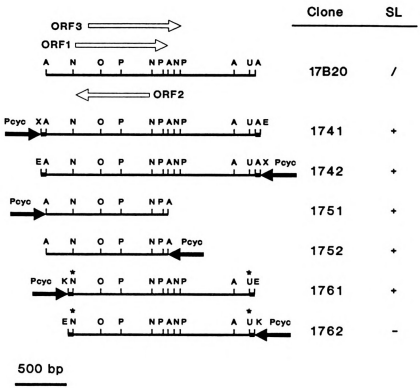


Figure 4-9: Complementation analysis using derivatives of pCHB17B20 cloned into the expression vector pCHB500. The 2.1 kb BamHI insert of pCHB17B20 is shown at the top with the orientation and location of three open reading frames (ORF1, ORF2, and ORF3) indicated by open arrows. For each clone shown, three independent constructs were prepared, which gave identical results in complementation assays. SL: presence (+) or absence of the sulfolipid in lipid extracts of the sulfolipid deficient *sqdA* mutant containing the indicated clone (clones are listed without the pCHB prefix); pCHB17B20 can not be tested (/). Pcyc: *cycA* promotor (filled arrows), the direction of transcription is indicated by the arrow head. Small stretches of Bluescript poly-linker region flanking the inserts are indicated with thicker line. Restriction sites: A, BamHI; K, KpnI; N, NaeI; O, XhoI; P, PstI; U, StuI; X, XbaI. Locations of NaeI and StuI sites, no longer cleavable due to the cloning process, are marked with an asterix (*).

in both orientations into the EcoRV site of the Bluescript vector. From there, the fragment was cloned in both orientations into the pCHB500 poly-linker using KpnI and EcoRI sites, giving rise to pCHB1761 and pCHB1762 (Figure 4-9). Due to the cloning procedure, the NaeI/StuI fragment was flanked by small stretches of Bluescript poly-linker sequence in pCHB500. In a complementation assay, only pCHB1761, as opposed to pCHB1762, was able to restore the wild type phenotype. In pCHB1761 the *cycA* promotor was located upstream of ORF1 reading in the direction of the putative *sqdA* gene (shown in Figure 4-7), while in pCHB1762 the situation was reversed. Therefore, only ORF1 but not ORF2 could represent the *sqdA* gene.

The complementation analysis of the pCHB17B20 derivatives, cloned into the expression vector pCHB500 as described above, clearly supported the findings of the sequence analysis data, and localized the *sqdA* coding sequence. The orientation dependence of complementation shown by clones pCHB1761 and pCHB1762 indicated the direction of gene expression, which coincided with the orientation of the putative *sqdA* coding sequence shown in Figure 4-7. The results of the deletion and complementation experiments excluded two alternative open reading frames as possible candidates for the coding sequence of the *sqdA* gene.

Amino acid sequence analysis of the putative *sqdA* gene product. The putative *sqdA* coding sequence, as suggested by the results discussed above, was translated into the corresponding amino acid sequence (Figure 4-7) consisting of 295 amino acids. A molecular weight of 33.6 kD could be

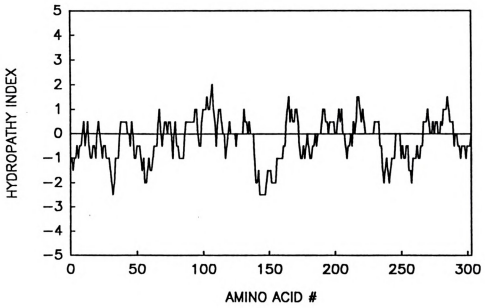


Figure 4-10: Hydropathy plot of the predicted *sqdA* amino acid sequence. Amino acids (304) from the start of the first initiation codon to the second stop codon (see Figure 4-7) were included in the analysis. The analysis was done according to Kyte and Doolittle (16). The segment size used was 10 amino acids.

calculated for the resulting protein. A hydropathy analysis was done on the predicted protein according to the Kyte and Doolittle method (16) to determine whether it might be membrane bound or soluble. Every amino acid is given a value according to its hydrophobicity. Using a moving-segment approach the average hydropathy within a given segment of predetermined length, usually between 7 and 11 amino acids, is determined along the amino acid sequence, and the consecutive scores are plotted along the sequence. A midpoint line is indicated that corresponds to the grand average of the hydropathy of the amino acid composition found in most of the sequenced proteins. Regions of the sequence scoring above the mid point line are hydrophobic, while those regions scoring below the midpoint line are hydrophilic. The analysis result for the putative *sqdA* gene product is given in Figure 4-10. The lack of extensive regions of continuous hydrophobicity, which are indicative of the membrane spanning regions of membrane proteins, suggested that the protein is soluble, rather than membrane bound.

To investigate the possible function of the predicted *sqdA* protein, the amino acid sequence was used to screen two protein data banks for homologous proteins of known function. No known protein showed sufficient homology to the putative *sqdA* protein, to draw any conclusions.

CONCLUSIONS

Two cosmids complementing the sulfolipid deficient *sqdA* mutant were isolated from a *R. sphaeroides* wild type DNA cosmid library. Complementation

was also observed with two smaller subclones, derived from one of the complementing cosmids. A fragment covering a region common to both of the two clones was sequenced. Analysis of the smallest fragment capable of complementing the *sqdA* mutation permitted the identification of a single open reading frame. The maximum size of the putative *sqdA* coding sequence was found to be determined by two flanking in-frame stop codons. Of the four possible initiation codons, only one was preceded by a potential ribosome binding site, satisfying the required criteria concerning complementarity to the 3'-end of the 16S rRNA as well as correct spacing. The respective GTG codon was the only putative initiation codon of any of the observed reading frames, preceded by a perfect ribosome binding site. Based on this circumstantial evidence, the protein sequence outlined in Figure 4-7 was proposed to be the *sqdA* gene product responsible for the complementation of the sulfolipid deficient *sqdA* mutant of *R. sphaeroides*. However, it should be pointed out that only purification of the protein from wild type *R.sphaeroides* cells, followed by sequencing of the N-terminus of the protein can clarify the exact location of the initiation codon.

The negative result obtained from screening protein data banks for homology with the putative *sqdA* gene product indicated that the *sqdA* gene encodes a novel protein involved at one, as yet, unknown level in sulfolipid biosynthesis. It is expected that the analysis of the ^{35}S-labelled, water-soluble compounds found to accumulate in the *sqdA* mutant (see chapter 3) will provide some clues about the function of *sqdA* gene product. With the *sqdA*

gene in hand isogenic lines can be constructed for comparative biochemical and physiological experiments. The *sqdA* coding sequence can be expressed using *E. coli* expression vectors, and antibodies raised against the product will provide valuable tools to study the *sqdA* gene product.

In summary, based on the data provided in this chapter, experiments can be designed to further confirm the identity of the *sqdA* gene as well as to elucidate the function of the *sqdA* gene product. It is expected that these experiments will provide an answer to the remaining questions about the exact role of the *sqdA* gene in sulfolipid biosynthesis in *R. sphaeroides* in the near future.

REFERENCES

1. Allen L.N., Hanson R.S. (1985). Construction of broad-host-range cosmid cloning vectors: Identification of genes necessary for growth of *Methylobacterium organophilum* on methanol. J. Bacteriol. 161:955-962

2. Bachmann B.J. (1987) In *Eschericha coli* and *Salmonella typhimurium*, Cellular and Molecular Biology. Ed. Neidhardt F.C. et al. ASM, pp 1190-1219

3. Barr P.J., Thayer R.M., Laybourn P., Najarian R.C., Seela F., Tolan D.R. (1986). 7-Deaza-2'-Deoxyguanosine-5'-Triphosphate: Enhanced Resolution in M13 Dideoxy Sequencing. Bio Techniques 4:428-432

4. Beaucage S.L., Caruthers, M.H. (1981). Deoxynucleoside Phosphoramidites - A new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett. 22:1859-1862

5. Beck E., Ludwig G., Auerswald E.A., Reiss B., Schaller H. (1982). Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene 19:327-336

6. Biggin, M.D., Gibson T.J., Hong G.F. (1983). Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA 80:3963-3965

7. Carle G.F, Olson M.V. (1984). Separation of DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. Nuc. Acid Res. 12:5647-5664

8. Donohue T.J., Hoger J.H., Kaplan S. (1986). Cloning and expression of the *R. sphaeroides* reaction center H gene. J. Bacteriol. 168:953-961

9. Feinberg A.P., Vogelstein B. (1984). Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137: 266-267

10. Figurski D., Helinski D.R.(1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function in *trans*. PNAS 76:1648-1652

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

12. Gold L., Pribnow D., Schneider T., Shinedling S., Swebilius Singer B., Stormo G. (1981). Translational initiation in prokaryotes. Ann. Rev. Microbiol. 35:365-403

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

14. Kiley P.J., Kaplan S. (1987). Cloning, DNA sequence, and expression of the *R. sphaeroides* light-harvesting B800-850-α and B800-850-ß genes. J. Bacteriol. 169:3268-3275

15. Kiley P.J., Kaplan S. (1988). Molecular genetics of photosynthetic membrane biosynthesis in *R. sphaeroides*. Microbiol. Rev. 52:50-69

16. Kyte J., Doolittle R. F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

17. Maniatis T., Fritsch E.F., Sambrook J. (1982). Molecular Cloning, A Laboratory Manual. Cold Spring Harbour Laboratory, New York.

18. New England BioLabs. Restriction of foreign DNA by *E. coli*. Catalog (1990-1991), pp 143-144

19. Parkinson C., Cheng S.Y. (1989). A convenient method to increase the number of readable bases in DNA sequencing. Bio Techniques 7:828-829

20. Raleigh E.A, Lech K., Brent R. (1989). In: Current Protocols in Molecular Biology. Eds. Ausubel F.M. et al. Publishing Associates and Wiley Intersience; New York, Unit 1.4

21. Sanger F., Nicklen S., Coulson A.R. (1977). DNA sequencing with chain termination inhibitors. Proc. natl. Acad. Sci. USA 74:5463-5467

22. Simon R., Priefer U., Puehler A. (1983). A broad host range mobilization system for *in vivo* engineering: transposon mutagenesis in gram negative bacteria. Bio/Technology 1:37-45

23. Tabor S., Richardson C.C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA 84:4767-4774

24. Williams J.C., Steiner L.A., Feher G., Simon M.I. (1984). Primary structure of the L subunit of the reaction center from *R. sphaeroides*. Proc. Natl. Acad. Sci. USA 81:7303-7307

25. Williams J.C., Steiner L.A., Ogden R.C., Simon M.I., Feher G. (1983). Primary structure of the M subunit of the reaction center from *R. sphaeroides*. Proc. Natl. Acad. Sci. USA 80:6505-6509

26. Zhou, C., Yang Y., Jong A.Y. (1990). Mini-prep in 10 minutes. Bio Techniques 8:172-173

CHAPTER 5

ISOLATION AND CHARACTERIZATION OF A CLONE COMPLEMENTING THE SULFOLIPID-DEFICIENT *sqdB*, AND *sqdC* MUTANTS OF *R. SPHAEROIDES*

ABSTRACT

Using a complementation assay, a cosmid (pCHB1601) was isolated which was able to restore the wild type phenotype when present in the *sqdB*, and *sqdC* mutants, but not when present in the *sqdA* mutant. A detailed restriction map for this clone was prepared. From this cosmid a 2.4 kb subfragment containing clone (pCHB16014) was isolated which complemented the *sqdB* mutant. Sequence analysis revealed an open reading frame encoding a putative protein of 404 amino acids. The N-terminus of this putative protein showed striking homology to the N-terminal amino acid sequence of UDP-glucose epimerase, encoded by the bacterial *galE* gene and the *gal10* gene of yeast. Further deletion and complementation analysis of the pCHB16014 subclone, as well as insertion of a promotor-less chloramphenicol-acetyl-transferase gene into the clone to determine the predominant direction of expression, supported the idea that the observed open reading frame might be the coding sequence of the *sqdB* gene.

A second subclone (pCHB16018) was isolated from the cosmid pCHB1601, which complemented the *sqdC* mutant. Restriction and Southern analysis revealed that this clone was different from the subclone complementing the *sqdB* mutant. Cloning of a restriction fragment from the

cosmid pCHB1601, which contained sequences common to the pCHB16018 subclone, into an expression vector and subsequent complementation analysis, revealed the location and orientation of the putative *sqdC* gene. It was proposed that the *sqdC* gene is located downstream of the *sqdB* gene and is transcribed in the same direction. This raises the possibility that *sqdB* and *sqdC* are two genes involved in sulfolipid biosynthesis, which might be organized in an operon.

INTRODUCTION

In the previous chapter the isolation and analysis of a cosmid clone (pCHB1701) was described, which complemented the *sqdA* mutant. For the same reasons described in the introduction to the previous chapter, it was desired to isolate the genes complementing the *sqdB* and *sqdC* mutants. It was hoped that detailed physical mapping of the DNA sequences complementing all three mutants might reveal that they affect three different genes involved in sulfolipid biosynthesis. In addition, it was assumed that the complementing clones for the mutants will provide the basis for the construction of isogenic lines, which are important for further biochemical and physiological analysis.

In this chapter, the analysis of a cosmid (pCHB1601) complementing the *sqdB* and *sqdC* mutants is described. The strategy employed was, in principle, the same as for the identification of the *sqdA* gene. However, the detailed analysis of the putative *sqdB* gene was complicated by the fact that

relatively few suitable restriction sites for the construction of deletion clones were present, and that stop codons rarely occurred in some of the reading frames, making it difficult to identify the coding sequence of the putative *sqdB* gene. The location and orientation for the *sqdC* gene with respect to the *sqdB* gene were determined, but attempts to find the putative coding sequence were not made during the course of this study.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains used for the experiments described in this chapter were the same as those described in Table 4-1 of the previous chapter, except that Bluescript plasmids were propagated in DH5α, and that for the preparation of single stranded DNA, DH5αF' was replaced by XL1-Blue ([F'::TN10 proA$^+$B$^+$ lacI^q Δ(lacZ)M15] recA1 endA1 gyrA96(nal^r) thi hsdR17 (r_K-m_K+) supE44 relA1 lac; reference 3).

In addition to the vectors pBS-KS$^+$, pBS-KS$^-$, pLA2917, pRK2013, and pCHB500 described in Table 4-2 in the previous chapter, the plasmids listed in Table 5-1 were used in the experiments described in this chapter.

Growth and media. Cultures of *R. sphaeroides* were grown as described in the method section of chapters 3 and 4. Growth conditions and media for *E. coli* were described in chapter 4.

Triparental mating and complementation assay. Isolation of the clones complementing the *sqdB* and *sqdC* mutants was done using triparental matings, followed by analysis of lipid extracts obtained from the exconjugants

Table 5-1: Plasmids used in the experiments described in chapter 5.

Plasmid	Description or Construction	Source or Reference
pBR328	pBR322 derivative containing the CAT gene	(1)
pCHB1601	cosmid clone complementing the *sqdB* and *sqdC* mutants	this study Figure 5-2
pCHB1611	pCHB1601 digested with HindIII and recircularized	this study Figure 5-3
pCHB1621	11 kb HindIII fragment of pCHB1621 in HindIII of pLA2917	this study Figure 5-3
pCHB16014	2.4 kb fragment of pCHB1601 in pLA2917, complementing the *sqdB* mutant	this study Figure 5-7
pCHB100, pCHB200	partial 2.9 kb PstI, HindIII fragment from pCHB16014 containing the insert and flanking pLA2917 sequences in the PstI, HindIII sites of pBS-KS$^+$ an -KS$^-$	this study
pCHB16014D-B	pCHB16014 cut with BglII and recircularized	this study Figure 5-7
pCHB16014D-AB	pCHB16014 cut with BamHI and BglII, and recircularized	this study Figure 5-7
pCHB16201, pCHB16202	2 kb SalI, EcoRI partial fragment of pCHB16014 first cloned into the SalI, EcoRI, sites of pBS-KS$^-$; excised from there with KpnI and EcoRI, and cloned into the KpnI, EcoRI sites of pCHB500 to give pCHB16201; or excised with KpnI and XbaI, and cloned into the KpnI, XbaI sites of pCHB500 to give pCHB16202 (opposite orientation)	this study Figure 5-7 this study Figure 5-7

Table 5-1 (cont'd)

pCHB16151, pCHB16152	1.5 kb SalI, EcoRI fragment of pCHB16014 first cloned into the SalI, EcoRI sites of pBS-KS$^-$; excised from there with KpnI and EcoRI, and cloned into the KpnI, EcoRI sites of pCHB500 to give pCHB16151; or excised with KpnI and XbaI, and cloned into the KpnI, XbaI sites of pCHB500 to give pCHB16152 (opposite orientation)	this study Figure 5-7
pCHB16014CAT-1, pCHB16014CAT-2	1 kb Sau3AI fragment of pBR328 pBR328, containing the CAT gene coding sequence, cloned in both orientations into the BamHI site of pCHB16014	this study Figure 5-8
pCHB16018	9.5 kb subclone of pCHB1601 complementing the *sqdC* mutant	this study Figure 5-11
pCHB181, pCHB182	4.2 kb BglII fragment of pCHB1601 cloned in both orientations into the BamHI site of pCHB500	this study Figure 5-11

as described in the method section of chapter 4.

DNA manipulations and analysis. The construction of small subclones by partial Sau3AI digestion from the complementing cosmid, as well as general methods for DNA manipulations such as preparation of plasmid DNA, cloning, restriction analysis, and Southern analysis were done following the protocols outlined in detail in chapter 4.

Sequence analysis. Single stranded + and - strands for sequencing were prepared from the Bluescript clones pCHB100 and pCHB200 respectively and their deletion derivatives as described in chapter 4, except that XL1-Blue was used as host strain instead of DH5αF'. All growth media for XL1-Blue contained 10 μg/ml tetracycline to maintain the F-episome. An incomplete, unidirectional series of deletions from one end of pCHB100 and of pCHB200 (for sequencing of the other strand) was prepared by exonuclease III/mung bean nuclease treatment as described by Henikoff (7). (Some restriction enzymes produce 3'-overhanging ends, which are resistant to exonuclease III attack, while 5'-overhangs are readily digested. The unique arrangement of restriction sites for enzymes creating the two kinds of sticky ends on the polylinker of the Bluescript vector allows the controlled unidirectional digestion of the insert with exonuclease III. Subsequent treatment of the ends with mung bean nuclease creates blunt ends, which can be readily ligated).

Sequencing reactions on single stranded templates were performed using the conditions described in the method section of chapter 4. Utilizing the

obtained sequence information, the remaining gaps were closed using synthetic oligonucleotides, which were prepared as described in chapter 4.

Determination of CAT-activity. *R. sphaeroides* cells harboring various plasmid constructs containing the chloramphenicol-acetyl-transferase (CAT) gene were grown photoheterotrophically in 200 ml Sistrom's medium containing 0.8 μg/ml tetracycline as described in chapter 3. Cells were harvested by centrifugation and resuspended in 5 ml enzyme buffer (0.5 M sucrose, 5 mM ascorbic acid, 0.1 M Tris-Cl (pH 7.8), 5 mM DTT). Following the addition of 0.5 mg DNAse I, the cells were broken in a French press at 18,000 psi. Aliquots of 1 ml were frozen in liquid N_2 and stored at -70°C. As needed, the extracts were thawed on ice, and the debris was removed by centrifugation for 5 min in a microfuge. The protein concentration in the supernatant was determined using the Bradford micro assay (2). For the determination of CAT-activity in the supernatants, 40 μl of enzyme (diluted appropriately with enzyme buffer) were mixed with 5 μl of 0.25 M Tris-Cl (pH 8.0), 1 μl of 10 mM acetyl-CoA, and 0.5 μl of ^{14}C-chloramphenicol (57 μCi/mmol; 25 μCi/ml). Following incubation at 37°C for 1 hour, the reaction mixture was extracted with 300 μl ethyl acetate. The ethyl acetate phase was dried down under a stream of N_2, and the material was resuspended with 10 μl ethyl acetate. Following spotting of the material onto a Si250 Silica plate (Baker), activated for 15 min at 120°C, the plates were developed with chloroform/methanol (95:5, v/v). Radioactivity was determined by scanning individual lanes using a System 200 Imaging Scanner (Bioscan). In addition,

autoradiograms of the TLC plates were obtained to confirm the location of the bands as detected by the scanner. The enzyme concentration was adjusted, such that only monoacetylated product was formed and not more than 80% of the substrate was converted.

RESULTS AND DISCUSSION

Isolation of a cosmid complementing the *sqdB* and *sqdC* mutants.

Individual cosmids containing *R. sphaeroides* wild type sequences were mated from *E. coli* (HB101) into the *sqdB* mutant, and the exconjugants were tested for their ability to synthesize sulfolipid. The cosmid library used was identical to the library described in chapter 4. Approximately 800 exconjugants were screened to isolate the first cosmid complementing the *sqdB* mutant (pCHB1601). Two other complementing cosmids were isolated by screening an additional 100 cosmid clones. Figure 5-1 shows the lipid phenotype of the *sqdB* mutant alone and carrying the cosmid pCHB1601. It was not clear whether all three cosmids were independent, since at least two were found in the same microtiter plate of the cosmid library at fairly close positions, suggesting the possibility of cross contamination. However, all further experiments were done focusing on the cosmid pCHB1601. To test whether this cosmid could complement not only the *sqdB* mutant, it was mated into the *sqdA* and *sqdC* mutants. Testing of the exconjugants for restoration of the ability to synthesize sulfolipid revealed that the cosmid pCHB1601 could not only complement the *sqdB* mutant, but also the *sqdC* mutant. However, the

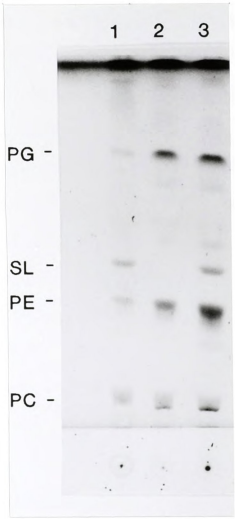


Figure 5-1: Lipid phenotype of the *sqdB* mutant lacking or containing the cosmid pCHB1601. The following lipid extracts were separated on an ammonium sulfate impregnated silica plate. (1) wild type, (2) *sqdB* mutant containing pLA2917, (3) *sqdB* mutant containing pCHB1601. PG, phosphatidyl glycerol; SL, sulfolipid; PE, phosphatidyl ethanolamine; PC, phosphatidyl choline. The lipids were visualized by charring.

sqdA mutant was not complemented. This result indicated that the *sqdA* mutant did not belong to the same complementation group as the other two mutants and, therefore, had to be affected in a different gene. At that stage of the analysis, the question had to be raised whether the *sqdB* and *sqdC* mutants were allelic. This question could be answered through the isolation of two distinctive classes of subclones of pCHB1601 complementing the two mutants as described below.

Construction of a physical map for the cosmid pCHB1601. To construct a restriction map for the cosmid pCHB1601, DNA purified on a CsCl gradient was cut with a series of restriction enzymes. The Enzymes KpnI, HindIII, and EcoRV had only one or two restriction sites internal to the insert. These enzymes were employed in various combinations to establish a simple, but complete map of pCHB1601 (Figure 5-2). The length of the cosmid was estimated to be about 48 kb. A centrally located HindIII site allowed the isolation of the right half of the cosmid as an 11-kb HindIII fragment, which was subcloned into the HindIII site of pLA2917 to give pCHB1612 (Figure 5-2). Recircularization of the remaining plasmid gave rise to pCHB1611, which contained the left half of pCHB1601 (Figure 5-2). These two subclones could be tested for complementation and allowed the mapping of multiple restriction sites for the enzymes PstI, BglII, and EcoRI. Mapping of these sites for clone pCHB1612 was unambiguously possible due to the central location of a unique KpnI site and a single EcoRV site internal to the insert and did not require further Southern analysis (Figure 5-3, B). However, because no unique

Figure 5-2: Simple restriction maps for pCHB1601, the two subclones pCHB1611 and pCHB1612, and pLA2917. Vector sequences are drawn with thick lines. Restriction sites: H, HindIII; K, KpnI; R, EcoRV. The plasmids are shown linearized at one of the EcoRV sites.

restriction sites in pCHB1611 were found, Southern analysis was used to establish the map shown in Figure 5-3 (A). Several combinations of enzymes were used to prepare double digests of the pCHB1611 DNA, which was transferred to nylon membrane filters following electrophoresis. These filters were probed and repeatedly stripped using the vector as well as various subfragments of pCHB1611 as indicated in Figure 5-3 (A). The order of the relatively small number of BglII fragments could be established based on identification of the left and right border fragments by hybridization with the vector probe, and due to the fact that one of the internal BglII fragments was cut with EcoRV, resulting in an unambiguous restriction pattern (Figure 5-2, A). The relatively high number of PstI and EcoRI sites located on the left half of pCHB1611 were mapped by probing the filters with two BglII fragments (B3 and B4; Figure 5-3, A) spanning the left half of the insert. The remaining PstI and EcoRI sites located on the right side of the insert could be deduced without further Southern analysis. Restriction maps of subclones of pCHB1611 (see further below) provided additional information confirming the map outlined in Figure 5-3 (A). It should be pointed out that sites closer together than 0.3 kb might have been missed, due to the resolution limits of the restriction mapping technique (e.g., two BglII sites about 50 bp apart were discovered by sequencing of one of the subclones).

Isolation and sequence analysis of a small subclone of pCHB1601, complementing the *sqdB* mutant. To isolate smaller subclones of pCHB1601, the purified DNA from pCHB1601 was partially digested, and fragments of the

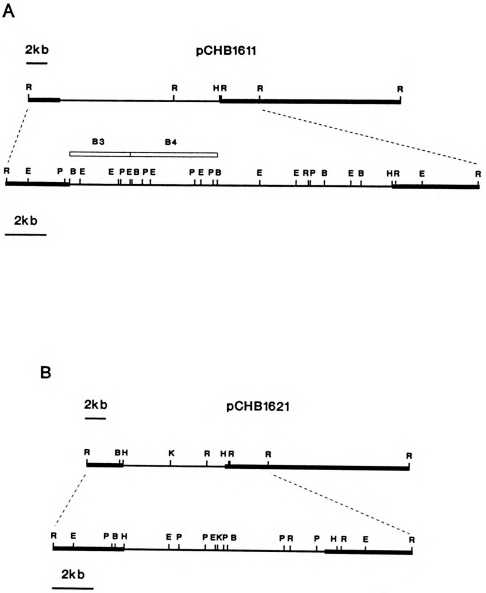


Figure 5-3: Detailed restriction maps for the pCHB1601 subclones pCHB1611 and pCHB1621. Vector sequences are drawn with thicker line. (A) Map for pCHB1611. Open boxes indicate the location of fragments (B3, and B4) used as probes in DNA hybridization experiments as mentioned in the text.(B) Map for pCHB1621. Restriction sites: B, BglII; E, EcoRI; H, HindIII; K, KpnI; P, PstI; and R, EcoRV.

size ranges 2-3 kb, 3-4 kb, and 4-6 kb were ligated into the BglII site of pLA2917 providing the material for three pCHB1601 sublibraries. Out of 96 clones from the 2-3 kb size range library, five could restore the wild type phenotype when mated into the *sqdB* mutant. The smallest, positive clone containing an insert of about 2.4 kb (pCHB16014) was further analyzed. A 2.9 kb fragment containing the insert as well as flanking pLA2917 sequences was excised using a HindIII, PstI partial digest, and was subsequently cloned into the Bluescript vectors pBS-KS$^+$ and pBS-KS$^-$ cut with HindIII and PstI. A restriction map of the 2.9 kb fragment is given in Figures 5-4 and 5-7. The two resulting Bluescript clones pCHB100 and pCHB200 were sequenced using a exonuclease III/mung bean nuclease deletion strategy as described in the methods section (Figure 5-4). Gaps in the sequence were filled using the synthetic oligonucleotide primers listed in Table 5-2.

Analysis of the sequence (Figure 5-5) for open reading frames revealed that two reading frames, located on opposite strands, contained only one or no stop codon, while the other four possible reading frames were blocked with stop codons allowing the presence of only relatively small open reading frames. Due to the high GC-content of 67 % of the examined sequence, which is typical for *R. sphaeroides*, one would expect a statistical underrepresentation of the 3 stop codons TGA, TAA, and TAG with their overall GC-content of only 22 %. The observed low number of stop codons in at least two of the six possible reading frames did not allow for a straightforward identification of the coding sequence for the *sqdB* gene.

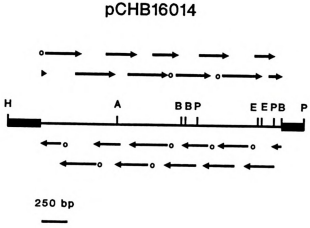


Figure 5-4: Sequencing strategy used to sequence the 2.4-kb insert of pCHB16014. The HindIII, PstI fragment containing the insert of pCHB16014 (thin line) as well as flanking pLA2917 vector sequences (thick line) shown, was cloned into the Bluescript vectors as described in the text. Sequencing runs (arrows) preceded by an open circle were performed using synthetic oligonucleotide primers. Restriction sites: A, BamHI; B, BglII; E, EcoRI; H, HindIII; P, PstI.

Table 5-2: List of synthetic oligonucleotide primers used for sequencing of the pCHB16014 insert.

Primer	Sequence
CB3	5' -GAC, AGG, TTG, TCG, ACG, ATA-3'
CB5	5' -AGA, GGC, ATC, GAG, ATC, GTC-3'
CB6	5' -CCG, CAG, CCC, GTC, GTT, CTG-3'
CB7	5' -ATC, CAG, TCG, GCG, ATC, GGT-3'
CB8	5' -TGC, CAG, ATC, CAA, TCA, GGA-3'
CB9	5' -TCC, TGA, TTG, GAT, CTG, GCA-3'
CB11	5' -TCT, TGC, CGC, CAA, GGA, TCT-3'
CB12	5' -CGA, CGC, AGC, GCA, CCG, AAT-3'
CB13	5' -CAC, GTT, CCG, CAG, CAC, CAG-3'

Figure 5-5: Nucleotide sequence of the insert found in pCHB16014. Only the coding strand for the putative *sqdB* gene is shown. The amino acid sequence deduced by translation of the open reading frame is given beneath the nucleotide sequence. Restriction sites which were used for subcloning are listed above the sequence. Three putative initiation codons (I^{1-3}), as well as a putative ribosome binding site (SD) are indicated above the sequence.

```
1     CCCGGCGGCACGGTGGGTGCCGACAAGATCCGGTCGAGCGAGGAGGCCGTGGCCTTCGTG

61    CGCGGCTTCGTCAGTGCGGGCAAGCCGGTGGCTGCCATCTGCCATGGCCCCTGGGCGCTG

121   GTTGAGGCGGACGTGCTCAAAGGGCGCGAAGTGACCTCCTACCCCTCGCTCGCCACCGAC
                          SalI
181   ATCCGCAACGCAGGGGGCCGCTGGGTCGACCGCGAGGTCGTTGTCGATTCGGGGCTCGTC

241   ACCAGCCGCAAGCCGGACGGATCTCGATGCCTTCTGCGCCAAGATGATCGAGGAATTTGC

301   CGAAGGCGTGCACGACGGCCAGCGCCGCAGCGCCTGAAAAAGCAGGCCCGCCATCGGCAC
                                      I1
361   AGCCTTGCCGAGCGTCGTGAAAAGTTCATTCGACCGGTGGACTTGCAGCCGATACGCACC
          I2                                                  SD
421   AACATGGTCACGAGGACGTGTGCCGCCCGTGAGGCGGCACGCGCGATTCACGAAGGAGGC
          I3
481   TTCATGCGCATCGCAGTTCTGGGTGGGGACGGGTTTGTCGGCTGGCCGACGGCGCTCCAC
1        M  R  I  A  V  L  G  G  D  G  F  V  G  W  P  T  A  L  H
                                   SalI
541   CTGTCGGACCTCGGACACGAGATCCATATCGTCGACAACCTGTCCCGGCGCTGGATCGAT
20     L  S  D  L  G  H  E  I  H  I  V  D  N  L  S  R  R  W  I  D

601   ACCGAGCTGGGCGTGCAGTCCCTGACCCCGATGGACTCGATCCAGGAGCGCTGCCGGATC
40     T  E  L  G  V  Q  S  L  T  P  M  D  S  I  Q  E  R  C  R  I

661   TGGCATCAGGAGACCGGCCAGCGGCTGCATTTCCACCTGCTCGACCTGCGCGAATACGAC
60     W  H  Q  E  T  G  Q  R  L  H  F  H  L  L  D  L  R  E  Y  D
      BamHI
721   CGGATCCGCGCCTGGCTCGCCGAATACCGCCCCGAGGCCATCATCCACTTCGCCGAGCAG
80     R  I  R  A  W  L  A  E  Y  R  P  E  A  I  I  H  F  A  E  Q

781   CGCGCCGCGCCCTATTCGATGAAGTCGGACCGCCACAAGGTCTACACCGTCAACAACAAC
100    R  A  A  P  Y  S  M  K  S  D  R  H  K  V  Y  T  V  N  N  N

841   GTCAATGCGACCCACAACCTGTTGGCCGCCATGGTCGAGACCGGCATCGATGCCCATCTC
120    V  N  A  T  H  N  L  L  A  A  M  V  E  T  G  I  D  A  H  L

901   GTCCATCTCGGCACCATGGGGGTCTATGGCTATTCGACCGTGGGCGCGCCCATCCCCGAG
140    V  H  L  G  T  M  G  V  Y  G  Y  S  T  V  G  A  P  I  P  E

961   GGTTATCTCGACGTCTCGGTCGAGACCCCGGCGGGCCCGAAAGAGCTCGAGATCCTCTAT
160    G  Y  L  D  V  S  V  E  T  P  A  G  P  K  E  L  E  I  L  Y
```

Figure 5-5 (cont'd)

```
1021  CCGACGCGGCCGGGCTCGGTCTATCACATGACCAAGAGCCTCGATCAGATCCTCTTCCAG
180    P  T  R  P  G  S  V  Y  H  M  T  K  S  L  D  Q  I  L  F  Q

1081  TATTATGCTCAGAACGACGGGCTGCGGATCACCGACCTGCATCAGGGCATCGTCTGGGGC
200    Y  Y  A  Q  N  D  G  L  R  I  T  D  L  H  Q  G  I  V  W  G

1141  ACGCACACGAACCAGACCCGCCGCCATCCGCAGCTCATCAACCGCTTCGACTATGACGGC
220    T  H  T  N  Q  T  R  R  H  P  Q  L  I  N  R  F  D  Y  D  G

1201  GACTATGGCACCGTCCTGAACCGCTTCCTGATCCAGTCGGCGATCGGTTATCCGCTGACC
240    D  Y  G  T  V  L  N  R  F  L  I  Q  S  A  I  G  Y  P  L  T

1261  GTGCACGGCACCGGCGGGCAGACCCGCGCCTTCATCCATATCCAGGATTCGGTGCGCTGC
260    V  H  G  T  G  G  Q  T  R  A  F  I  H  I  Q  D  S  V  R  C
                                                  BglII
1321  GTCGAGCTCGCGCTGTCGGATGCGCCCAAGGCCGGCGAACGGGTGAAGATCTTCAACCAG
280    V  E  L  A  L  S  D  A  P  K  A  G  E  R  V  K  I  F  N  Q

1381  ATGACCGAGACGCACCGCGTGCGAGATCTGGCCGAACTGGTGGCGAAGATGACCGGCGCC
300    M  T  E  T  H  R  V  R  D  L  A  E  L  V  A  K  M  T  G  A

1441  AAGGTCTCGTTCCTGCCCAACCCCCGCAAGGAGGCCGACGAGAACGAGCTCGTGGTCAGG
320    K  V  S  F  L  P  N  P  R  K  E  A  D  E  N  E  L  V  V  R

1501  AACGACCAGTTCCTCGCGCTGGGGCTGAAGCCCATCACGCTGCAGGAAGGGCTGCTGGGC
340    N  D  Q  F  L  A  L  G  L  K  P  I  T  L  Q  E  G  L  L  G

1561  GAGGTGGTCGATGTGGCGAAGAAGTTCGCCCACCGGATCGACCGCAGCCGCGTGCCCTGC
360    E  V  V  D  V  A  K  K  F  A  H  R  I  D  R  S  R  V  P  C

1621  GTCTCGGCCTGGACGAAGGACATCGCGCAGCGGGTCGAGCACGATCCCGAGGGGCGGCGG
380    V  S  A  W  T  K  D  I  A  Q  R  V  E  H  D  P  E  G  R  R

1681  CTGCGCTCGGTGTCCTGATTGGATCTGGCAGCCGAAGGACCGGCTGGGTCCGAGCGGGCC
400    L  R  S  V  S

1741  TATGTCACGCTCGTGACCAATGCCGATTATGCGCTGGGGGCACGGGCGCTGCTGCGCTCG

1801  CTGGCCCTCAGCGGCACCACGGCCGACCGGGTGGTGCTGCATACCGACGTGCCCGAGGAG

1861  GCGCTGGCGCCGCTCCGGGCGCTCGGCGCGCGGCTCGTGCGGGTGGAGCTTCTGCCCACC

1921  TCGCCCGAGTTCAACGCGGCCCATGCGCGAGAGGCGCTTCATGCGCGCGCGGCCTTCACC

1981  AAGGGCGGCAAGCCGCCCTTCCACACGCCGCTCGATAATTTCGCCAAGCTCCGGCTCTGG

2041  CAGCTCGTGGACTATCGCTCGGTCGTCTTCATCGACGCGGATGCGCTGGTGCTGCGGAAC
                          EcoRI
2101  GTGGACCGGCTGTTCGACTATCCCGAATTCTGCGCCGCGCCGAATGTCTACGAGAGCCTG
```

Figure 5-5 (cont'd)

```
2161  TCGGACTTTCACCGGATGAATTCGGGCGTCTTCACCGCCCGGCCCTCGACGGACACCTAC

2221  GCCCGGATGCTCGAAGCACTCGACGTGCCGGGGGGCCTTCTGGCGGCGGACCGACCAGAGC

2281  TTCCTGCAGCAGTTCTTCCCCGACTGGCAGGGCCTGCCGGTCTTCTGCAACATGCTCCAG

2341  TATGTCTGGTTCGCGATGCCCGAGCTCTGGAGCTGGGAGC
```

Figure 5-6: Comparison of the putative *sqdB* ribosome binding site to binding sites from various *R. sphaeroides* genes (A) and to the 3'-end of the 16S rRNA from *R. sphaeroides* (B). Bases matching the complementary sequence of the 16S rRNA 3'-end representing the ribosome binding site are shaded. Sequences were taken from: *sqdA* (this study), *pucB* and *pucA* (10), *puhA* (5), *pufL* (13), *pufM* (14), *pufB* and *pufA* (9), and *R. sphaeroides* 16S rRNA 3'-end (6).

A

```
          -15       -10       -5        0
sqdB       C A C G A A G G A G G C T T C A T G
sqdA       A G C A G G A G G C C C C G G G T G
pucB       T T G G G A G A C G A C A C A G T G
pucA       T A G G A G A A G A C T G A C A T G
puhA       C G G A G G A C C C C T T A A A T G
pufL       A G C G G A G A G G G A A G C A T G
pufM       T C C C G G G A G G C A T C A A T G
pufB       G A T C C G G A G G A T A G C A T G
pufA       G A G G A G A A C A A G A T C A T G
```

B

```
16S 3'     HO-U C U U U C C U C C U C C A C U A  5'
rRNA                     * * * * * *
sqdB         5' C A C G G A A G G A G G C T T C A T G  3'
```

However, the reading frame containing one stop codon looked promising, because one of the in-frame ATG codons (Figure 5-5, I^3) was preceded by a putative ribosome binding site showing the AGGAGG motif complementary to the 3' end of the 16S rRNA of *R. sphaeroides* (Figure 5-6, B). The spacing between the putative ribosome binding site and the ATG initiation codon fell into the range observed for other genes of *R. sphaeroides* (Figure 5-6, A). None of the other possible initiation codons showed a comparable ribosome binding site. Beginning at the initiation codon described above (Figure 5-5, I^3), the open reading frame could encode a 46-kD protein. However, the absence of a stop codon upstream of the respective initiation codon inside the pCHB16014 insert did not allow the determination of an upper limit for the length of the coding sequence. Since at least one stop codon is required for the termination of translation, it seemed unlikely that the other reading frame, that lacked a stop codon, could represent the coding sequence. Given the indicated uncertainties about the location of the *sqdB* gene on the pCHB16014 sequence, additional experiments were required to determine the correct open reading frame.

Complementation analysis using various deletion clones derived from pCHB16014. Two simple deletion clones of pCHB16014 were constructed. First, digestion of pCHB16014 with BglII followed by recircularization resulted in a 1 kb deletion expanding from the right end of the insert into the 3'-end of the putative open reading frame for the *sqdB* gene (Figure 5-7, pCHB16014D-B). Second, double digestion of pCHB16014 with BglII and BamHI followed by recircularization gave rise to clone pCHB16014D-AB, which carried a deletion of

1.7 kb beginning at the right insert end (Figure 5-7). Both deletion clones were unable to restore the wild type phenotype when present in the *sqdB* mutant, indicating that parts of the open reading frame or an essential promotor sequence were lost. As described for the analysis of the putative open reading frame for the *sqdA* gene in chapter 4, various fragments of the pCHB16014 insert were cloned behind a cytochrome-c promotor from *R. capsulata* located on the pCHB500 vector (see Appendix B for the construction and analysis of this vector). Two SalI sites, one upstream of the putative initiation codon (Figure 5-5, I[3]) and one downstream in the putative coding region for the *sqdB* gene were particularly useful for the analysis. A 1.9 kb fragment spanning the region from the upstream SalI site to the nearest EcoRI site downstream of the putative open reading frame was obtained by a partial SalI, EcoRI digest and cloned into the Bluescript vector cut with SalI and EcoRI. From the resulting clone, the fragment was excised with KpnI and EcoRI, and cloned into pCHB500 cut with the same enzymes to construct pCHB16201 (Figure 5-7). To obtain the opposite orientation, the fragment was excised from the Bluescript clone with KpnI and XbaI, and cloned into the respective sites of pCHB500 giving rise to clone pCHB16202 (Figure 5-7). For the construction of the two clones pCHB16151 and pCHB16152 (Figure 5-7), a shorter fragment of 1.55 kb from the second SalI site to the EcoRI site nearest to the 3'-end of the putative open reading frame was cloned in both orientations into pCHB500 using the same strategy as described above. In a subsequent complementation experiment, both of the clones containing the larger insert in either orientation (pCHB16201,

and pCHB16202) were able to restore the capability of the *sqdB* mutant to synthesize sulfolipid, while neither of the smaller clones (pCHB16151, and pCHB16152) could restore the wild type phenotype when present in the *sqdB* mutant. Therefore, it was concluded that the *sqdB* gene must be located within the 1.9 kb SalI/EcoRI fragment of pCHB16014. In addition, this fragment must contain an endogenous promotor driving the expression of the *sqdB* gene, since complementation was independent of the orientation of the insert relative to the external promotor provided by the pCHB500 vector. The fact that the smaller SalI, EcoRI fragment cloned in both orientations into pCHB500 failed to provide sufficient sequence for complementation, was consistent with the idea that this fragment did not include the N-terminal portion of the putative coding sequence for the *sqdB* gene as shown in Figure 5-7, and that it was missing a putative promotor proposed to be located on the 350 bp fragment between the two SalI sites. An open reading frame analysis performed on the sequence contained in the 1.9 kb SalI/EcoRI fragment of pCHB16014 revealed only one large open reading frame identical to the one described above (Figure 5-5), with no or relatively small open reading frames, as defined by the presence of an initiation as well as a stop codon, located on the opposite strand. From the deletion and complementation experiment described above, it was concluded that the initiation codon had to be located between the two SalI sites. It was proposed that only one of the three initiation codons located on this fragment (Figure 5-5, I^3), was a good candidate, since it was preceded by a ribosome binding site (Figure 5-6).

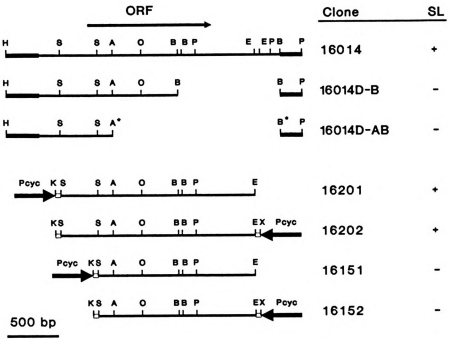


Figure 5-7: Complementation analysis using various subclones of pCHB16014. the putative open reading frame (ORF) for the *sqdB* gene is indicated above the pCHB16014 sequence (the pCHB prefix of the clone names was omitted). The complementation result is indicated as absence (-) or presence (+) of the sulfolipid (SL). Three independently constructed clones for each of the constructs shown were tested for complementation, and gave identical results. Vector sequences derived from pLA2917 are drawn with a thicker line, while sequences derived from the Bluescript polylinker are shown as open boxes. The location and direction of the cytochrome-c promotor (Pcyc) is indicated by a thick arrow. Restriction sites: A, BamHI; B, BglII; E, EcoRI; H, HindIII; K, KpnI; O, XhoI; P, PstI; S, SalI; X, XbaI. Sites which were lost during the cloning procedure are marked (*).

However, due to the lack of suitable restriction sites, it was not possible to isolate a fragment which would contain the putative *sqdB* coding sequence without being preceded by an internal promotor. It was therefore not feasible to confirm the orientation of complementation by cloning the promotor-less coding sequence in both orientations behind the promotor provided by pCHB500. In the case of the *sqdA* gene (chapter 4), orientation dependent complementation was observed by cloning the putative coding sequence lacking a promotor into pCHB500, thereby confirming the coding strand. A different approach was chosen to confirm the predominant orientation of gene expression in the insert of pCHB16014 containing the *sqdB* gene.

Expression of a promotor-less CAT-gene inserted into pCHB16014. To determine the predominant orientation of gene expression and, thereby, the orientation of the *sqdB* gene on pCHB16014, a 0.9 kb Sau3AI fragment containing the chloramphenicol acetyl transferase coding sequence but not a promotor was excised from pBR328 and cloned in both orientations into the unique BamHI site of pCHB16014. The two resulting clones pCHB16014CAT-1 and -2 (Figure 5-8, A) were mated into the *sqdB* mutant along with the pLA2917 vector and pCHB16014. Upon testing of the lipid composition of the resulting strains, it was noted that both clones containing the insertions lost their ability to complement the defective *sqdB* gene, indicating that the insertion resulted in the disruption of the *sqdB* gene located on pCHB16014. Measurements of the CAT-activity in extracts obtained from the four cell lines described above revealed that the presence of the clone pCHB16014CAT-1 in the *sqdB* mutant

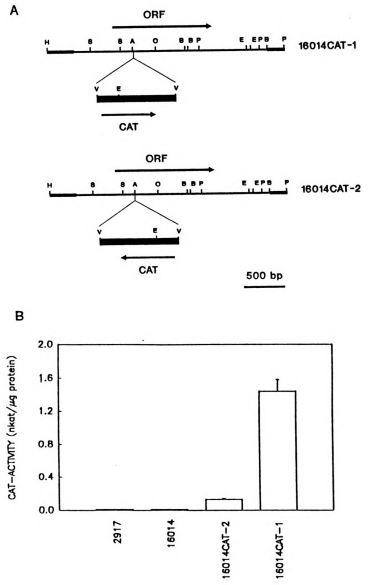


Figure 5-8: Cat-gene insertion clones of pCHB16014 and CAT-activity observed in cells containing these clones. (A) Location and orientation of the fragment containing the CAT coding sequence (filled box) in clones pCHB16014CAT-1 and -2. The exact location and orientation of the putative *sqdB* open reading frame (ORF) and the coding sequence for the CAT-gene (CAT) are indicated by arrows. Vector sequences derived from pLA2917 are drawn with a thicker line. Restriction sites: A, BamHI; B, BglII; E, EcoRI; H, HindIII; O, XhoI; P, PstI; V, Sau3AI. (B) Cat-activity measured in extracts from *sqdB* mutant cells containing the plasmids pLA2917, pCHB16014, pCHB16014CAT-1, and pCHB16014CAT-1. Three replicates were measured and averaged for each of the extracts, and the calculated standard errors are indicated.

resulted in ten-fold higher activity compared to pCHB16014CAT-2. No CAT-activity could be detected in extracts obtained from cells harboring pLA2917 or pCHB16014 (Figure 5-8, B). Since in pCHB16014CAT-1 the orientation of the CAT coding sequence was the same as for the putative *sqdB* coding sequence, it was concluded that the main orientation of gene expression coincided with the orientation of the putative *sqdB* gene and its promotor, supporting the idea that the coding sequence as depicted in Figure 5-5 indeed represents the *sqdB* gene. However, since small but significant amounts of CAT-activity were observed when the CAT coding sequence and the putative *sqdB* coding sequence were facing in opposite direction in pCHB16014CAT-2 (Figure 5-8), one or the other strand could not be definitively excluded as the location for the *sqdB* gene coding sequence.

Analysis of the amino acid sequence encoded by the putative *sqdB* gene. The open reading frame, which was proposed to contain the *sqdB* coding sequence based on the experiments described above, was translated into the amino acid sequence (Figure 5-5). The resulting sequence contained 404 amino acids and could give rise to a 46-kD protein. The amino acid sequence was subjected to a hydropathy analysis according to the method described by Kyte and Doolittle (11). The result (Figure 5-9) led to the conclusion that the putative protein is probably soluble rather than membrane bound, because no major hydrophobic domains typical for membrane spanning regions could be detected. Most proposed reactions involved in sulfolipid biosynthesis (see Figure 1-2, chapter 1), would not require membrane

bound proteins, except for the enzyme catalyzing the last reaction.

Using the amino acid sequence deduced from the putative coding sequence of the *sqdB* gene in a homology search involving two protein data bases, it was discovered that the amino acid sequence for the enzyme UDP-glucose epimerase from various fungi and bacteria showed homology to the N-terminal region of the putative *sqdB* gene product (Figure 5-10). Out of 100 amino-terminal amino acid residues 13 were identical in all three sequences (Figure 5-10, shaded amino acids). Given the degree of identity of about 25% observed in the stretches of homology, and given the number of conservative substitutions, which did not change the overall character of the protein, the pattern of homology between the UDP-glucose epimerase genes from various organisms and the putative *sqdB* gene product opened the possibility for speculation about the function of the putative *sqdB* gene. The enzyme UDP-glucose epimerase catalyzes the conversion of UDP-glucose to UDP-galactose (8). The fact that a class of enzymes involved in sugar nucleotide metabolism showed homology to the putative *sqdB* gene product, thought to be involved in the biosynthesis of the plant sulfolipid, was very exciting, because UDP-sulfoquinovose belonging to the same class of compounds was proposed to be a precursor of the plant sulfolipid (chapter 1). It is therefore conceivable that the *sqdB* gene product may be involved in sulfolipid biosynthesis at the level of UDP-sulfoquinovose or some other sugar nucleotide. In addition, this finding of homology supported the idea that the coding sequence proposed for the *sqdB* gene (Figure 5-5), actually constituted the

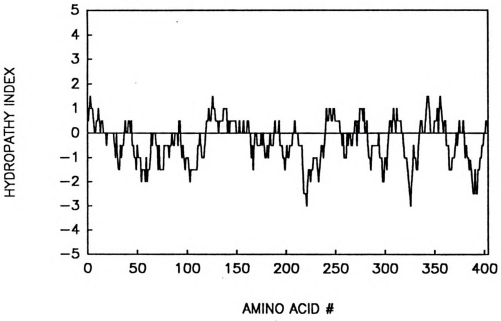


Figure 5-9: Hydropathy plot for the putative *sqdB* gene product. Included in the analysis were 404 amino acids, beginning at the initiation codon I[3] as shown in Figure 5-5.

Figure 5-10: Comparison of the amino acid sequences of the putative *sqdB* gene product and UDP-glucose epimerases from *S. cerevisiae* and *E. coli.* (A) The amino acid sequences were translated from the nucleotide sequence of the *gal10* gene of *S. cerevisiae* (4), and (B) from the *galE* gene of *E. coli* (12). Single dots indicate conservative substitutions, while colons represent perfect identity. Amino acids conserved in all three sequences are shaded.

A

```
                    10        20        30        40        50        60
sqdB        MRIAVLGGDGFVGWPTALHLSDLGHEIHIVDNLSRRWIDTELGVQSLTPMDSIQERCRIW
              . : ::.:..: .:...: . :..  ..:::...: .. .. .. . . .. ..
gal10       KIVLVTGGAGYIGSHTVVELIENGYDCVVADNLSNSWSSVRVTGRRAGDVLNLTAKPDRA
                   20        30        40        50        60        70

                    70        80        90       100       110
sqdB        HQETGQRLHFHLLDLREYDRIRAWLAEYRPEAIIHFAEQRAAPYSMKSD-RHKVYTVNNN
            ..:   . .... :  ... .. : .: . .  .. .:.: .. .:. : :  . .....
gal10       KRELKWQTELQVED--SCKDLWKWTTENPFGYQLRGVEARFSAEDMRYDARFVTIGAGTR
                   80          90       100       110       120

          120        130       140       150
sqdB        VNAT-HNLLAAMVETGIDAHLVHLGTMGVYGY
             .::  :: :..:.  .... : ::  .  ::
gal10       FQATFANLGASIVDLKVNGQSVVLGYENEEGY
          130        140       150       160
```

B

```
                    10        20        30        40        50        60
sqdB        MRIAVLGGDGFVGWPTALHLSDLGHEIHIVDNLSRRWIDTELGVQSLTPMDSIQERCRIW
            ::. : ::.:..: .: ..: . ::.. :.:::...  ..   ...:. ....     :.
galE        MRVLVTGGSGYIGSHTCVQLLQNGHDVIILDNLCNSKRSVLPVIERLGGKHPTFVEGDIR
                    10        20        30        40        50        60

                     70         80        90       100       110
sqdB        HQE-TGQRLHFHLLD-LREYDRIRAWLAEYRPEAIIHFAEQRAAPYSMKSDRH
            ...  .. :: : .: . ... ..: ..:  .... ..... ... .. :...
galE        NEALMTEILHDHAIDTVIHFAGLKA-VGESVQKPLEYYDNNVNGTLRLISAMR
                    70        80         90       100       110
```

sqdB gene, since it seemed unlikely that a random sequence could show the observed degree of homology.

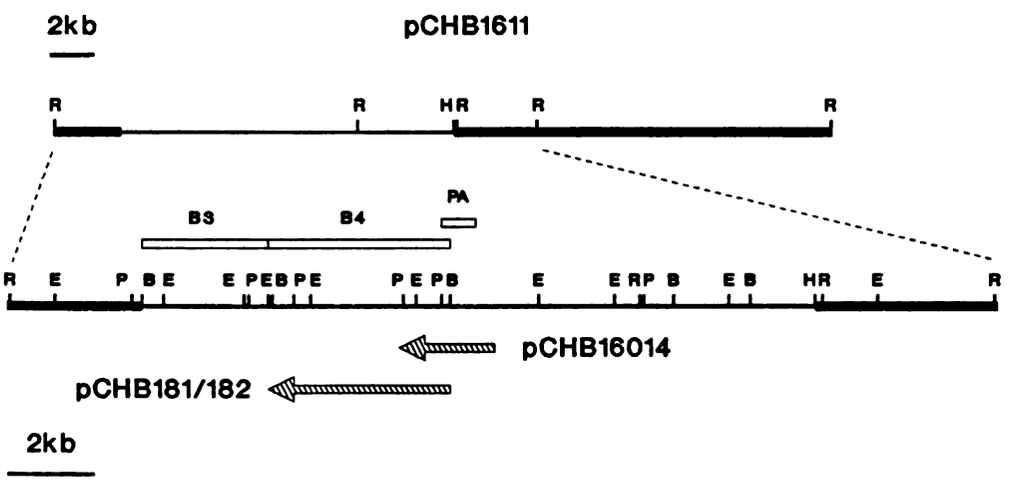
Localization and orientation of the *sqdB* and *sqdC* genes relative to each other on pCHB1601. To determine the position of the putative *sqdB* gene on the map of the original cosmid pCHB1601, an 0.8-kb BamHI, PstI fragment internal to the *sqdB* coding region (fragment PA; Figure 5-11, A) was used to probe DNA from pCHB1611 and pCHB1621 digested with PstI and BglII, containing the left and right half of the cosmid pCHB1601 (Figure 5-2). Only fragments of pCHB1611 hybridized to this *sqdB* specific probe. From the results of this Southern experiment, it was concluded that the putative *sqdB* gene was located in the center of pCHB1611, and by comparison of the restriction maps of pCHB16014 (Figure 5-7) and pCHB1611, its orientation was deduced (Figure 5-11).

From initial experiments it was known that the *sqdC* mutant was complemented by the cosmid pCHB1601. Subclones pCHB1611 and pCHB1621 (Figure 5-2, 5-3) were tested for their ability to restore the wild type phenotype when present in the *sqdC* mutant. Only pCHB1611 which contained the left half of pCHB1601 on which the putative *sqdB* gene was also located, was capable of complementing the *sqdC* mutant. At this stage of the analysis, it was not clear that the two mutants carried mutations in two different genes. However, screening the sublibrary of pCHB1601, containing 2-3 kb fragments, did not yield a positive clone which could complement the *sqdC* mutant. This was the first genetic evidence that the two mutants, indeed, represented two

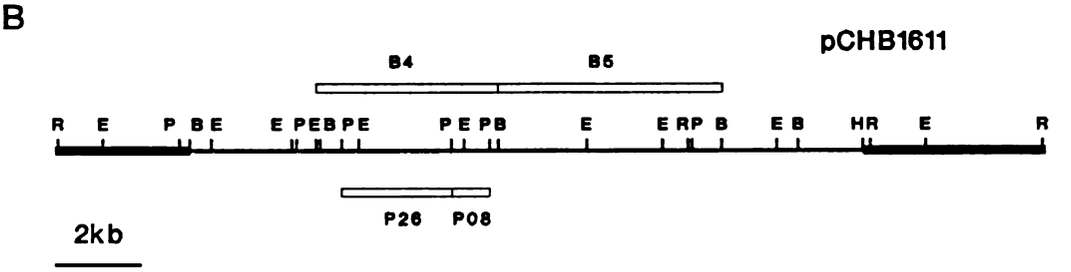
different genes, since the same sublibrary did yield five positive clones complementing the *sqdB* mutant (see above). Upon screening of the sublibrary of pCHB1601 that contained the 4-6 kb fragments, two clones could be isolated which were able to restore the wild type phenotype of the *sqdC* mutant. Restriction analysis of these two clones revealed that they were identical, possibly clonal. Further analysis was, therefore, restricted to one of the clones, pCHB16018. Detailed restriction mapping and Southern analysis revealed that this clone contained two independent fragments resulting in an insert size of about 10 kb (Figure 5-11, C). Using the DNA of the 2.6-kb PstI band of pCHB1611, which contained two comigrating PstI fragments, one extending into the left vector border and one designated P26 (Figure 5-11, B) as a probe in a Southern experiment, the central 2.6-kb PstI fragment of pCHB16018 showed a positive signal. The P26 fragment was thought to be located directly downstream of the *sqdB* gene, and its neighboring 0.8-kb PstI fragment P08 (Figure 5-11, B) was thought to be contained in pCHB16014 extending into the putative coding sequence of the *sqdB* gene (compare Figure 5-11, A). Probing restricted DNA of pCHB16018 using the 0.8-kb PstI fragment of pCHB16014 (identical to P08; Figure 5-11, B), gave a positive result suggesting that a contiguous stretch of DNA covering the central 2.6-kb PstI fragment of pCHB16018 and portions of its flanking sequences (indicated as open box in Figure 5-11, C) were located on pCHB1611 directly downstream from the *sqdB* gene. Insert sequences of pCHB16018 located to the left of this central region (Figure 5-11, C) were found to hybridize to the 5.3-kb BglII fragment

Figure 5-11: Location and orientation of the *sqdB* and *sqdC* genes on pCHB1611. (A) Restriction map of pCHB1611; the location of the two subclones pCHB16014 and pCHB181 that complement the mutants *sqdB* and *sqdC,* respectively, is indicated as cross hatched arrows. The presumed orientation of gene expression is indicated by the arrow heads. PA (open box) depicts the location of a 0.8-kb PstI, BamHI fragment from pCHB16014 as determined by Southern analysis. (B) Restriction map of pCHB1611 showing the location of various fragments (B4, B5, P26, and P08; open boxes) as described in the text. (C) Restriction maps of the three subclones pCHB16018, pCHB181, and pCHB182 containing the *sqdC* gene. Indicated above the pCHB16018 map (jagged, open bar) is the approximate extent of the DNA stretch common to pCHB181 and pCHB182, and pCHB16018. The three maps were aligned at the sites located in this region. The orientation and location of the external cytochrome-c promotor, located on the pCHB500 vector, is indicated by filled arrows. Restriction sites: B, BglII; E, EcoRI; H, HindIII; P, PstI; R, EcoRI. Sites marked (*) were lost due to the cloning procedure.

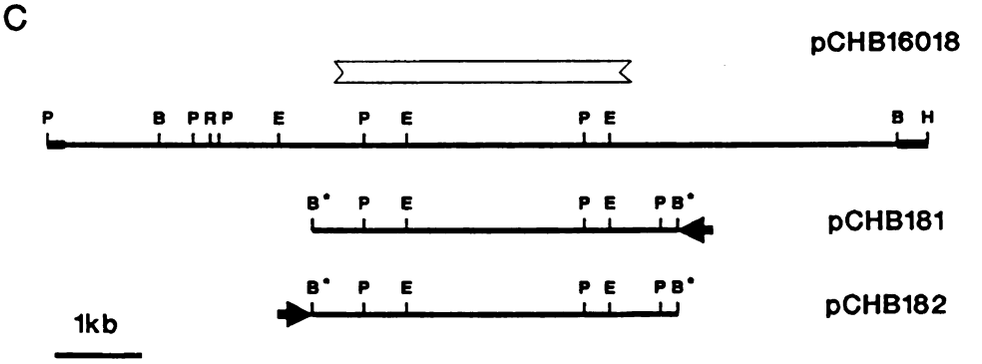
A


2kb
pCHB1611
R
R
HR
R
R
PA
B3
B4
R E P B E E PEB PE P E PB E E RP B E B HR E R
pCHB16014
pCHB181/182
2kb

B


pCHB1611
B4
B5
R E P B E E PEB PE P E PB E E RP B E B HR E R
P26
P08
2kb

C


pCHB16018
P B PRP E P E P E B H
B* P E P E PB*
pCHB181
B* P E P E PB*
pCHB182
1kb

(B5; Figure 5-11, B) of pCHB1611. To test the possibility that the *sqdC* gene was located directly downstream from the *sqdB* gene, a 4.3-kb BglII fragment (B4; Figure 5-11, B) of pCHB1611 beginning inside the *sqdB* coding sequence, and containing the two PstI fragments designated P26 and P08 (Figure 5-11, B), was cloned in both directions into the BamHI site of the pCHB500 expression vector, giving rise to clones pCHB181 and pCHB182 (Figure 5-11, C). Mating these two clones into the *sqdC* mutant indicated that only pCHB181, but not pCHB182 could restore the wild type phenotype. This result clearly located the *sqdC* gene on a defined fragment of pCHB1611 which overlapped with the putative *sqdB* gene. The orientation dependence of the fragment relative to the cytochrome-c promotor in pCHB500 suggested that the *sqdC* gene was expressed in the same orientation as had been determined for the *sqdB* gene (Figure 5-11, A). In addition, the result indicated that the *sqdC* gene was not directly preceded by a promotor, and therefore the ability of small DNA fragments to restore the wild type phenotype of the *sqdC* mutant depended on an external promotor as in pCHB181. The fact that using the sublibrary approach the only complementing clone (pCHB16018), that could be isolated from pCHB1611 contained two large fragments, was consistent with this idea, because the second insert might have brought a promotor sequence close to the *sqdC* coding sequence in pCHB16018 to allow for complementation.

Location of the *sqdB* and *sqdC* genes relative to the *sqdA* gene. As described above, the cosmid pCHB1601 did not complement the *sqdA*

mutant described in the previous chapter. However, the reciprocal experiment, mating the cosmid pCHB1701 complementing the *sqdA* mutant into the *sqdB* and *sqdC* mutants and testing for complementation has not been done. Using the left border fragment of pCHB1611 (fragment B3; Figure 5-3, A) to probe pCHB1701 DNA in a Southern experiment, no hybridizing fragments could be detected. The right end was not tested. Despite this incomplete Southern and complementation analysis, a comparison of two distinct restriction maps of pCHB1701 and pCHB1601, made it appear unlikely that the two cosmids share common DNA sequences, suggesting that the three genes belong to two relatively distant gene clusters or operons. However, further analysis is required to carefully test this idea.

CONCLUSIONS

Using a complementation assay, a cosmid complementing the *sqdB* and *sqdC* mutants was isolated. Two subclones containing the left half (pCHB1611) and the right half (pCHB1612) of pCHB1601 were constructed, and a physical map was prepared by restriction and Southern analysis (Figure 5-3). Further subcloning allowed the isolation of a 2.4-kb subclone (pCHB16014), which could complement the *sqdB* mutant. The location and orientation of the coding sequence for the *sqdB* gene was identified using sequence, deletion, and insertion analysis as described above. Based on these data, an open reading frame for the *sqdB* gene (Figure 5-5) preceded by a ribosome binding site (Figure 5-6) was proposed. The proposed open

reading frame could encode a 46-kD protein. The N-terminal quarter of the amino acid sequence showed homology to the N-terminal portion of the UDP-glucose epimerase from the yeast and *E. coli* (Figure 5-10). This finding suggested the possibility that the product of the *sqdB* gene is involved in sulfolipid biosynthesis at the level of UDP-sulfoquinovose or some other sugar nucleotide.

A subclone containing two fragments derived from pCHB1601 was isolated (pCHB16018) which was able to complement the *sqdC* mutant. Southern and restriction analysis of this clone suggested that the central portion contained a sequence, which was located directly downstream of the *sqdB* gene on the physical map of pCHB1611 (Figure 5-11). A fragment of pCHB1611 partially overlapping the *sqdB* gene was able to complement the *sqdC* mutant in an orientation-dependent fashion, when cloned into the pCHB500 expression vector. The orientation of complementation coincided with the proposed orientation of the *sqdB* gene. Based on the results discussed above, it is proposed that the *sqdB* gene and the *sqdC* genes may be organized in an operon which is transcribed from a promotor upstream of the *sqdB* gene. Experiments can be designed to test this idea. Deletion and sequence analysis should allow the delineation of the location of the *sqdC* coding sequence with respect to the *sqdB* gene. Analysis of the RNA species hybridizing to the two genes as well as mapping of the transcriptional start site(s) should clarify whether the *sqdB* and sqdC genes are present in a single operon. In addition, other genes involved in sulfolipid biosynthesis may

be located nearby. To find these genes, the construction of a deletion mutation of *R. sphaeroides* lacking the region spanned by pCHB1611 will be useful. This would allow the construction of defined mutations (deletion, insertions) in the clone pCHB1611, which could be tested in *trans*. Alternatively, a series of gene replacements spanning the region of the cosmid could be constructed.

Based on the available evidence, the *sqdA* mutant described in the previous chapter may mark a second operon or gene cluster involved in sulfolipid biosynthesis in *R. sphaeroides*. Further experiments will be required to test this idea. Some of the mutants described in chapter 3, which have not been analyzed, might possibly carry mutations in additional genes. The fact that all three mutants analyzed in detail are affected in different genes, suggests that the experiments did not saturate the pathway for sulfolipid biosynthesis with mutations.

REFERENCES

1. Balbas P., Soberon X. Merino E., Zurita M., Lomeli H., Valle F., Flores N., Bolivar F. (1986). Plasmid vector pBR322 and its special-purpose derivatives - a review. Gene 50:3-40

2. Bradford M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-252

3. Bullock W.O., Fernandez J.M, Short J.M (1987). XL1-Blue: A high efficiency plasmid transforming recA *E. coli* strain with beta-galactosidase selection. Bio Techniques 5;376-378

4. Citron. B.A., Donelson J.E. (1984). Sequence of the *Saccharomyces GAL* region and its transcription *in vivo*. J. Bacteriol. 158:269-278

5. Donohue T.J., Hoger J.H., Kaplan S. (1986). Cloning and expression of the *R. sphaeroides* reaction center H gene. J. Bacteriol. 168:953-961

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

7. Henikoff Steven (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351-359

8. Kalckar H.M. (1958). Uridine diphosphogalactose: metabolism, enzymology, and biology. Adv. Enzym. 20:111-134

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

10. Kiley P.J., Kaplan S. (1987). Cloning, DNA sequence, and expression of the *R. sphaeroides* light-harvesting B800-850-α and B800-850-ß genes. J. Bacteriol. 169:3268-3275

11. Kyte J., Doolittle R. F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

12. Lemaire H.G., Muller-Hill B. (1986). Nucleotide sequences of the *galE* gene and the *galT* gene of *E. coli*. Nuc. Acid Res. 14:7705-7711

13. Williams J.C., Steiner L.A., Ogden R.C., Simon M.I., Feher G. (1983). Primary structure of the M subunit of the reaction center from *R. sphaeroides*. Proc. Natl. Acad. Sci. USA 80:6505-6509

14. Williams J.C., Steiner L.A., Feher G., Simon M.I. (1984). Primary structure of the L subunit of the reaction center from *R. sphaeroides*. Proc. Natl. Acad. Sci. USA 81:7303-7307

CHAPTER 6

IN VITRO SULFOLIPID BIOSYNTHESIS FROM ^{35}S-PAPS IN EXTRACTS FROM *R. SPHAEROIDES*

ABSTRACT

Crude extracts of *R. sphaeroides* cells were tested for their capability to convert ^{35}S-PAPS into sulfolipid. Incubating labelled PAPS with crude extracts from wild type or *sqdB* mutant cells resulted in the production of several labelled, chloroform-soluble compounds, which were usually not found in extracts from cells labelled *in vivo*. However, a labelled compound comigrating with sulfolipid in different TLC systems was observed in extracts from wild type but not from *sqdB* mutant cells. The identity of the sulfolipid was confirmed by comigration of the deacylation product with deacylated sulfolipid prepared *in vivo*. Intact cells, which cannot take up PAPS, were not able to label sulfolipid from PAPS, indicating that the conversion of PAPS into sulfolipid by extracts was truly an *in vitro* reaction. Maximal rates of sulfolipid synthesis from PAPS *in vitro* were about 30 pmol/h mg protein. Addition of water soluble compounds extracted from wild type cells, as well as a mixture of ATP, $MgCl_2$, NADPH, 3-phosphoglycerol, 3-phosphoglycerate, and glucose-6-phosphate did not stimulate the synthesis of sulfolipid in the extracts. In a preliminary series of experiments it was determined that possible intermediates or substrate analogues such as cysteic acid or DL-glyceraldehyde did not inhibit *in vitro* sulfolipid biosynthesis, but that addition of UDP-glucose had a small

inhibitory effect. Reduced sulfur compounds such as DTT and sulfite had a more pronounced inhibitory effect.

INTRODUCTION

As a first step for the initiation of a biochemical analysis of the enzymes involved in sulfolipid biosynthesis, a cell-free system to study the process *in vitro* is required. Such a cell-free system allows the analysis of the effects of potential inhibitors and precursors of sulfolipid biosynthesis, without regard to whether the respective compound is taken up into the cell or not. Separating the enzymes catalyzing the various reactions into membrane-bound and soluble proteins and dissecting the reaction sequence with inhibitors and substrates can then eventually lead to the identification and purification of individual enzymes. Although several investigators have attempted to study the reactions involved in sulfolipid biosynthesis in extracts of broken chloroplasts, no success has been reported using chloroplast extracts. However, there is one published study by Hoppe and Schwenn (3), who reported the incorporation of labelled sulfate, sulfite, and PAPS into sulfolipid by cell-free extracts from *Chlamydomonas reinhardii*. Incorporation of labelled sulfate required ATP and magnesium, and was stimulated by the addition of glutathione, while addition of cysteic acid did not affect the incorporation. PAPS behaved as a legitimate substrate, since the reaction followed a saturation curve with increasing amounts of PAPS. On the other hand, the incorporation of sulfite into sulfolipid could not be saturated, even with very

high amounts of sulfite, suggesting a non-enzymatic artifact.

Although mutants of *R. sphaeroides* deficient in sulfolipid biosynthesis, and their complementing genes, were isolated by genetic and molecular-biological techniques as described in the previous chapters, the function of the genes has to be elucidated with biochemical means to determine the pathway for sulfolipid biosynthesis. The accumulation of water-soluble, ^{35}S-labelled compounds in some of the mutants provides an opportunity for such a biochemical analysis since elucidation of the structure of the compounds may reveal the enzymatic reaction affected in the respective mutant. However, the confirmation of their precursor nature has to be by feeding these compounds to cell extracts from *R. sphaeroides*. To complete the analysis, it was, therefore, necessary to develop an *in vitro* sulfolipid biosynthesis system using extracts of wild type *R. sphaeroides* cells. In this chapter, exploratory experiments are described concerning the development of a cell-free system to study sulfolipid biosynthesis.

MATERIALS AND METHODS

Bacterial strains. The experiments described in this chapter involved two strains of *R. sphaeroides*, the wild type 2.4.1 obtained from R.L. Uffen's lab at MSU, and the sulfolipid-deficient *sqdB* mutant described in chapter 3.

Preparation of the sulfur-activating enzymes from yeast. A crude enzyme preparation from yeast was employed to prepare ^{35}S-3'-phosphoadenosine-5'- phosphosulfate (PAPS) using a modified method

according to Robbins (5). Three enzymes are required to prepare PAPS from ATP and ^{35}S-sulfate. ATP-sulfurylase catalyzes the reaction: $ATP+SO_4+H^+ \rightarrow$ adenosine phosphosulfate (APS) + pyrophosphate (PP). The energetically unfavored synthesis of APS is driven by the removal of pyrophosphate from the equilibrium by pyrophosphatase, and by the concomitant conversion of APS to PAPS by APS-kinase. The yeast preparation described below contained all three enzymes. As starting material 500 g dry yeast, obtained from a bakery supplier, were resuspended in 1.5 l of 50 mM K_2PO_4 (pH 9.2) and the suspension was poured into 4 l of liquid N_2. After thawing and addition of 0.5 l of 50 mM K_2PO_4 (pH 9.2), the suspension was stirred at 4°C over night. All of the following steps were done at 4°C. The debris was sedimented at 9000 rpm for 10 min using a Sorvall GS-3 rotor (Dupont). The extract was acidified with 1 N acetic acid to pH 5.1 and, the material centrifuged again as described above. To 1 l of supernatant 1.5 l of saturated ammonium sulfate solution (1400 g of ammonium sulfate dissolved in 2 l of water, 0.2 mM EDTA, pH 7.5) were added. Following equilibration, the material was centrifuged as described above. The supernatant was discarded and the pellet was resuspended in 25 ml of 20 mM Tris-Cl (pH 8.0). The samples were frozen in 5 ml aliquots in liquid N_2 and stored at -20°C. Prior to use, the enzyme solution was dialyzed against 4 X 1 l of 20 mM Tris-Cl (pH 8.0) to remove ammonium sulfate.

Preparation of ^{35}S-PAPS using the sulfur activating enzyme isolated from yeast. Reaction mixtures of 1 ml (usually 5 separate reactions) were set

up for the synthesis of ^{35}S-PAPS from ^{35}S-sulfate. The reaction mixtures contained (final concentrations): 100 mM Tris-Cl (pH 8.5), 10 mM ATP (1 M stock solution in 0.1 M Tris; pH 8.0), 10 mM $MgCl_2$, about 600 µl dialyzed enzyme suspension, and 500 µCi ^{35}S-sulfate (approximate specific activity of 10 µCi/nmol, adjusted by the addition of cold sulfate to carrier free ^{35}S-sulfate; it was assumed that only negligible amounts of sulfate were carried over by the dialyzed enzyme suspension). The mixtures were incubated for 2.5 h at 37°C and stopped by addition of 2 ml of ethanol. The proteins were allowed to precipitate for 10 min on ice and then sedimented in a clinical centrifuge at top speed. The conversion of ^{35}S-sulfate and ATP into ^{35}S-PAPS was tested by TLC on polyethylene-imine cellulose plates (see below), and was repeatedly found to be about 60 %. Only two labelled compounds were detected, assumed to be sulfate and PAPS as determined by comigration with standards . The supernatant solutions were stored at -20°C until they were used for the purification of labelled PAPS.

Analysis and purification of the ^{35}S-PAPS by anion exchange chromatography. To test the quality of the PAPS preparation, the material was analyzed by anion exchange chromatography on polyethylene-imine plates (0.1 mm thick on plastic sheets with fluorescence indicator, obtained from Baker). The plates were soaked in methanol and air-dried prior to spotting, and were developed using 0.9 M guanidine-HCl dissolved in water as running buffer. Following development, the plates were soaked again in methanol to remove excess salt from the plates prior to autoradiography. Unlabelled

standards were detected by placing the plate under a UV lamp. The R_f-values observed for various compounds were: ATP, 0.06; PAPS, 0.25; ADP, 0.31; AMP, 0.72; and sulfate, 0.81.

For preparative purposes, labelled PAPS was separated from ^{35}S-sulfate by chromatography on ECTEOLA-cellulose (coarse; obtained from Sigma) columns as described by Vargas (7). The column material was prepared by sequential washes with 0.1 N HCl, H_2O, 0.1 N NaOH, H_2O, 0.5 N formic acid, and H_2O (25 ml/g material). Ethanolic supernatants from labelling reactions (described above), which were stored at -20°C, were centrifuged in a clinical centrifuge to remove precipitated material. The supernatants were dried to 500 µl under a stream of N_2. Following the addition of 2 ml 10 mM $(NH_4)HCO_3$, the sample was loaded onto a 2-ml column, prepared in a Pasteur pipet and equilibrated with 20 ml of 10 mM $(NH_4)HCO_3$. Sulfate was washed off the column with 12 ml of 10 mM $(NH_4)HCO_3$ and PAPS was eluted with 2 ml of 0.5 M $(NH_4)HCO_3$. The buffer was removed under vacuum following freezing of the sample in liquid N_2. The material was resuspended in 300 µl of 50 % (v/v) ethanol, dried down under a stream of N_2, and was finally resuspended in 100 µl 0.1 M Hepes-Na buffer (pH 7.5). Since PAPS is very unstable in water, the purified ^{35}S-PAPS was not stored longer than one week at -70°C. No other contaminating, ^{35}S-labelled compounds were detected by TLC immediately following purification. For practical reasons, the specific activity of the ^{35}S-PAPS was assumed to be the same as that of the substrate ^{35}S-sulfate (10 µCi/nmol). It was assumed that no significant amounts of non-radioactive

sulfate were carried over with the enzyme preparation, since the sulfate concentration in the enzyme suspension was reduced about 10^9-fold by repeated dialysis as described above. In principle, it would have been possible to quantify the amount of PAPS in the sample using a very cumbersome method involving an enzymatic transfer reaction described by Vargas (7).

Preparation of concentrated extracts from *R. sphaeroides*. Cells of *R. sphaeroides* were grown photoheterotrophically in 200 ml Sistrom's medium as described in chapter 3. Late log cultures were harvested by centrifugation, and the cells were resuspended in 1.5 ml of 0.1 M Hepes (pH 7.2). Subsequent to the addition of 1000 units of DNAse I, the cells were broken in a French press at 16,000 psi. Using the Bradford micro assay (1), the protein concentration in the extracts was determined to be about 20-30 mg/ml, depending on the amount of cells in the 200 ml culture.

Assay for *in vitro* sulfolipid biosynthesis. To study sulfolipid biosynthesis in extracts of *R. sphaeroides*, reaction mixtures (100 µl) were set up in polypropylene microfuge tubes, containing 0.1 M Hepes (pH 7.2), 1 µCi ^{35}S-PAPS (approximate specific activity 10 µCi/nmol) and 25 µl cell extract (about 0.5 mg protein). Initially, the following compounds were added to circumvent the possible dilution of precursors: 20 mM ATP, 1 mM NADPH, 5 mM $MgCl_2$, 1 mM 3-phosphoglycerol, 1 mM 6-phosphoglucose, and 1 mM 3-phosphoglycerate. The addition of these compounds was later omitted, since they did not affect the synthesis of sulfolipid in the described

experiments. Several mixtures were incubated in parallel at 32°C. The reactions were stopped at different times, usually not later than 12 min following the addition of the substrate, by addition of 800 µl chloroform/methanol (1:1, v/v) on ice. To the extract 260 µl of 0.2 M H_3PO_4, 1 M KCl were added. Mixing, followed by centrifugation, resulted in phase-partitioning of the lipids into the lower chloroform phase. The chloroform phase was spotted onto an ammonium sulfate-treated silica plate, which was activated and run as described in chapter 3. Due to the high lipid concentration in these extracts, the plates were overloaded on some occasions, resulting in smearing or broadening of the bands. Labelled lipids were detected by autoradiography, and the bands were transferred into scintillation cocktail. The radioactivity was determined by scintillation counting. Taking into account the counting efficiency, and based on the specific activity of the substrate and the amount of protein in the sample, rates of incorporation of PAPS into sulfolipid were expressed in some cases as (pmol PAPS incorporated)/(h mg protein).

Preparation and TLC of labelled sulfolipid standard. To identify *in vitro* synthesized sulfolipid on TLC plates, lipid extracts containing ^{35}S-labelled sulfolipid prepared by labelling intact *R. sphaeroides* wild type cells, were run on the same TLC plates as the lipid extracts from the *in vitro* reaction mixtures. The labelled sulfolipid standard was essentially prepared as described in chapter 3. During the preparation of water soluble, ^{35}S-labelled compounds from wild type cells (methods section of chapter 3), a large

amount of labelled sulfolipid could be obtained in the chloroform phase as byproduct, which was used for this purpose. The two TLC systems used were ammonium sulfate plates, prepared and run as described in chapter 3, and regular Baker Si250 silica plates activated for 15 min at 120°C and developed with chloroform/acetone/methanol/acetic acid/water (50:20:10:10:5, v/v).

Deacylation of sulfolipid and analysis of sulfoquinovose glycerol by TLC. Following chromatography of the labelled sulfolipid on TLC plates and detection by autoradiography, the respective bands (about 5000 cpm) were scraped from the plates and the material was transferred to a Pasteur pipet plugged with glass wool. Sulfolipid was eluted from the silica with 2 ml chloroform/methanol (1:1, v/v). The solvent was removed under a stream of N_2 and the lipid was resuspended in 100 µl of the same solvent. Addition of 25 µl 0.2 M H_3PO_4, 1 M KCl resulted in phase-partitioning of the sulfolipid into the chloroform phase, which was transferred to a new tube, and dried with N_2. This material was resuspended in 100 µl chloroform/methanol (2:3, v/v), and 100 µl of 0.2 N NaOH in methanol were added. Following incubation at 37°C for 20 min, 40 µl methanol, 160 µl chloroform, and 180 µl water were added. The aqueous upper phase was transferred to a new tube containing about 100 µl Dowex $50H^+$ beads in methanol/water (10:9, v/v), to neutralize the solution. The beads were spun down and the supernatant was transferred to a new tube. To maximize the yield, the chloroform phase, obtained in the previous step, was washed with 200 µl methanol/water (10:9 v/v), and the resulting upper phase was used to wash the Dowex $50H^+$ beads. The

supernatant solutions from the first treatment and the wash were combined and the solvent was evaporated under a stream of N_2. The material was resuspended with 10 µl methanol/water (10:9) and the sample was spotted into the lower right hand corner of a cellulose plate (0.1 mm thick on plastic, Merck). The plate was developed in the first dimension with water-saturated phenol, and in the second dimension with methanol/formic acid/water (80:13:7, v/v).

Extraction of possible precursors from wild type cells with hot ethanol. To prepare an extract from wild type cells containing small molecules, which could stimulate the *in vitro* sulfolipid biosynthesis, cells from a 200-ml photoheterotrophically grown culture of *R. sphaeroides* (for growth conditions see chapter 3) were sedimented and resuspended in 5 ml 80% (v/v) ethanol, and the sample was incubated at 65°C for 5 min. Following centrifugation in a clinical centrifuge, the supernatant solution was transferred to a new tube, and the pellet was reextracted with 5 mL of 100% (v/v) ethanol for 5 min at 65°C. Following removal of the supernatant the pellet was extracted again with 50% (v/v) ethanol at 65°C for 5 min. The three supernatant solutions were combined and dried to 1 ml at about 50°C under a stream of N_2. Following the addition of 3 ml of chloroform and 1 ml of water, the upper, aqueous phase was transferred to a new tube and dried down under a stream of N_2 at about 50 °C. The material was resuspended in 1 ml of water and stored at -20°C. The pH value in the solution was determined to be pH 7.3.

RESULTS AND DISCUSSION

^{35}S-PAPS is only incorporated into sulfolipid by broken cells, but not by intact cells. Labelled PAPS was chosen as a substrate, because it is the only known precursor for sulfolipid biosynthesis besides sulfate, and has the advantage that it can not be taken up by intact cells. Therefore, it was assumed that any incorporation of label from PAPS into sulfate by extracts of *R. sphaeroides* was truly *in vitro* and could not be explained as being due to the presence of unbroken cells in the extract. This was demonstrated as follows. A cell suspension of *R. sphaeroides* wild type cells was split and one aliquot was run through the French press at 16,000 psi as described in the methods section. Various amounts of broken and intact cells were incubated in parallel with ^{35}S-PAPS. Only in lipid extracts obtained from incubation mixtures containing broken cells was a labelled lipid detected that comigrated with authentic sulfolipid (Figure 6-1). In addition, several labelled compounds that did not comigrate with sulfolipid were detected in reaction mixtures containing broken cells but not in those containing intact cells. The sulfolipid standard shown in Figure 6-1 was actually not purified sulfolipid, but a complete lipid extract obtained from cells labelled *in vivo* with ^{35}S-sulfate. No additional sulfur- labelled bands except the sulfolipid band were visible, indicating that the additional bands observed in extracts from *in vitro* reactions were artifacts. Two conclusions from this experiments were that, since only broken cells could label chloroform-soluble material, including sulfolipid, from PAPS, the labelling reactions must have been *in vitro* and

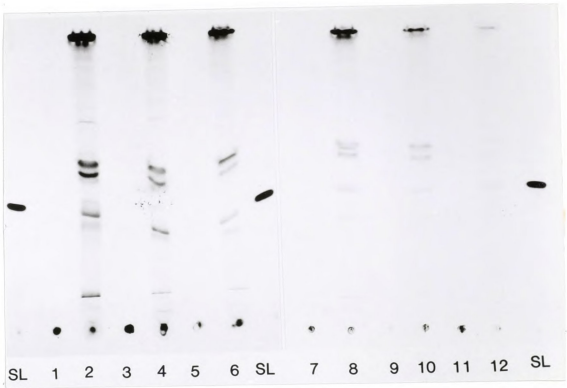


Figure 6-1: Incorporation of ^{35}S-PAPS into sulfolipid by broken and intact wild type cells of *R. sphaeroides*. Shown are lipid extracts obtained from reaction mixtures, separated on ammonium sulfate-treated plates. Labelled lipids were detected by autoradiography. Various amounts of broken cells (lanes with even numbers) or intact cells (lanes with odd numbers) were incubated for 12 min in the presence of ^{35}S-PAPS. Lanes marked with SL show lipid extracts containing authentic sulfolipid obtained from cells labelled with sulfate. The various protein concentrations used in the reactions were (in mg): (lanes 1, 2) 1.8; (lanes 3, 4) 1.5; (lanes 5, 6) 1.2; (lanes 7, 8) 0.9; (lanes 9, 10) 0.6; and (lanes 11, 12) 0.3.

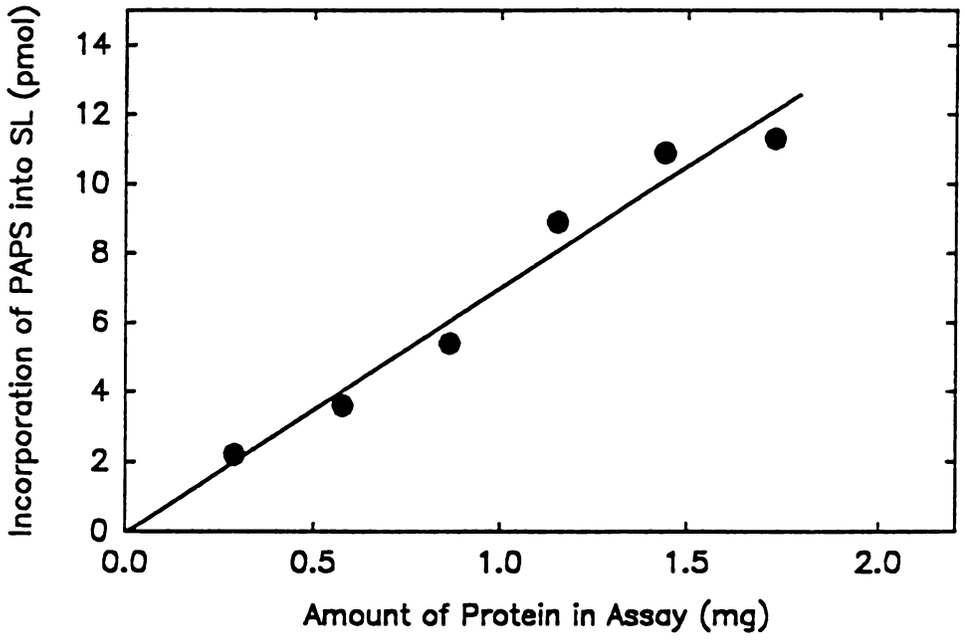


Figure 6-2: Dependence of the incorporation of ^{35}S-PAPS into sulfolipid on the amount of extract present in the reaction mixture. The radioactivity in the putative sulfolipid bands observed in the experiment described in Figure 6-1, was determined, and used to calculate the incorporation of ^{35}S-PAPS into the sulfolipid. The incorporation shown on the Y-axis was plotted against the protein concentration in the reaction mixtures (X-axis).

contaminations of extracts with intact cells would probably not bias the interpretation of results. Attempts to quantify the fraction of intact cells in French press extracts by plating the extract onto plates containing Sistrom's medium and counting of the resulting colonies, were hampered by the fact that the cells had a tendency to clump and colonies derived from single cells could not be obtained, but it was obvious that there was still a substantial amount of viable, unbroken cells present in the extracts.

When the amount of extract, expressed as mg of protein, used in the experiment described above was plotted against the incorporation of PAPS into the putative sulfolipid, a linear relationship was observed (Figure 6-2), indicating that the system was not saturated under the chosen conditions. The amounts of extract used in the reactions, ranging from 0.3 mg of protein to 1.8 mg (Figure 6-1), were relatively high. However, since it was expected that several unknown enzymatic reactions and intermediates were involved to catalyze the incorporation of PAPS into sulfolipid, the extracts were kept as concentrated as possible to avoid the dilution of intermediates.

Characterization of the sulfolipid labelled *in vitro*. To confirm that the lipid labelled from ^{35}S-PAPS by extracts of *R. sphaeroides* wild type cells, which seemed to comigrate with authentic sulfolipid (Figure 6-1), indeed represented sulfolipid, a series of experiments was performed. When extracts obtained from the sulfolipid deficient *sqdB* mutant (see chapter 3) and extracts from wild type cells were incubated in the presence of ^{35}S-PAPS, no band comigrating with the sulfolipid standard was observed in mutant extracts, while

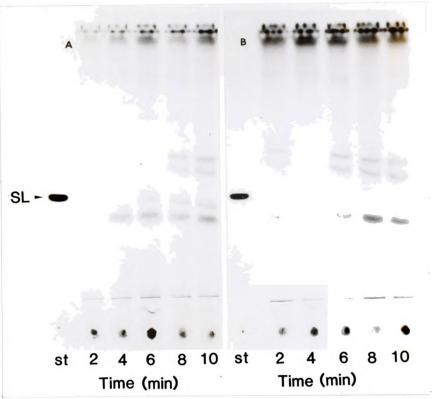


Figure 6-3: Comparison of extracts from wild type and sulfolipid deficient *sqdB* mutant cells for their capability to label sulfolipid from ^{35}S-PAPS. (A) wild type cell extracts, and (B) *sqdB* mutant cell extracts were used in the reaction mixtures. Lipid extracts from *in vitro* reaction mixtures separated on ammonium sulfate treated-plates are shown. The lipids were detected by autoradiography. Lanes marked (st) show *in vivo* labelled sulfolipid. Lipids were extracted from the reaction mixtures at various times as indicated.

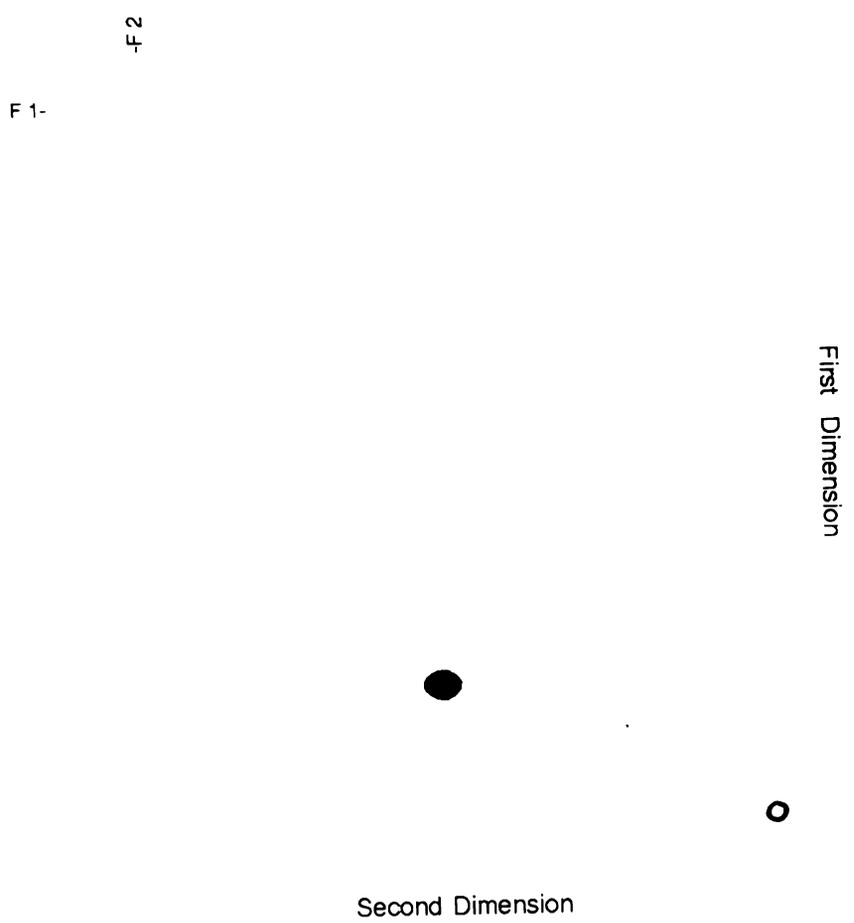


Figure 6-4: Cochromatography of deacylated sulfolipid labelled *in vitro* and *in vivo*. Presumptive sulfoquinovose-glycerol prepared from *in vitro* (2000 cpm) and *in vivo* labelled sulfolipid (4000 cpm) were mixed together and analyzed in a two dimensional TLC-system as described in the method section. The open circle marks the origin. The two solvent fronts are indicated in the upper left corner (F1, solvent front of the first dimension; F2, solvent front of the second dimension).

the same number and intensity of artifactual bands was observed as for the wild type extracts (Figure 6-3). This result provided strong evidence that wild type extracts were indeed capable of synthesizing sulfolipid from labelled PAPS, and that the labelled compounds, only observed in *in vitro* labelling experiments, were most likely unrelated to sulfolipid biosynthesis. The nature of these compounds, specific to the *in vitro* system, has not been further analyzed. However, invoking the analogy to a different *in vitro* system for studying sulfolipid biosynthesis, in lipid extracts from isolated spinach chloroplasts, which were labelled with ^{35}S-sulfate, a labelled, chloroform-soluble compound was found in addition to sulfolipid (4). This compound was not observed in extracts obtained from labelled leaves. Analysis of this compound identified it as elemental sulfur which is chloroform-soluble (4). With this result in mind, it was assumed that the labelled compounds observed besides sulfolipid in lipid extracts which, were prepared from the *in vitro* reaction mixtures employing broken cells of *R. sphaeroides*, represented some form of reduced sulfur multimers. However, further experiments are required to test this idea.

The identity of the putative sulfolipid labelled in *in vitro* reactions from ^{35}S-PAPS was further confirmed by comigration of the isolated lipid from the respective TLC-band with *in vivo* labelled sulfolipid in two different solvent systems. In addition, *in vivo* and presumptive *in vitro* labelled sulfolipid were deacylated, and the two resulting preparations, were found to comigrate in a two dimensional TLC-system, when spotted together in one mixed sample

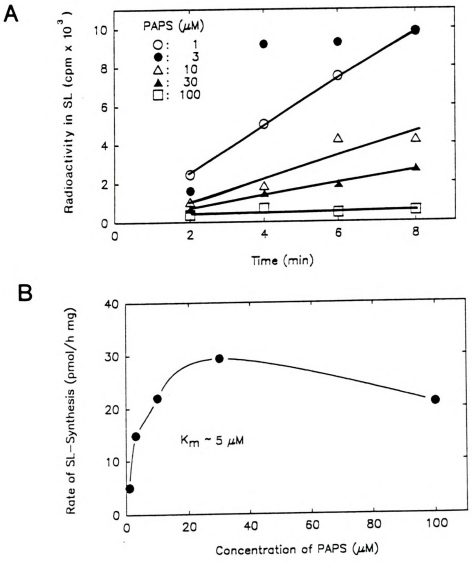


Figure 6-5: Substrate saturation for the incorporation of ^{35}S-PAPS into sulfolipid. (A) Apparent rates of sulfolipid biosynthesis for various total concentrations of PAPS as indicated. The radioactivity in the sulfolipid bands as determined by scintillation counting was plotted against the incubation time. The amount of protein used in each reaction mixture was 0.85 mg. (B) Absolute rates of PAPS incorporation into sulfolipid, calculated on the basis of data shown in (A), were plotted against the concentration of PAPS in the reaction mixtures (Michaelis-Menten plot). The substrate concentration, at which the process is half saturated, is indicated (k_M).

(Figure 6-4).

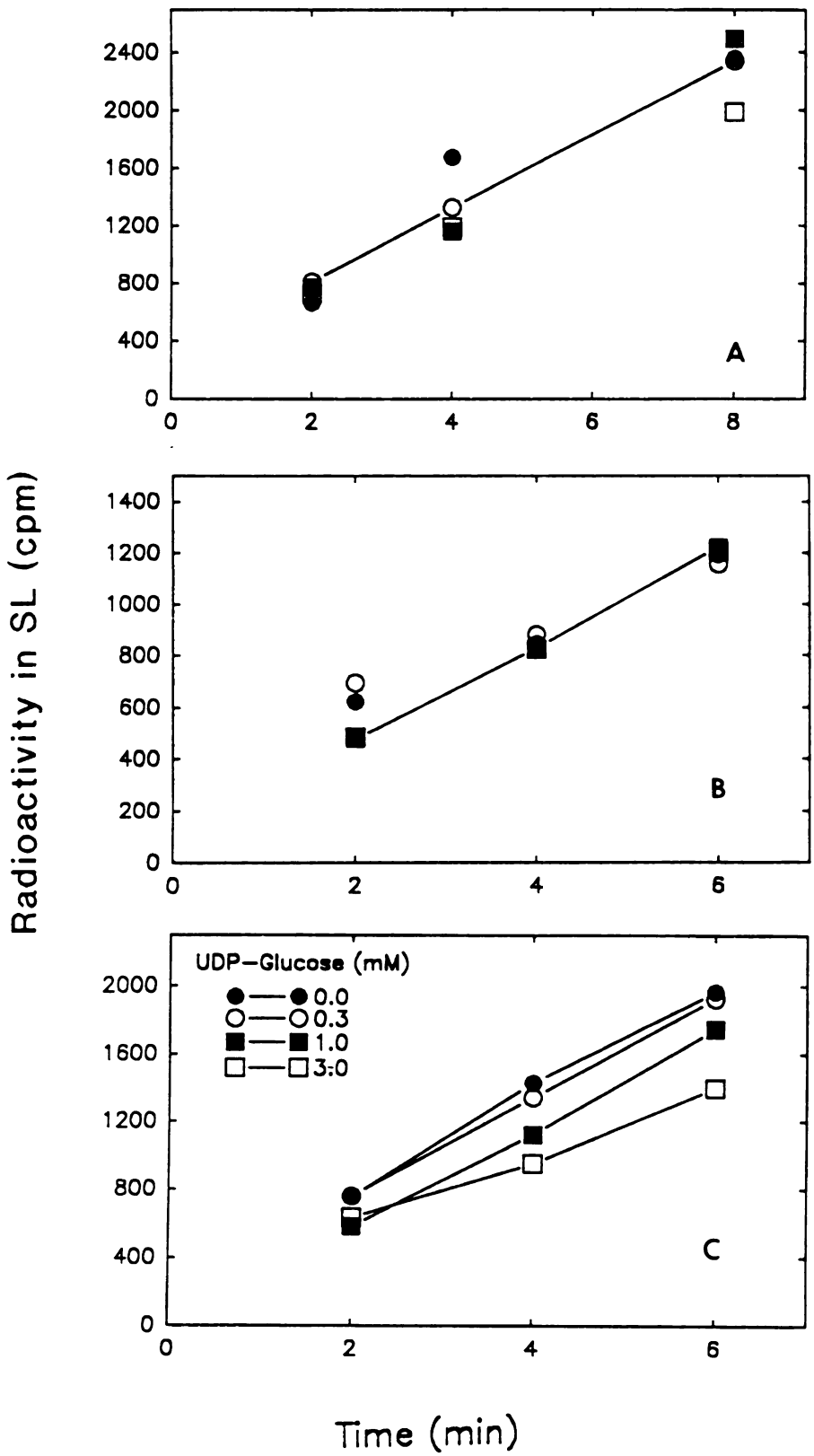
Saturation of the incorporation of ^{35}S-PAPS into sulfolipid by addition of unlabelled PAPS. To determine, whether PAPS behaved as a legitimate substrate for an enzymatic reaction in the *in vitro* label experiments, the rate of incorporation of ^{35}S-PAPS into sulfolipid was determined at a range of substrate concentrations. The total amount of PAPS in the reaction mixture was varied while the amount of label (1 μCi) per reaction was kept constant. The apparent rates, measured as increase in radioactivity in the respective sulfolipid bands isolated from TLC-plates over time, were found to be linear, and stable PAPS could compete effectively with labelled PAPS under the chosen conditions (Figure 6-5, A). From the apparent rates, absolute rates for the incorporation of PAPS into sulfolipid were calculated and plotted against the total concentration of PAPS present in the reaction mixtures (Michaelis-Menten plot; Figure 6-5, B).

Based on this experiment, the maximal rate of sulfolipid biosynthesis from PAPS in extracts from *R. sphaeroides* was estimated to be about 25-30 pmol/h mg protein, and the process was half saturated at PAPS concentrations of about 5 μM (k_M in Figure 6-5, B). Unfortunately, no data for the rate of sulfolipid biosynthesis *in vivo* in *R. sphaeroides* were available for comparison. The value of 5 μM for the substrate concentration at which the process was found to be half saturated, fell well into the range of known k_M constants for various enzymes. This result confirmed that PAPS behaved as a legitimate substrate for an enzymatic reaction. This result was not in

agreement with a nonenzymatic, artifactual incorporation of label from PAPS into sulfolipid by extracts of *R. sphaeroides*.

Effect of various compounds added to the reaction mixtures. In a series of exploratory experiments, various compounds were added to the *in vitro* reaction mixtures, to screen for those which could either stimulate or inhibit sulfolipid biosynthesis. First of all, to stimulate the reactions in the extract, a basic mixture of metabolic intermediates commonly found in cells was added, as well as a concentrated, aqueous extract from wild type cells, prepared by extraction of cells with hot ethanol (see methods section), which was thought to contain metabolic intermediates possibly required for the reactions involved in sulfolipid biosynthesis. The basic metabolite mixture contained (final concentrations in the reaction mixture): 1 mM ATP, 1 mM NADPH, 5 mM $MgCl_2$, 1 mM glycerol-3-phosphate, 1 mM glucose-6-phosphate, and 1 mM 3-phospho-glycerate. Neither the addition of these compounds, the metabolite extract from the wild type cells, or any combination of these was able to stimulate the rate of incorporation of ^{35}S-PAPS into sulfolipid (Figure 6-6, A). Additions of the listed compounds individually also had no effect. It was concluded that either the availability of intermediates in the highly concentrated, crude cell extracts used was not the limiting factor under the chosen conditions, or that the enzymes involved in sulfolipid were organized in a complex resulting in metabolite channeling. Some of the Calvin cycle enzymes are thought to form such a complex (2). It is assumed that the respective metabolites do not freely exchange with the surrounding medium,

Figure 6-6: Effect of the addition of various compounds on the incorporation of ^{35}S-PAPS into sulfolipid. (A) Addition of a metabolite mixture as mentioned in the text, and of a concentrated extract (25 µl/reaction) from wild type cells prepared as described in the method section. Symbols: closed circles, no additions; open circles, extract from wild type cells; closed squares, metabolite mixture; open squares, metabolite mixture and extract. Aliquots were incubated for 2, 4, and 8 min prior to analysis of the reaction products. (B) Effect of different concentrations of DL-glycerate in the reaction mixture. Symbols: closed circles, 0 mM; open circles, 1 mM; closed squares, 3 mM; and open squares 10 mM of DL-glyceraldehyde present in the reaction mixture. The reaction products were analyzed after 2, 4, and 6 min of incubation. (C) Effect of the addition of UDP-glucose. The concentrations in the reaction mixtures are indicated in the Figure. Reaction products were analyzed after 2, 4, and 6 min of incubation. The amount of radioactivity found in the sulfolipid band on TLC plates as determined by scintillation counting is indicated on the y-axis for all three data sets. In all three experiments 10 µM concentrations of PAPS were used per reaction mixture (specific activity about 1 µCi/nmol).


Radioactivity in SL (cpm)
2400
2000
1600
1200
800
400
0
0
2
4
6
8
A
1400
1200
1000
800
600
400
200
0
0
2
4
6
B
UDP–Glucose (mM)
0.0
0.3
1.0
3.0
2000
1600
1200
800
400
0
0
2
4
6
C
Time (min)

but instead are passed along from reaction to reaction by the enzymes in the complex. Such a process would imply that intermediates could not be easily fed to the enzymes. Based on this result, the metabolite mixture mentioned above, which was added to the reaction mixtures in the initial experiments, was omitted in later experiments.

Two compounds, DL-glyceraldehyde and UDP-glucose, were added to the reaction mixtures, to test whether they could inhibit the incorporation of label from PAPS into sulfolipid. DL-glyceraldehyde is a competitive inhibitor of the aldolase in the Calvin cycle (6). This, or a similar enzyme was proposed to catalyze the synthesis of sulfoquinovose-1-phosphate from sulfolactaldehyde and dihydroxyacetone-3-phosphate (see chapter 1, Figure 1-2). However, in one exploratory experiment, addition of DL-glyceraldehyde up to 10 mM did not inhibit the incorporation of ^{35}S-PAPS into sulfolipid (Figure 6-6, B). On the other hand, the addition of 3 mM UDP-glucose resulted in slight reduction of the rate of incorporation of PAPS into sulfolipid (Figure 6-6,C). The sugar nucleotide UDP-glucose was tested, because it resembles UDP-sulfoquinovose, which was thought to be a precursor of sulfolipid biosynthesis (chapter 1). No firm conclusions should be drawn from a single, exploratory experiment. However, future experiments may show that UDP-glucose plays a role as competitive inhibitor or substrate of sulfolipid biosynthesis in *R. sphaeroides*.

In extracts of *Chlamydomonas*, the only other *in vitro* system described until now (3), labelled sulfite could be incorporated into sulfolipid. However,

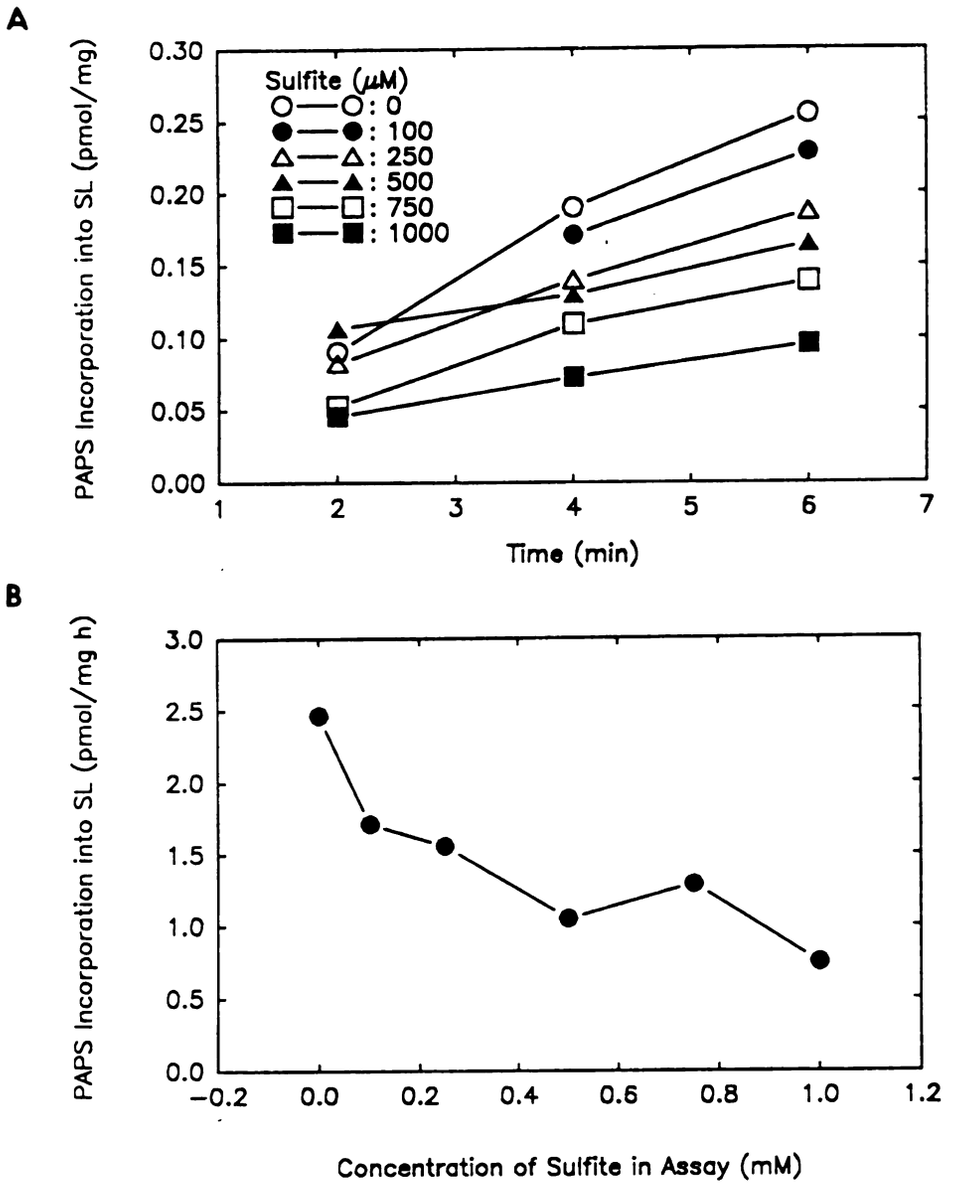


Figure 6-7: Effect of sulfite on the incorporation of ^{35}S-PAPS into sulfolipid. (A) Rates of sulfolipid synthesis from PAPS in the presence of various concentrations of sulfite as indicated in the Figure. The total concentration of PAPS in the reaction mixtures was 10 μM (specific activity 1 μCi/nmol). Samples were analyzed after 2, 4, and 6 min of incubation in the presence of labelled substrate. The rates were calculated from the radioactivity found in the sulfolipid bands upon separation of the reaction products by TLC. (B) Inhibition of sulfolipid biosynthesis by sulfite. The rates of sulfolipid synthesis calculated from the data shown in (A), were plotted against the concentration of sulfite in the reaction mixtures.

Table 6-1: Effect of sulfur containing compounds and aspartic acid on the incorporation of ^{35}S-PAPS into sulfolipid. Samples were extracted after 2, 4, and 6 min of incubation in the presence of ^{35}S-PAPS. The rates, presented as pmol PAPS incorporated into sulfolipid per hour and mg protein, were calculated from the radioactivity associated with the sulfolipid bands detected upon separation of the reaction products by TLC. The concentration of PAPS in the reaction mixtures was 10 μM (specific activity 1 μCi/ nmol). The results shown were obtained from a single experiment, and therefore have to be considered preliminary.

Compound into	Concentration (mM)	Rate of PAPS Incorporation Sulfolipid (pmol/h mg protein)
control	1	4.3
cysteine	1	3.2
cysteic acid	1	3.2
aspartic acid	1	2.1
reduced glutathione	1	3.7
dithiothreitol	1	1.1

the process could not be saturated, indicating that sulfite was not behaving as a legitimate substrate for an enzymatic reaction. To test the effect of sulfite in the *R. sphaeroides* system, varying amounts of sulfite were added to the reaction mixtures and its effect on the incorporation of ^{35}S-PAPS into sulfolipid was examined. Addition of 1 mM sulfite resulted in 5-fold reduction of the rate of sulfolipid biosynthesis from PAPS. It was concluded that this inhibitory effect of sulfite, which was observed repeatedly in different experiments, could be due to the fact that sulfite was either an intermediate in the pathway of sulfolipid biosynthesis from PAPS, or alternatively that sulfite was poisoning the system in an unexplained way.

Addition of 1 mM DTT repeatedly inhibited the incorporation of PAPS into sulfolipid, while 1 mM concentrations of cysteine, cysteic acid, and reduced glutathione, tested only once in a preliminary experiment (Table 6-1), showed no pronounced inhibitory effect. Cysteic acid was proposed to be a possible intermediate of sulfolipid biosynthesis (compare chapter 1), but had very little effect on sulfolipid biosynthesis in cell extracts. Based on this single experiment, it seemed likely that cysteic acid was not a precursor of sulfolipid biosynthesis in *R. sphaeroides*. However, further experiments are required to confirm the preliminary data. Aspartic acid, which resembles cysteic acid in its structure and physical properties, showed some inhibition of sulfolipid biosynthesis in extracts (Table 6-1). Glutathione was reported to stimulate *in vitro* sulfolipid biosynthesis in extracts from *Chlamydomonas* (3), but appeared to have little effect in the *R. sphaeroides* system (Table 6-1).

CONCLUSIONS

The series of experiments described above demonstrated that crude extracts from *R. sphaeroides* cells are capable of catalyzing the incorporation of label from ^{35}S-PAPS into sulfolipid. Amongst other unidentified, chloroform-soluble compounds, which were believed to be an artifact of the *in vitro* system, sulfolipid could be identified. The evidence for the identity of sulfolipid was threefold: first, extracts isolated from the sulfolipid deficient *sqdB* mutant could not label the putative sulfolipid from ^{35}S-PAPS (Figure 6-3); second, the putative sulfolipid band comigrated with authentic sulfolipid labelled from ^{35}S-sulfate *in vivo* in two different TLC-systems; and third, the deacylation product of *presumptive in vitro* and *in vivo* labelled sulfolipid comigrated in a two dimensional TLC-system (Figure 6-4). The fact that only broken but not intact cells could label sulfolipid from ^{35}S-PAPS (Figure 6-1) confirmed that unbroken cells contaminating the extracts could not be responsible for the observed synthesis of sulfolipid, and supported the idea that the reactions indeed proceeded *in vitro*. Dependence of the process on the extract (protein) concentration (Figure 6-2) as well as substrate saturation (Figure 6-5; A, B) indicated that ^{35}S-PAPS behaved as a legitimate substrate for a series of enzymatic reactions involved in sulfolipid biosynthesis and excluded the artifactual, chemical incorporation of label from ^{35}S-PAPS into sulfolipid. Maximal rates of 20-30 pmol/h mg protein were observed for the synthesis of sulfolipid from PAPS (Figure 6-5, B). Under suboptimal substrate concentrations mostly used during the experiments to avoid the dilution of the

label with non-radioactive substrate, rates of about 3 pmol/h mg protein were observed (Figure 6-7, B, Table 6-1). In a series of preliminary experiments, it was determined that sulfite (Figure 6-7; A, B) and DTT (Table 6-1) inhibited sulfolipid biosynthesis from PAPS. A substrate analog of aldolase, DL-glyceraldehyde, had no effect (Figure 6-6, B), while addition of UDP-glucose, a compound similar to UDP-sulfoquinovose, inhibited the process slightly (Figure 6-6, C). Cysteic acid, proposed as a possible precursor of sulfolipid, had little effect, as did cysteine and glutathione (Table 6-1). Aspartic acid, on the other hand, which chemically resembles cysteic acid, appeared to reduce the rate of sulfolipid biosynthesis. However, these preliminary should be used only as guidance for future research, rather as a basis for conclusions.

The synthesis of sulfolipid in cell extracts was not stimulated by the addition of metabolites such as ATP , NADPH, and various sugar phosphate esters (Figure 6-6, A). This result was not surprising, since highly concentrated cell extracts were used in the reactions, which might have provided all the required metabolites in non-limiting concentrations. Serious attempts to purify individual components of the system have not yet been undertaken. Given the limited information available about the nature of precursors of sulfolipid biosynthesis, and the unavailability of labelled, putative precursors, the resulting approach of trial and error seems to be cumbersome and time consuming. However, with the isolation of sulfolipid deficient mutants of *R. sphaeroides* described in this study, which were shown to accumulate

possible precursors, the analytical process of dissecting the components of sulfolipid biosynthesis assumes a new dimension. The next step should be to isolate and purify these compounds from the mutants and to test them as precursors in the *in vitro* system described in this chapter. Information gained from structural elucidation of these compounds will be useful to design experiments involving new substrates and inhibitors of sulfolipid biosynthesis, which can be tested in the *in vitro* system described. Therefore, the data presented in this chapter provide a crucial cornerstone for further analysis of the pathway of sulfolipid biosynthesis in *R. sphaeroides*, and ultimately in higher plants.

REFERENCES

1. Bradford M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-252

2. Gontero M., Cardenas M.L., Ricard J. (1988). A functional five-enzyme complex of chloroplasts involved in the Calvin cycle. Eur. J. Biochem. 173:437-443

3. Hoppe W., Schwenn J.D. (1981). *In vitro* biosynthesis of the plant sulpholipid: on the origin of the sulphonate head group. Z. Naturforsch. 36c:820-826

4. Joyard J., Forest E., Blee E., Douce R. (1988). Characterization of elemental sulfur in isolated intact spinach chloroplasts. Plant Physiol. 88:961-964

5. Robbins P.W. (1962). Sulfate activating enzymes. Methods Enzymol. 5:964-977

6. Stokes D.M., Walker D.A. (1972). Photosynthesis by isolated chloroplasts. Inhibition by DL-glyceraldehyde of carbon dioxide assimilation. Biochem. J. 128:1147-11573.

7. Vargas F. (1988). Preparation and quantification of 3'-phosphoadenosine 5'-phospho[^{35}S]sulfate with high specific activity. Anal. Biochem. 172:82-88

CHAPTER 7

CONCLUSIONS AND PERSPECTIVES

Since the discovery and structural elucidation of the plant sulfolipid by Benson over thirty years ago (2), relatively little progress has been made towards the elucidation of the biosynthetic pathway and the function of this unusual lipid. This is surprising, since the occurrence of substantial amounts of the plant sulfolipid in all photosynthetic organisms examined make it one of the most abundant, sulfur-containing organic compounds on earth. Therefore, this compound plays an important role in the global sulfur cycle and is a substantial ingredient of the human and animal diet. Lack of progress in our understanding of sulfolipid biosynthesis and function cannot be blamed entirely on lack of interest by the scientific community, since a number of researchers tried to solve the problem during the past thirty years. Among the reasons for slow progress might have been the fact that our general knowledge about sulfur metabolism in plants is not very deep. Not much knowledge can be derived from animal systems, since mammals are not capable of assimilating inorganic sulfur. This process is unique to plants and bacteria and must, therefore, be studied in these organisms. Studying sulfur metabolism in general seems to be very complicated, due to the diversity and reactivity of sulfur-containing compounds. Sulfolipid biosynthesis brings together two plant-specific processes: first, photosynthetic sugar biosynthesis and second, photosynthetic sulfur assimilation. At present, we have quite a deep

understanding of sugar biosynthesis in plants, and many of the ideas about sulfolipid biosynthesis, as outlined in chapter 1, were derived from our knowledge about this process. The elucidation of the pathway for the sugar biosynthesis took many years and was facilitated by the fact that most of the intermediates are highly abundant in photosynthetically active leaves. Intermediates of sulfolipid biosynthesis, on the other hand, are less abundant. Given these facts, it becomes clear that the study of sulfolipid biosynthesis provides an intellectual challenge, and that novel ideas are required to resolve this problem.

The objective of the study presented here was to develop an experimental system which would facilitate a solution of the classical problem of sulfolipid biosynthesis. A genetic approach to dissect the components of sulfolipid biosynthesis had never been tried before. In classical studies, complicated biosynthetic pathways, such as the biosynthesis of histidine, were elucidated with the help of bacterial mutants, which were blocked at various steps of the pathway (1, 4). It was assumed that studying mutants that showed blocks in the pathway of sulfolipid biosynthesis would provide the means to elucidate the pathway for sulfolipid biosynthesis. The organism of choice for such a study was the purple non-sulfur bacterium *R. sphaeroides*. It contained the plant sulfolipid with its typical 6-sulfo-6-deoxy hexose head group, as was confirmed by comparative mass spectroscopy of isolated sulfolipid from spinach and *R. sphaeroides*. As described in chapter 2, a characteristic, comparable fragmentation pattern for the sulfolipid head group from spinach

and *R. sphaeroides* could be observed employing negative FAB-MS/MS techniques.

The purple non-sulfur bacterium *R. sphaeroides* had distinct advantages over other organisms for a genetic approach, since common genetic techniques were available for this bacterium, and, more importantly, this organism was able to grow in the absence of photosynthesis. Although the function of sulfolipid is not clearly established, there was considerable concern that obligately photoautotrophic organisms, such as higher plants or many cyanobacteria, would not be able to grow in the absence of sulfolipid, causing lethality of possible sulfolipid-deficient mutants.

Using a novel screening procedure, the examination of lipid extracts from chemically mutagenized *R. sphaeroides* cells by TLC made it possible to isolate a series of mutants deficient in sulfolipid biosynthesis (chapter 3). The analysis was focused on three mutants (*sqdA, sqdB,* and *sqdC*), two of which accumulated ^{35}S-labelled, water-soluble compounds that may be precursors of sulfolipid biosynthesis. Structural analysis of these compounds as well as the elucidation of their possible role as precursors in sulfolipid biosynthesis, will provide the crucial data to establish a pathway.

However, the next logical step in developing the *R. sphaeroides* model system for the study of sulfolipid biosynthesis, following the isolation and phenotypic characterization of the sulfolipid deficient mutants (Chapter 3), was thought to be the isolation and characterization of the genes complementing the mutations. All strains containing the mutations causing the sulfolipid

deficient phenotype contained background mutations due to the extensive chemical mutagenesis process. To determine that the observed phenotype of the mutants such as accumulation of compounds or growth changes as compared to wild type cells was due to the mutation causing the sulfolipid deficiency, isogenic lines had to be created. This is normally accomplished either by gene replacement of the respective wild type gene with a defective copy, or by comparing the mutants to the same mutant strain carrying the complementing gene on a single-copy plasmid. Both approaches required the isolation of the respective genes. In addition, it was assumed that the isolation and analysis of the genes would provide information about the number of complementation groups involved and the structure and possible function of the respective gene products.

To isolate the genes complementing the different mutants, a cosmid library was constructed and a complementation assay was devised, through which complementing cosmids were identified (Chapter 4). By screening the cosmid library, two cosmid classes were identified, one of which was able to complement the *sqdB* and *sqdC* mutants, while the second complemented the *sqdA* mutant. From these two cosmid classes distinct subclones were isolated, which were able to complement each of the three mutants. Sequence analysis of the clone complementing the *sqdA* mutant revealed an open reading frame which could encode a 33.6-kD protein. Further deletion and complementation analysis, and localization of a ribosome binding site in front of the open reading frame allowed the identification of the putative *sqdA* gene (Chapter 4). The

analysis was facilitated by the construction of an expression vector for *R. sphaeroides* (Appendix B), which allowed the orientation-dependent expression of reading frames to determine the coding strand for the *sqdA* gene. The role of the *sqdA* gene product in sulfolipid biosynthesis could not be deduced from homology of the translated protein sequence with protein sequences of known enzymes. Further experiments will be required to elucidate the function of the *sqdA* gene product. Analysis of the compounds accumulating in the *sqdA* mutant seems to be the most promising strategy to accomplish this goal. The availability of the *sqdA* gene will make it possible to construct isogenic lines, which are crucial for the analysis of the precursor role of the accumulating compounds, and the elucidation of the function of sulfolipid by comparative, physiological experiments. In addition, expression of the *sqdA* gene in *E. coli* and production of antibodies against the purified product will provide further tools to study the role of the *sqdA* gene in sulfolipid biosynthesis.

The putative gene complementing the *sqdB* mutant which was found to be located within a different set of cosmids than the *sqdA* gene, was similarly analyzed. The putative *sqdB* open reading frame, identified by sequence analysis of a small complementing subclone derived from the complementing cosmid, could encode for a 46-kD protein (Chapter 5). The identity of this open reading frame with the *sqdB* gene was further confirmed by deletion and complementation analysis using a novel *R. sphaeroides* expression vector (Appendix B). Most interestingly, The translated amino acid sequence for the putative *sqdB* gene product showed N-terminal homology with UDP-glucose

epimerase from bacteria and fungi. This result was thought-stimulating, because UDP-glucose resembles UDP-sulfoquinovose, a proposed precursor of sulfolipid biosynthesis. Based on this homology, it was suggested that the *sqdB* gene product might be involved in sulfolipid biosynthesis at the level of UDP-sulfoquinovose or another nucleotide sugar (Chapter 5).

The *sqdC* gene was found to be complemented by the same cosmid that complemented the *sqdB* gene. A subclone derived from this cosmid which contained two inserts was able to complement this mutant (Chapter 5). Southern analysis indicated that parts of this subclone were located directly downstream of the *sqdB* gene. A fragment of the cosmid downstream of and partially overlapping the *sqdB* gene, was found to complement the *sqdC* mutant in an orientation-dependent fashion when cloned in two orientations into an expression vector. The conclusion drawn from this result was that the *sqdC* gene is located directly downstream of the *sqdB* gene, and that its direction of expression coincides with that assumed for the *sqdB* gene. Based on this observation, it is proposed that the *sqdB* and *sqdC* genes are organized in an operon, coding for genes involved in sulfolipid biosynthesis. However, this idea has to be confirmed by further experiments. The *sqdA* gene on the other hand, does not seem to be closely linked to the other two genes and may therefore belong to a second operon or gene cluster involved in sulfolipid biosynthesis. Using a recently established, physical map of the complete *R. sphaeroides* genome (5, 6), it should be possible to determine the relative location of the two sets of genes on this map by Southern hybridization.

Most likely, there are more genes involved in sulfolipid biosynthesis than the three which have been described in this study. Some of the putative mutants which were isolated during this study (Chapter 3), but have not been further analyzed, may mark additional genes involved in sulfolipid biosynthesis. A detailed analysis of the DNA sequences surrounding the two gene clusters may reveal novel *sqd* genes, since genes catalyzing a metabolic sequence frequently occur in operons or clusters in bacteria. Using the cosmid DNA, it should be possible to construct large deletions in the genome of *R. sphaeroides* by gene replacement techniques, covering the sequences surrounding the known *sqd* genes. Cosmids mutagenized *in vitro* could than be analyzed in *trans* for their capability to complement the deleted genes, thereby possibly revealing the location of novel *sqd* genes. Alternatively, a series of gene replacements spanning the region of the cosmid could be constructed.

Finally, to elucidate the pathway for the biosynthesis of the plant sulfolipid, the genetic analysis has to go hand-in-hand with biochemical analysis. The clue to solving the pathway seems to be the analysis of the compounds accumulating in the *sqdA* and *sqdC* mutants. To determine their precursor function for sulfolipid biosynthesis, an *in vitro* sulfolipid synthesis system using extracts from wild type *R. sphaeroides* cells was established (Chapter 6). Although the system is still crude and not very well developed, it can be useful to test the capability of the labelled compounds accumulating in the mutants, thereby establishing their precursor function. Experiments with intact cells would not necessarily be able to provide conclusive results, since it

is unknown whether these compounds can be taken up by the cells. In addition, inhibitors and putative substrate metabolites can be tested *in vitro*, unaffected by problems of uptake. The *in vitro* system described in Chapter 6, provides the basis for the biochemical dissection of the pathway for sulfolipid biosynthesis.

To summarize the findings of this study, one can say that *R. sphaeroides* was established as a genetic model system for the study of sulfolipid biosynthesis. First, sulfolipid deficient mutants were obtained and characterized; second, three genes were isolated, which seem to be directly involved in sulfolipid biosynthesis *in R. sphaeroides*; and third, the biochemical basis for solving the pathway of sulfolipid biosynthesis has been provided by establishing an *in vitro* sulfolipid synthesis system. Based on the data and materials provided by this study, a series of experiments can be designed, as pointed out throughout the text, which will provide avenues to elucidate the pathway for sulfolipid biosynthesis in *R. sphaeroides* in the near future. Solving the structure of the novel compounds accumulating in the mutants and determining their role in sulfolipid biosynthesis seems to be the key to the solution of the problem. Efforts to purify and characterize these compounds have just started.

It is anticipated that exploration of the described genetic model system will invigorate the field of sulfolipid research. The recent finding that sulfoquinovose diacylglycerol isolated from cyanobacteria shows AIDS-antiviral action (3) may increase interest in this compound in the future. It may turn out

that the study of plant sulfolipid not only represents an intellectual challenge for scientists, but also has an application in everyday life.

REFERENCES

1. Ames B.N., Martin R.G., Garry B.J. (1961). The first step in histidine biosynthesis J. Biol. Chem. 236:2019-2026

2. Benson A.A. (1963). The plant sulfolipid. Adv. Lipid Res. 1:387-394

3. Gustafson K.R., Cardellina II J.H., Fuller R.W., Weislow O.S., Kiser R.F., Snader K.M., Patterson G.L.M., Boyd M.R. (1989). AIDS-antiviral sulfolipids from cyanobacteria. J. Natl. Canc. Inst. 81:1254-12583.

4. Smith D.W.E., Ames B.N. (1964). Intermediates in the early steps of histidine biosynthesis. J. Biol. Chem. 239:1848-1855

5. Suwanto A., Kaplan S. (1989). Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: genome size, fragment identification, and gene localization. J. Bacteriol. 171:5840-5849

6. Suwanto A., Kaplan S. (1989). Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. J. Bacteriol. 171:5850-5859

APPENDIX A

THE EFFECT OF SULFOLIPID DEFICIENCY ON GROWTH OF *R. SPHAEROIDES*

With the isolation of sulfolipid-deficient mutants of *R. sphaeroides* and their complementing genes, it became possible to create isogenic lines, that would allow a direct test of the role of sulfolipid. In a preliminary experiment two isogenic lines of *R. sphaeroides* containing different amounts of sulfolipid in their membranes were compared for their ability to grow photoheterotrophically. The sulfolipid-deficient *sqdB* mutant (strain CHB16) carrying the wild type *sqdB gene* on a plasmid (pCHB16201; Chapter 5, Figure 5-7), which restored the wild type phenotype with respect to the sulfolipid content, was compared to the *sqdB* mutant carrying the vector plasmid alone (pCHB500; Appendix B), resulting in a sulfolipid-deficient strain (Figure A-1). These two strains were carrying the same mutations in their genome, except that one of the strains contained the wild type *sqdB* gene in *trans*, while the other did not. Hence any phenotypic differences observed, should be due to the absence or presence of an intact copy of the *sqdB* gene. To examine whether the plant sulfolipid plays a role in photosynthesis (see Chapter 1 for a complete discussion of sulfolipid function) the two strains were incubated under conditions which would allow them to grow photoheterotrophically. For this purpose, 8 ml of Sistrom's medium, containing 0.8 μg/ml tetracycline for selection of the plasmids, were inoculated with the

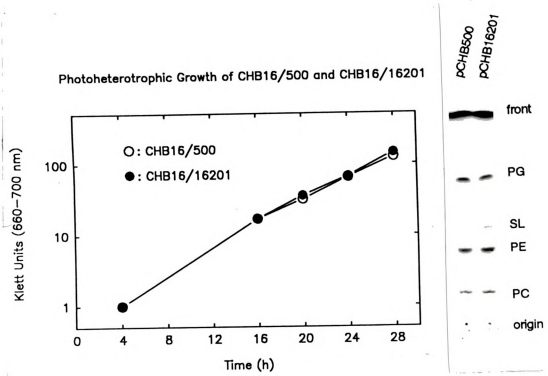


Figure A-1: Comparison of two isogenic lines of *R. sphaeroides*, containing different amounts of sulfolipid, for their heterotrophic growth. Two lines of the sulfolipid deficient *sqdB* mutant were compared, one containing the pCHB500 vector (CHB16/500) and the other, carrying the wild type *sqdB* gene on the plasmid pCHB16201 (CHB16/16201). Lipid extracts obtained from the two strains, separated on ammonium sulfate-treated silica plates are shown on the right. PG, phosphatidyl glycerol; SL, sulfolipid; PE, phosphatidyl ethanolamine; PC, phosphatidyl choline. Growth curves for both strains, obtained during logarithmic growth, are shown on the left. The cultures were grown as described in the text.

two strains described above. Screw-cap tubes were used, which would fit into a Klett photometer equipped with a filter of 660-700 nm cutoff. The tubes were percolated with a steady stream of a gas mixture of 5% CO_2 and 95% N_2, and the cultures were incubated in a water bath at 30°C under a 60 Watt incandescent bulb (100 μE/s per m^2). The growth of the cultures was monitored at various times by measuring the absorption with the aid of a Klett photometer. Replicates for each strain were monitored, and were not found to be very variable.

The result of the experiment described (Figure A-1) indicated that the two strains showed indistinguishable growth rates under the chosen conditions, despite drastic differences in their sulfolipid content. However, it would be premature to conclude from this experiment that sulfolipid plays no role in photosynthesis. As indicated in Chapter 3, the *sqdB* mutant (strain CHB16) was found to be leaky, and the residual amounts of sulfolipid present in the membranes of the *sqdB* mutant may be sufficient to maintain a vital role of sulfolipid in the *sqdB* mutant. To clarify this point, it will be necessary to create a null mutation by gene replacement. If the resulting mutant still grows with the same rate as the wild type under the conditions described, this may indicate that sulfolipid is not important for photosynthetic growth in *R. sphaeroides*. If in such an experiment a difference between the wild type and the mutant were to be observed, a comparative physiological and biochemical study could be initiated to study the function of plant sulfolipid.

APPENDIX B

CONSTRUCTION OF THE EXPRESSION VECTOR pCHB500

Crucial for the analysis of the coding sequence for the *sqdA* and *sqdB* genes as described in Chapters 4 and 5 was the construction of a vector which would allow the construction of a transcriptional fusion between a promotor functional in *R. sphaeroides* and a potential reading frame. Promotors of *E. coli* were shown not to work in *R. sphaeroides* (F. Daldal, personal communications), but endogenous promotors of *R. sphaeroides* have not yet been described in detail, and were not available. F. Daldal suggested the use of a fragment upstream from the *cycA* gene of *R. capsulata* (2), which he knew was able to promote transcription in *R. sphaeroides*. Based on this information, a 350 bp fragment beginning just upstream of the ribosome binding site of the *cycA* gene of *R. capsulata* was isolated by PCR (Figure B-1). The two 30-bp oligonucleotide PCR primers were designed such that one end of the PCR product could be cut by HindIII, while the other end contained a PstI site (Figure B-2). This design allowed the directional cloning of the promotor fragment into the polylinker of the broad host range vector pRK415 (3), obtained from N.T. Keen, giving rise to the expression vector pCHB500 (Figure B-3), which was used throughout much of this work. The PCR reactions were performed using the linearized plasmid pSH3 (cut with EcoRI) as template, which was kindly provided by F. Daldal and contained the *cycA* gene complete with its promotor sequences from

Figure B-1: Sequence of the *cycA* gene from *R. capsulata*, upstream of the initiation codon (taken from 2). The location of the two PCR primers CB22 and CB23 is indicated.

```
          CB22
1   GTCGACCCGGATCTGAGCCATTTTTCCGGCATCGTGCCCTGCGGCATCCGGAGCATGGCG

61  TGACCTCGCTCGTCGATCTGGGGCTGCCGGTGACGATGGCCGATCTCGACCTGGCGCTGC

121 GGGCCCGGTTCGAGACCGATTTCACCCCGCGCCTGCGCGGCTGAGCCGCCGCCGCAAGCG

181 CTGCCGGAGGGCGGCATTTCATCTGGCATTCACTTTGCGTCCGGATTCTTGCCCCGCGGG

241 TCATGGAATTGAAAGCAAAGGGATATGCGCGACCTTTTGCCTTGTGGTAAAGCGTGGACC
                                                  CB23
301 GGGTAAGGAATCTGCCCCTTGGCCCCGGCTCCAGACCGGGCAAGCAAACAGACGAAGGAT

361 ACAGCGATG
```

Figure B-2: Sequence of the two PCR primers used to clone the *cycA* promotor from *R. capsulata*. The HindIII (H3) and PstI(P1) sites are indicated above the sequence.

```
                 H3
CB22    5'-ATCGAAGCTTGTCGACCCGGATCTGAGCCA-3'

                 P1
CB23    5'-ATCGCTGCAGCGTCTGTTTGCTTGCCCGGT-3'
```

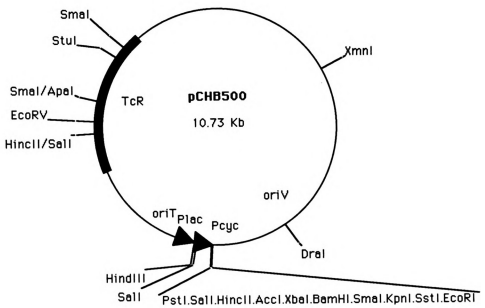


Figure B-3: Circular restriction map of pCHB500. The orientation and location of two promotors, Plac (promotor of the *E. coli* lac-operon) and Pcyc (putative *R. capsulata* promotor of the *cycA* gene), are indicated by thick arrows. The tetracycline gene (TcR) is indicated as thick a line. The locations of the origin of transfer (oriT) and origin of replication (oriV) are shown. Restriction sites are shown with their common names. The sites of the polylinker, downstream from the Pcyc promotor, are shown in an expanded view.

R. capsulata. The PCR reactions were set up on a Perkin Elmer thermocycler using the buffers, templates, and primers at concentrations recommended by that company. Reactions of 100µl were run using the following temperature program: 3 min at 95°C; 30 cycles of 1 min at 94°C, 3 min at 60°C, and 3 min at 72°C; 7 min at 72°C. The reaction product was gel-purified and digested with HindIII and PstI. Following ligation with pRK415 cut with the same enzymes, and transformation of HB101 cells, several clones were identified in DNA minipreps which contained the expected fragment, as was confirmed by Southern analysis.

To test the capability of pCHB500 to express open reading frames in an orientation-dependent fashion, a 977 bp Sau3AI fragment obtained from pBR328 (1) containing the open reading frame for the chloramphenicol acetyl transferase (CAT) gene, was cloned in both orientations into the BamHI site of the pCHB500 polylinker. Upon transfer of the resulting clones into *R. sphaeroides* and plating of the cells onto agar plates containing various concentrations of chloramphenicol, it was observed that clones, which contained the CAT gene in the same orientation as the *cycA* promotor, could confer resistance in the presence of 10 µg/ml chloramphenicol, while clones containing the CAT gene in the opposite orientation could not. A similar result was observed , when a cassette containing the NPT open reading frame, conferring resistance to kanamycin, was cloned in both orientations behind the *cycA* promotor. Although the orientation-dependent expression was not precisely quantitated, the data were in agreement with proper transcription

from the *cycA* promotor in *R. sphaeroides*. As described in Chapters 4, and 5, the pCHB500 vector was successfully used to test the orientation of the coding sequence for the *sqdA* and *sqdC* genes in complementation experiments. The vector has been given to F. Daldal, who is interested in analyzing the expression from the *cycA* promotor in more detail.

REFERENCES

1. Balbas P., Soberon X. Merino E., Zurita M., Lomeli H., Valle F., Flores N., Bolivar F. (1986). Plasmid vector pBR322 and its special-purpose derivatives - a review. Gene 50:3-40

2. Daldal F., Cheng S., Appelbaum J., Davidson E., Prince R.C. (1986). Cytochrome c_2 is not essential for photosynthetic growth of *Rhodopseudomonas capsulata*. Proc. Natl. Acad. Sci. USA 83:2012-2016

3. Keen N.T., Tamaki S., Kobayashi D., Trollinger D. (1988). Improved broad-host-range vector for DNA cloning in Gram-negative bacteria. Gene 70:191-197