

MOLECULAR GENETICS OF SULPHATE  
ASSIMILATION IN 'ARABIDOPSIS THALIANA'

Michael Austin Roberts

A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews



1997

Full metadata for this item is available in  
St Andrews Research Repository  
at:

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/14347>

This item is protected by original copyright

Tk  
C256

ProQuest Number: 10167161

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

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



ProQuest 10167161

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

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

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

Date: 1/05/1997      Signature of candidate

MOLECULAR GENETICS OF SULPHATE ASSIMILATION  
IN *ARABIDOPSIS THALIANA*

by

Michael Austin Roberts

A thesis submitted to the  
University of St Andrews  
for the degree of  
Doctor of Philosophy

29 January 1997



## Abstract

The steps involved in plant reductive assimilation of sulphate to sulphide for incorporation into cysteine are not clear. The aim of the project described in this thesis was the isolation of genes encoding sulphate assimilation enzymes which will provide molecular tools for unravelling this key metabolic pathway.

Functional complementation of the *Escherichia coli* cysteine auxotrophic strain JM15 (serine acetyltransferase deficient) using an *Arabidopsis thaliana* cDNA library in the expression vector  $\lambda$ YES resulted in the isolation of at least three members of an *A. thaliana* multigene family encoding serine acetyltransferase. Characterisation of one clone, Sat-1, showed that it conferred serine acetyltransferase activity (with apparent  $K_m$  for substrates acetyl CoA and L-serine of 0.043 and 3.47 mM, respectively) on strain JM15. The 1515 bp full-length cDNA encodes a deduced protein of 391 amino acids, SAT-1, that has significant identity with bacterial and plant serine acetyltransferases, and that contains a putative N-terminal organellar targeting peptide. Southern hybridisation indicated that Sat-1 is present as a single copy in *A. thaliana*, while northern analysis revealed a single message of 1.5 kb.

Using the *A. thaliana* cDNA library in the expression vector  $\lambda$ YES, cDNAs encoding a novel putative "APS reductase" were obtained by functional complementation of *E. coli* cysteine auxotrophic strains JM81A (adenosine 5'-phosphosulphate [APS] kinase deficient) and JM96 (3'-phosphoadenosine-5'-phosphosulphate [PAPS] reductase deficient). Retransformation of three clones (Papsr-19, Papsr-26 and Papsr-43) encoding different putative APS kinase isoforms into strain JM96 conferred low PAPS reductase activity on the mutant, although this activity was thioredoxin-independent unlike wild-type bacterial activity. The putative APS reductase has a PAPS reductase-like C-terminal domain, but further analysis

demonstrated that the enzyme accepts APS in preference to PAPS as substrate and has a thioredoxin-like C-terminal domain (Gutierrez-Marcos et al. 1996, Proc. Natl Acad. Sci. 93: 13377-13382).

Isolation and characterisation of these genes invites a new hypothesis for plant reductive sulphate assimilation and provides direction for future research.

Declarations

(i) I, Michael Austin Roberts, hereby certify that this thesis, which is approximately 68500 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date: 27/01/1997 Signature of candidate

(ii) I was admitted as a research student in January 1993 and as a candidate for the degree of Doctor of Philosophy in January 1994; the higher study for which this is a record was carried out in the University of St Andrews between 1993 and 1996.

Date: 27/01/1997 Signature of candidate

(iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 27/1/97 Signature of supervisor: (

## Acknowledgements

I thank my supervisor, Dr John L. Wray, for conceptualising an interesting and rewarding project, for providing research facilities, and for critically reading this thesis. I thank also those staff and students at the University of St Andrews, particularly in the Sir Harold Mitchell Building, who provided me with assistance during the course of the work.

I am grateful to members of the Cereals Department, John Innes Centre, Norwich, for allowing me to use their facilities during the preparation of the thesis and for their patience.

I thank my family for their long-distance support and encouragement, without which the task would have been much more difficult.

Financial support was provided by a University of St Andrews Research Studentship and an Overseas Research Student Award.

## List of Contents

	<u>Page</u>
<u>CHAPTER 1: INTRODUCTION</u>	1
<u>1.1 The relevance of sulphur for plants</u>	1
<u>1.2 Biochemistry and molecular genetics of sulphate assimilation</u>	3
<b>1.2.1 Sulphate uptake and transport</b>	3
<b>1.2.2 Sulphate activation by ATP sulphurylase</b>	6
<b>1.2.3 Reduction of activated sulphate to sulphide - carrier-bound or free intermediate pathway?</b>	8
<b>1.2.3.1 Bound intermediate pathway</b>	9
<i>Transfer of the sulpho group of APS to a carrier molecule by APS sulphotransferase</i>	9
<i>Reduction of bound sulphite to bound sulphide by thiosulphonate reductase</i>	11
<i>Incorporation of sulphide into cysteine</i>	12
<i>Conversion of APS to PAPS as a supply of activated sulphate</i>	12
<i>Summary of chemical reactions</i>	13
<b>1.2.3.2 Free intermediate pathway</b>	15
<i>Phosphorylation of APS to PAPS by APS kinase</i>	16
<i>Reduction of PAPS by PAPS reductase to form sulphite</i>	17
<i>Reduction of sulphite to sulphide by sulphite reductase</i>	19
<i>Incorporation of sulphide into cysteine</i>	20
<i>Summary of chemical reactions</i>	20
<b>1.2.3.3 Comparison of the free and bound intermediate pathways of sulphate assimilation</b>	21
<i>Reactivity of sulphur-containing compounds</i>	21
<i>Evaluation of sulphate assimilation enzyme assays</i>	22
<i>Is PAPS an obligatory intermediate in plant sulphate assimilation?</i>	24
<i>Conclusions</i>	29

	<u>Page</u>
<b>1.2.4 Biosynthesis of cysteine</b>	29
<i>Formation of O-acetylserine by serine acetyltransferase</i>	30
<i>Cysteine formation by O-acetylserine (thiol)-lyase</i>	31
<i>A multifunctional complex for cysteine biosynthesis</i>	32
<u>1.3 Functional complementation - the cloning method of choice?</u>	34
<u>1.4 Aims of the investigation</u>	37
<u>CHAPTER 2: MATERIALS AND METHODS</u>	38
<u>2.1 Reagents</u>	38
<u>2.2 Bacterial strains and their maintenance</u>	39
<u>2.3 Plant material</u>	41
<u>2.4 The <math>\lambda</math>YES <i>Arabidopsis thaliana</i> cDNA expression vector for functional complementation</u>	42
2.4.1 Description of the $\lambda$ YES cDNA library	42
2.4.2 Titrating $\lambda$ phages	44
<i>Preparation of plating cells</i>	44
<i>Plating phage <math>\lambda</math></i>	44
2.4.3 Induction of $\lambda$ KC in <i>Escherichia coli</i> strain BNN132	45
2.4.4 Creating $\lambda$ KC lysogens of <i>Escherichia coli</i> cysteine auxotrophs	46
2.4.5 Functional complementation of <i>Escherichia coli</i> cysteine auxotrophic strains with the $\lambda$ YES cDNA library	47
2.4.6 Production of empty vector pYES from clone pSAT1	48
<u>2.5 Plasmid DNA</u>	49
2.5.1 Plasmids employed	49
2.5.2 Preparation of plasmid DNA	49

	<u>Page</u>
2.5.3 Transformation of <i>Escherichia coli</i> with plasmid DNA	52
2.5.4 Restriction enzyme digestion of DNA	54
2.5.5 Agarose gel electrophoresis of DNA fragments	55
<u>2.6 Isolation and analysis of plant genomic DNA</u>	56
2.6.1 Extraction of plant genomic DNA	56
2.6.2 Restriction enzyme digestion of plant genomic DNA	57
2.6.3 Southern analysis of DNA	58
<i>Agarose gel electrophoresis</i>	58
<i>Southern transfer</i>	58
<i>Probe isolation and labelling</i>	59
<i>Prehybridisation and hybridisation</i>	61
<i>Washing and autoradiography</i>	61
<i>Gene copy number determination</i>	61
<u>2.7 Isolation and analysis of <i>Arabidopsis thaliana</i> RNA</u>	63
2.7.1 General precautions against RNases	63
2.7.2 Extraction of <i>Arabidopsis thaliana</i> total RNA	63
2.7.3 Extraction of <i>Arabidopsis thaliana</i> mRNA	64
2.7.4 Northern analysis of RNA	65
<i>Agarose gel electrophoresis</i>	65
<i>Northern transfer</i>	66
<i>Prehybridisation and hybridisation</i>	66
<i>Washing and autoradiography</i>	67
2.7.5 Synthesis of radiolabelled first-strand cDNA to assess poly(A) RNA quality	67
<u>2.8 DNA sequence determination</u>	68
2.8.1 Introduction	68
2.8.2 Selection, manufacture and preparation of oligonucleotide primers	69
2.8.3 DNA sequencing with the Sequenase Version 2.0 DNA Sequencing kit	70
<i>Denaturation of double-stranded DNA template</i>	71
<i>Annealing template and primer</i>	71

	<u>Page</u>
<i>Labelling and extension/termination reactions</i>	71
2.8.4 DNA sequencing with the TaqTrack Sequencing System	72
<i>Alkaline denaturation of plasmid DNA</i>	72
<i>Annealing template and primer</i>	73
<i>Labelling reaction</i>	73
<i>Extension/termination reaction</i>	73
<i>Chase step to eliminate false bands</i>	73
2.8.5 Direct DNA sequencing of PCR products	74
2.8.6 End-labelled primer DNA sequencing	74
2.8.7 Preparation and running of polyacrylamide sequencing gels	74
2.8.8 Gel drying and autoradiography	75
<u>2.9 Rapid amplification of cDNA ends (RACE)</u>	77
2.9.1 Primers	77
2.9.2 cDNA synthesis	77
2.9.3 RNA hydrolysis, cDNA purification and precipitation	78
2.9.4 Anchor ligation	78
2.9.5 PCR amplification	79
2.9.6 Purification of PCR products	79
<u>2.10 Ligation of DNA into plasmid vectors</u>	80
<u>2.11 Determination of mRNA 5' ends by primer extension</u>	81
2.11.1 Primer labelling	81
2.11.2 Primer hybridisation and extension	81
2.11.3 Electrophoresis	82
<u>2.12 Analysis of DNA and protein sequences</u>	83
<u>2.13 Expression and purification of GST-PRH-26 fusion proteins</u>	85
2.13.1 Production of transformants	85
2.13.2 Analysis of GST recombinants for fusion protein expression	86
2.13.3 Large-scale purification of GST-PRH-26 fusion protein	88

	<u>Page</u>
<u>2.14 Preparation of rabbit anti-GST-PRH-26 antibodies</u>	89
<u>2.15 Western analysis of proteins</u>	90
2.15.1 Protein fractionation by SDS-PAGE	90
<i>Gel preparation and running</i>	90
<i>Coomassie Brilliant Blue staining</i>	91
2.15.2 Western transfer	92
2.15.3 Staining of proteins immobilised on nitrocellulose filters	92
2.15.4 Immunodetection of proteins immobilised on nitrocellulose filters	93
<u>2.16 Protein quantification</u>	94
<u>2.17 Enzyme assays</u>	95
2.17.1 $\beta$ -galactosidase	95
<i>Preparation of enzyme extract</i>	95
<i>Enzyme assay</i>	95
2.17.2 Serine acetyltransferase	96
<i>Preparation of enzyme extract</i>	96
<i>Enzyme assay</i>	96
2.17.3 PAPS reductase	97
<i>Preparation of enzyme extract</i>	97
<i>Enzyme assay</i>	97
<u>2.18 Photography</u>	99
<u>CHAPTER 3: CHARACTERISATION OF ARABIDOPSIS THALIANA cDNAS ENCODING SERINE ACETYLTRANSFERASE CLONED BY FUNCTIONAL COMPLEMENTATION OF ESCHERICHIA COLI CYSE<sup>-</sup> STRAIN JM15</u>	100
<u>3.1 Introduction</u>	100
<u>3.2 Results and Discussion</u>	103
2.15.1 Complementation of the <i>Escherichia coli</i> cysE mutant JM15 to prototrophy with the $\lambda$ YES <i>Arabidopsis thaliana</i> cDNA library	103

	<u>Page</u>
<i>Growth characteristics of Escherichia coli strain JM15</i>	103
<i>Induction of helper phage <math>\lambda</math>KC and preparation of <math>\lambda</math>KC lysogens of Escherichia coli strain JM15</i>	104
<i>Complementation of Escherichia coli strain JM15/<math>\lambda</math>KC with the <math>\lambda</math>YES cDNA library</i>	106
<i>Retransformation of pSAT1 into Escherichia coli strain JM15</i>	107
<b>3.2.2 Enzymatic analysis of JM15, wild-type and complemented SAT strains</b>	112
<i><math>\beta</math>-galactosidase activity of crude extracts from Escherichia coli strain JM15</i>	112
<i>Serine acetyltransferase activity in complemented, mutant and wild-type Escherichia coli strains</i>	113
<i>Assay for serine acetyltransferase activity in Arabidopsis thaliana extracts</i>	118
<b>3.2.3 Analysis of Sat clones and cDNA sequence determination</b>	119
<b>3.2.3.1 Analysis of clone Sat-2</b>	122
<i>Method development for DNA sequencing of clone Sat-2</i>	122
<i>Characterisation of clone Sat-2 and sequence comparisons</i>	125
<b>3.2.3.2 Analysis of clone Sat-1</b>	129
<i>Method development for DNA sequencing of clone Sat-1</i>	129
<i>Characterisation of clone Sat-1</i>	132
<i>Primer extension analysis of EST clone 89K20T7</i>	134
<i>5'RACE amplification of an Arabidopsis thaliana putative full-length Sat-1 cDNA clone</i>	135
<b>3.2.3.3 Analysis of clone Sat-1a</b>	150
<i>Characterisation of clone Sat-1a and sequence comparisons</i>	150
<i>Subcellular location of SAT-1</i>	163

	<u>Page</u>
<i>Is a second polypeptide expressed from the Sat-1a sequence?</i>	164
<b>3.2.3.4 Sequence analysis of Sat-80</b>	168
<b>3.2.4 Nucleic acid blotting with the Sat-1 probe</b>	168
<b>3.2.4.1 Southern analysis</b>	168
<b>3.2.4.2 Northern analysis</b>	176
<b><u>3.3 Conclusions</u></b>	178
<b><u>CHAPTER 4: CHARACTERISATION OF ARABIDOPSIS THALIANA cDNAS ENCODING PAPS REDUCTASE HOMOLOGUE ISOFORMS OBTAINED BY FUNCTIONAL COMPLEMENTATION OF ESCHERICHIA COLI STRAINS JM81A (DEFICIENT IN APS KINASE) AND JM96 (DEFICIENT IN PAPS REDUCTASE)</u></b>	180
<b><u>4.1 Introduction</u></b>	180
<b><u>4.2 Results and Discussion</u></b>	182
<b>4.2.1 Complementation of the <i>Escherichia coli</i> <i>cysC</i> mutant JM81A to prototrophy with the <math>\lambda</math>YES <i>Arabidopsis thaliana</i> cDNA library</b>	182
<i>Growth characteristics of Escherichia coli strain JM81A and preparation of strain JM81A/<math>\lambda</math>KC lysogens</i>	182
<i>Complementation of Escherichia coli strain JM81A/<math>\lambda</math>KC with the <math>\lambda</math>YES cDNA library</i>	184
<i>Retransformation of pAPSK14 and pAPSK15 into Escherichia coli strain JM81A</i>	185
<b>4.2.2 Analysis of Apsk clones and cDNA sequence determination</b>	187
<b>4.2.2.1 Analysis of clones Apsk-14 and Apsk-15</b>	187
<i>Characterisation of clones Apsk-14 and Apsk-15</i>	187
<i>Subcellular location of APSK-15</i>	192
<i>Sequence comparisons with APSK-14 and APSK-15</i>	194
<b>4.2.2.2 Are other enzyme types present amongst the Apsk clones?</b>	198

	<u>Page</u>
<b>4.2.3 Complementation of the <i>Escherichia coli</i> <i>cysC</i> mutant JM96 to prototrophy with the <math>\lambda</math>YES <i>Arabidopsis thaliana</i> cDNA library</b>	198
<i>Growth characteristics of Escherichia coli strain JM96 and preparation of strain JM96/<math>\lambda</math>KC lysogens</i>	199
<i>Complementation of Escherichia coli strain JM96/<math>\lambda</math>KC with the <math>\lambda</math>YES cDNA library</i>	201
<i>Retransformation of pPAPSR19 and pPAPSR26 into Escherichia coli strain JM96</i>	202
<b>4.2.4 Analysis of Papsr clones and cDNA sequence determination</b>	203
<b>4.2.4.1 Analysis of clones Papsr-19 and Papsr-20</b>	205
<i>Characterisation of clones Papsr-19 and Papsr-20</i>	208
<b>4.2.5 Comparison of Apsk and Papsr clones</b>	212
<b>4.2.6 Analysis and sequence comparisons of full-length APSK/PAPSR isoforms</b>	214
<i>Characterisation of PRH-19, PRH-26 and PRH-43 isoforms and comparison to APSK-14, APSK-15, PAPSR-19 and PAPSR-20</i>	214
<i>Sequence comparisons using PRH-19</i>	217
<b>4.2.7 Enzymatic analysis of JM96, wild-type and pPAPSR-complemented <i>Escherichia coli</i> strains, <i>Arabidopsis thaliana</i> and <i>Spinacia oleracea</i></b>	223
<i>PAPS reductase activity</i>	223
<i>APS reductase activity</i>	228
<b>4.2.8 Expression and purification of a GST-PRH-26 fusion protein and preparation of rabbit anti-GST-PRH-26 polyclonal antibodies</b>	230
<i>Formation of a GST-PRH-26 fusion protein and optimisation of expression</i>	231
<i>Production of rabbit anti-GST-PRH-26 polyclonal antibodies</i>	236
<b><u>4.3 Conclusions</u></b>	239

	<u>Page</u>
<u>CHAPTER 5: CONCLUDING REMARKS</u>	241
<i>Serine acetyltransferase</i>	241
<i>APS reductase and a new hypothesis for a pathway of     higher plant reductive sulphate assimilation</i>	242
<i>Productive years in plant sulphate assimilation     research</i>	244
 <u>REFERENCES</u>	 245

## List of Tables

	<u>Page</u>
<b><u>Table 3.1</u></b> UV induction of $\lambda$ KC from <i>Escherichia coli</i> strain BNN132	105
<b><u>Table 3.2</u></b> Efficiency of cloning plant genes by functional complementation of <i>Escherichia coli</i> mutant strains	108
<b><u>Table 3.3</u></b> Transformation and complementation of <i>Escherichia coli</i> strain JM15 by pSAT1	111
<b><u>Table 3.4</u></b> Serine acetyltransferase activity of <i>Escherichia coli</i> strains SAT1, SAT2, JM15 and LE392 (wild-type)	117
<b><u>Table 4.1</u></b> Papsr clone size determined by agarose gel fractionation of <i>Eco</i> RI-digested plasmids	206
<b><u>Table 4.2</u></b> PAPS reductase activity in cell-free extracts of <i>Escherichia coli</i> JM96 ( <i>cysH</i> ), TB1 (wild-type) and pPAPSR-complemented JM96 strains, and in leaf extracts of <i>Arabidopsis thaliana</i> and <i>Spinacia oleracea</i>	226

## List of Figures

	<u>Page</u>
<b><u>Figure 2.1</u></b> The $\lambda$ YES system for <i>cre-lox</i> -mediated automatic subcloning of phage $\lambda$ to plasmid clones	43
<b><u>Figure 3.1</u></b> Growth characteristics of <i>Escherichia coli</i> cyst(e)ine auxotroph JM15 and the transformed strains JM15/pYES and JM15/pSAT1	110
<b><u>Figure 3.2</u></b> $\beta$ -galactosidase activity in crude cell-free extracts of <i>Escherichia coli</i> strain JM15	114
<b><u>Figure 3.3</u></b> Dependence of serine acetyltransferase activity in strain SAT1 extracts on acetyl CoA and L-serine concentration	115
<b><u>Figure 3.4</u></b> Serine acetyltransferase assay using crude extracts of <i>Arabidopsis thaliana</i>	120
<b><u>Figure 3.5</u></b> Electrophoretic analysis of pYES-derived plasmids extracted from complemented <i>Escherichia coli</i> JM15 strains	121
<b><u>Figure 3.6</u></b> DNA sequence obtained from pSAT2 purified by Promega Magic Minipreps and single-stranded control DNA	123
<b><u>Figure 3.7</u></b> Nucleotide sequence and deduced amino acid sequence of clone Sat-2	126
<b><u>Figure 3.8</u></b> Amino acid sequence alignment of ATHSAT-2 with prokaryotic serine acetyltransferases	128

	<u>Page</u>
<b><u>Figure 3.9</u></b> Optimisation of DNA sequencing protocols for sequencing the 5' end of the noncoding strand of clone Sat-1	130
<b><u>Figure 3.10</u></b> Nucleotide sequence comparison of clones Sat-1, Sat-2 and EST 89K20T7, and deduced amino acid sequence of clone Sat-1	133
<b><u>Figure 3.11</u></b> Determination of the putative transcription initiation site of EST clone 89K20T7 by primer extension	136
<b><u>Figure 3.12</u></b> PCR amplification of the 5' end of the human TFR gene using the 5'-AmpliFINDER RACE kit	138
<b><u>Figure 3.13</u></b> RACE amplification of the 5' end of clone Sat-1. I. Use of primer PR2 and cDNA synthesised from poly(A) RNA	140
<b><u>Figure 3.14</u></b> RACE amplification of the 5' end of clone Sat-1. II. Use of primer P9 and cDNA synthesised from poly(A) RNA	142
<b><u>Figure 3.15</u></b> DNA sequence and deduced reading frames of clone Sat-1.5.1	144
<b><u>Figure 3.16</u></b> RACE amplification of the 5' end of clone Sat-1. III. Use of cDNA synthesised from total RNA and primers PR2 and P9	146
<b><u>Figure 3.17</u></b> Electrophoretic analysis of Sat-1.5.2 and Sat-1.5.3 fragments ligated into pGEM-7Zf(+)	147
<b><u>Figure 3.18</u></b> Nucleotide sequence and deduced amino acid sequences of clone Sat-1a	148

	<u>Page</u>
<b><u>Figure 3.19</u></b> Amino acid sequence alignment of plant and bacterial serine acetyltransferases	152
<b><u>Figure 3.20</u></b> Phylogenetic tree of plant and bacterial serine acetyltransferases	154
<b><u>Figure 3.21</u></b> Dot-plot comparisons of SAT-1 against itself, related enzymes and an unrelated sequence	160
<b><u>Figure 3.22</u></b> Amino acid alignment of the hypothetical proteins SAT-X and CysX	166
<b><u>Figure 3.23</u></b> Alignment of the hypothetical proteins SAT-X and CysX with corresponding reverse translations of four serine acetyltransferases	167
<b><u>Figure 3.24</u></b> Nucleotide sequence of the 5' and 3' ends of clone Sat-80	169
<b><u>Figure 3.25</u></b> Agarose gel analysis of <i>Arabidopsis thaliana</i> genomic DNA	170
<b><u>Figure 3.26</u></b> Southern blot analysis of <i>Arabidopsis thaliana</i> genomic DNA with the clone Sat-1	172
<b><u>Figure 3.27</u></b> Gene copy number reconstruction experiment with the clone Sat-1	174
<b><u>Figure 3.28</u></b> Southern blot analysis of <i>Brassica napus</i> genomic DNA with the clone Sat-1	175
<b><u>Figure 3.29</u></b> Northern blot analysis of <i>Arabidopsis thaliana</i> total RNA with the clone Sat-1	177

	<u>Page</u>
<b><u>Figure 4.1</u></b> Growth characteristics of <i>Escherichia coli</i> cysteine auxotrophic strain JM81A	183
<b><u>Figure 4.2</u></b> Electrophoretic analysis of pYES-derived plasmids extracted from complemented <i>Escherichia coli</i> JM81A strains	188
<b><u>Figure 4.3</u></b> Nucleotide sequence of the 5' and 3' end of clone Apsk-14 and deduced amino acid sequence	189
<b><u>Figure 4.4</u></b> Nucleotide sequence of the 5' and 3' end of clone Apsk-15 and deduced amino acid sequence	191
<b><u>Figure 4.5</u></b> Amino acid comparison of the APSK-14 and APSK-15 polypeptides	193
<b><u>Figure 4.6</u></b> Amino acid sequence alignment of APSK-14, APSK-15 and PAPS reductases	195
<b><u>Figure 4.7</u></b> Growth characteristics of <i>Escherichia coli</i> cysteine auxotrophic strain JM96	200
<b><u>Figure 4.8</u></b> Growth characteristics of <i>Escherichia coli</i> strains JM96, JM96/pYES and JM96/pPAPSR19	204
<b><u>Figure 4.9</u></b> Electrophoretic analysis of pYES-derived plasmids extracted from complemented <i>Escherichia coli</i> JM96 strains	207
<b><u>Figure 4.10</u></b> Nucleotide sequence of the 5' and 3' end of clone Papsr-19 and deduced amino acid sequence	209
<b><u>Figure 4.11</u></b> Nucleotide sequence of the 5' and 3' end of clone Papsr-20 and deduced amino acid sequence	211

	<u>Page</u>
<b><u>Figure 4.12</u></b> Amino acid comparison of APSK-14, APSK-15, PAPSR-19 and PAPSR-20 polypeptides	213
<b><u>Figure 4.13</u></b> Amino acid alignment of PRH-19, PRH-26 and PRH-43 proteins	215
<b><u>Figure 4.14</u></b> Amino acid alignment of the C-terminus of PRH-19 with thioredoxin and thioredoxin-like domains of representative members of the PDI family	220
<b><u>Figure 4.15</u></b> SDS-PAGE analysis of expression and purification of GST and the fusion protein GST-PRH-26 in transformed <i>Escherichia coli</i> strain TB1 cells	233
<b><u>Figure 4.16</u></b> SDS-PAGE analysis of purified and eluted fusion protein GST-PRH-26	235
<b><u>Figure 4.17</u></b> Dot-blot western analysis using rabbit pre-immune serum and primary antiserum raised against the GST-PRH-26 fusion protein	237

## CHAPTER 1: INTRODUCTION

### 1.1 The relevance of sulphur for plants

Sulphur is a plant macronutrient required for biosynthesis of the amino acids cysteine and methionine, coenzymes and primary and secondary metabolites (Metzler 1977; Mifflin & Lea 1977; Schmidt & Jäger 1992). Sulphur-containing compounds such as thiols, glucosinolates, sulpholipids and thiazoles are components of the defence mechanism of plants against disease and herbivores (Kleppinger-Sparace et al. 1990a, b; Schnug 1993), and also affect the nutritional status of several crop species (Bennett & Wallsgrove 1994). Early and more recent studies on the effects of sulphur deficiency showed that plants with suboptimal sulphur nutrition became chlorotic, had smaller leaves, thinner stems, and decreased quality and quantity of proteins (Eaton 1935, 1941, 1942, 1951; Eppendorfer 1968; Millard et al. 1985). Whereas previously the sulphur requirement of crop plants was satisfied by sulphur-containing fertilizers and atmospheric deposition, both of these sources have been diminished and sulphur deficiencies are becoming more frequent (Jolivet 1993). Since plants provide nonruminant animals with most of the dietary requirement for the sulphur-containing essential amino acid methionine (Giovanelli et al. 1980; Leustek 1996), a need to understand the biochemical pathways of sulphur nutrition and their regulation in plants is becoming increasingly important.

Inorganic sulphate is the form of sulphur most commonly available in the soil for plant metabolism (Anderson 1980; Schiff et al. 1993) and excluding small amounts of sulphur dioxide absorbed by leaves in polluted areas, sulphate provides all the sulphur necessary for plant growth (Peck & Lissolo 1988; Salisbury & Ross 1992). Sulphate-sulphur has an oxidation state of +6, whereas the sulphide-sulphur contained

in most organic compounds has an oxidation state of -2. Plants therefore catalyse an eight electron reduction of sulphur from sulphate to sulphide, a pathway known as sulphate assimilation (Anderson 1980). Sulphate assimilation is an energetically expensive process for plants, requiring almost double the amount of energy as carbon dioxide fixation or nitrate reduction to ammonium (Schwenn 1994). A distinction should be noted here between sulphate assimilation, the reduction of sulphate to sulphide carried out by photosynthetic and several nonphotosynthetic organisms, and dissimilatory sulphate reduction, where sulphate serves as a terminal respiratory electron acceptor for a class of anaerobic bacteria, the sulphate reducing bacteria (Peck & Lissolo 1988).

This introductory chapter will:

- (i) provide an overview of historical and current views on reductive sulphate assimilation into cysteine, the primary organic sulphur compound formed in plants. The emphasis of this section will be explanation and critical analysis of proposed biosynthetic pathways for sulphate assimilation. Less importance will be placed on other aspects of sulphate assimilation such as regulation and flux (reviewed in Anderson 1980; Giovanelli et al. 1980; Schmidt 1986; Giovanelli 1990; Kredich 1993; Stulen & De Kok 1993). Where useful for understanding the process in plants, and to provide a background to the methodology of this thesis, comparison will be made with the well-characterised pathways of sulphate assimilation in lower eukaryotes and prokaryotes;
- (ii) discuss the appropriateness of using functional complementation to isolate plant sulphate assimilation genes; and
- (iii) outline the objectives of this study.

## 1.2 Biochemistry and molecular genetics of sulphate assimilation

### **1.2.1 Sulphate uptake and transport**

Early studies with higher plants indicated that sulphate uptake was an energy-dependent process utilising respiratory (in preference to photosynthetic) reductants (Anderson 1980; Ferrari et al. 1990). Sulphate uptake showed complex kinetics that several authors found consistent with a single Michaelis-Menten type permease with a  $K_m$  (sulphate) of 10 - 100  $\mu\text{M}$  (Schiff & Hodson 1973; Anderson 1980; Clarkson et al. 1993), as has been found in several bacterial and fungal species (Furlong 1987; Kredich 1987; Gilmore et al. 1989; Woodin & Wang 1989). Others suggested a multiphasic mechanism or inferred the presence of two separate permeases with different  $K_m$  (Anderson 1980; Schmidt 1986), similar to the high and low affinity permeases found in lower eukaryotes such as *Saccharomyces cerevisiae* (yeast) (Breton & Surdin-Kerjan 1977) and *Neurospora crassa* (Marzluf 1970a, b; Ketter & Marzluf 1988; Ketter et al. 1991). Some variation may be explained by the different experimental systems such as isolated cells or excised roots employed for measuring sulphate uptake, from which extrapolation to *in vivo* uptake mechanisms may not be valid (Stulen & De Kok 1993). Two genes encoding different root-expressed sulphate transporters from the legume *Stylosanthes hamata* have been isolated recently (Smith et al. 1995). Both genes encode high-affinity permeases, with  $K_m$  for sulphate of 10 - 11  $\mu\text{M}$  (Smith et al. 1995), consistent with the notion that plants have a single, well regulated mechanism type for sulphate uptake (Clarkson et al. 1993). However, a cDNA encoding a putative low-affinity sulphate transporter from *Arabidopsis thaliana* has been isolated (Takahashi et al. 1996). Hybridisation analysis using that cDNA and two similar *A. thaliana* Expressed Sequencing Tag (EST) clones suggests that the plant contains three sulphate permease genes that are expressed in an organ-

specific manner (Takahashi et al. 1996).

A ubiquitous finding has been that sulphate uptake ability increased after prolonged sulphate starvation (Smith 1980; Saccomani et al. 1984; Clarkson et al. 1989; Cram 1990; Yildiz et al. 1994). In *Chlamydomonas reinhardtii*, an initial response to sulphate starvation is the *de novo* synthesis of a new high-affinity sulphate transporting system, while a secondary response, reported also for higher plants, involves the production of arylsulfatase to cleave sulphate off aromatic compounds (Davies et al. 1994; Yildiz et al. 1994). Northern analysis of the *S. hamata* high-affinity sulphate transporters indicated that expression of the genes is upregulated by sulphate deprivation (Smith et al. 1995). Furthermore, in *Zea mays* (maize) seedlings, energy coupling to sulphate uptake was found to be modulated by the availability of sulphate during seedling growth (Ferrari et al. 1990). It has been suggested that intracellular sulphate levels, which decrease after sulphate deprivation, regulate sulphate uptake (Smith 1975, 1980). Although a sulphate signal is not required for initial induction of plant sulphate transporters (Clarkson et al. 1993; Smith et al. 1995), the ability to detect sulphate and regulate sulphate uptake in response to shortages plays a key role in plant cultivar productivity (Saccomani et al. 1984).

The molecular mechanisms of sulphate uptake are not well known, even in prokaryotes (Schmidt & Jäger 1992; Schiff et al. 1993). Cram (1990) provided evidence for a proton co-transport system, derived mainly from experiments with microorganisms rather than higher plants. The proton motive force model for sulphate uptake was elaborated by Clarkson et al. (1993), albeit with little unequivocal experimental evidence from plants. In the proton motive force model, co-transport of one sulphate ion and three protons depolarises the membrane potential, and these protons then acidify the cytoplasm to activate plasma membrane ATPase which drives

protons out (Clarkson et al. 1993). The proton motive force model has been supported from work on *C. reinhardtii* (Yildiz et al. 1994), and structural analysis of the high affinity sulphate transporters from *S. hamata* reveals similarity with a family of proton/sulphate cotransporters (Smith et al. 1995). Further studies on cloned plant sulphate permeases will provide understanding of the mechanisms of sulphate uptake.

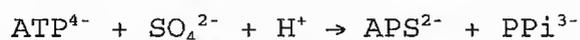
Sulphate taken up by plant roots is reduced primarily in leaf chloroplasts (Anderson 1980; Schwenn 1994), but can be distributed for storage in several organs (Dallam 1987). Also, roots of *Z. mays* seedlings are capable of rapidly reducing inorganic sulphate to organic forms (Dallam 1987), and excised *Macroptilium atropurpureum* root tissue reduced 5 % of the <sup>35</sup>S-sulphate taken up across the plasmalemma to organic form, compared with 20 % in excised leaf tissue (Bell et al. 1994). Whichever organ is the site of reduction *in vivo*, it is suggested that sulphate taken up from the external environment can be stored in the vacuole (Schmidt 1986; Clarkson et al. 1993). A low affinity permease ( $K_m$  for sulphate of 100  $\mu$ M) sensitive to high pH and which may be involved in internal transport of sulphate across the tonoplast has been cloned in *S. hamata* (Smith et al. 1995), but has not been extensively characterised. Measurement of sulphate flux rates from leaf tissue vacuoles indicates that the loss of sulphate from the vacuole may limit sulphate redistribution during sulphate stress (Bell et al. 1994).

Regulation of sulphate uptake and transport in plants is a complicated process because the site of uptake is separated from the normal site of reduction (Warrilow et al. 1995). Furthermore, sulphate uptake must be correlated with nitrate uptake to ensure the correct ratio of sulphur and nitrogen in biological compounds (Smith 1980; Cacco et al. 1983; Clarkson et al. 1989; Cram 1990). It is possible that genes analogous to *Escherichia coli cysB* (Hryniewicz et al. 1989) or

*Neurospora crassa cys3* (Marzluf 1970b), which regulate expression of sulphate transporters, will be present in plants to effect the necessary modulation of sulphate uptake and distribution.

### 1.2.2 Sulphate activation by ATP sulphurylase

Sulphate is a relatively inert compound and must be activated before it can be further metabolised in biochemical reactions (Schiff & Hodson 1973; Schiff & Saidha 1987; Lowe 1991; Schmidt & Jäger 1992). In all organisms studied thus far, the first step of sulphate assimilation is therefore the reaction of sulphate with ATP to produce adenosine 5'-phosphosulphate (APS) and pyrophosphate (PPi), catalysed by ATP sulphurylase (sulphate adenylyltransferase, EC 2.7.7.4) (Bandurski et al. 1956; Anderson 1980; Brunold 1990; Schiff et al. 1993):



The reaction is highly unfavourable for APS formation ( $K_{\text{eq}}$  is  $10^{-7}$  to  $10^{-8}$ ) (Peck & Lissolo 1988; Schwenn 1994), because the free energy of hydrolysis of the sulphate group of APS is higher than the free energy of the phosphate linkage of ATP (Anderson 1980; Schiff & Saidha 1987; Schmidt & Jäger 1992). The forward reaction is accelerated by rapid irreversible catalysis of PPi to orthophosphate (Pi) by the enzyme inorganic pyrophosphatase (EC 3.6.1.1) (Anderson 1980; Salisbury & Ross 1992). In bacteria and mammals, APS formation by ATP sulphurylase and subsequent APS catalysis by the enzyme APS kinase (adenylylsulphate kinase, EC 2.7.1.25) are energetically coupled reactions (Kredich 1987; Lyle et al. 1994). The energetically favourable coupled reaction has been proposed also for higher plants (Schiff & Hodson 1973; Schwenn 1994; Section 1.2.3.2).

The enzymology of ATP sulphurylase has been well studied in

several organisms (Segel et al. 1987; Lowe 1991). In plants, the enzyme has a pH optimum of 8.0 - 9.0, and reported  $K_m$  values of 0.5 mM (sulphate), 0.38 mM (ATP), 0.47  $\mu$ M (APS) and 3  $\mu$ M (PPi) (Balharry & Nicholas 1970; Anderson 1980; Schmidt 1986), similar to those reported for rat chondrosarcoma ATP sulphurylase (Lyle et al. 1994). Those  $K_m$  data emphasize that the enzyme exhibits substantial product inhibition, and requires removal of product for APS formation to occur (Schwenn 1994). The kinetic mechanism of the enzyme is an ordered steady-state single displacement, with MgATP the leading substrate for the forward reaction and APS the leading substrate in the reverse reaction (Lyle et al. 1994). The kinetically active enzyme in *A. thaliana* exists as a homotetramer of about 200 kD (Murillo & Leustek 1995).

ATP sulphurylase has been reported in both photosynthetic and nonphotosynthetic plant tissue, but is associated predominantly with the chloroplast (Ellis 1969; Balharry & Nicholas 1970; Burnell & Anderson 1973; Anderson 1980; Schmidt & Jäger 1992). Two forms of ATP sulphurylase were isolated from *Spinacia oleracea* (spinach) leaves, a major chloroplastic form and a minor cytoplasmic form (Lunn et al. 1990). The existence of a small gene family encoding ATP sulphurylase has been shown in *A. thaliana* (consisting of at least three, and possibly five, members) (Leustek et al. 1994; Klonus et al. 1995; Murillo & Leustek 1995; Logan et al. 1996) and *Solanum tuberosum* (potato) (at least two members) (Klonus et al. 1994). The three cloned *A. thaliana* genes encode deduced proteins which are all targeted to the chloroplast, but whether this duplication represents gene redundancy or whether each of the three isoforms has a different physiological role is unknown at present (Leustek et al. 1994; Klonus et al. 1995; Murillo & Leustek 1995). In chloroplasts, increases in stromal pH during light periods may regulate ATP sulphurylase activity (Schmidt 1986). The presence of a cytoplasmic form of the enzyme appears to be species-specific, and its role is unclear because conditions

in the cytoplasm are extremely unfavourable for the forward reaction (Lunn et al. 1990; Schiff et al. 1993; Murillo & Leustek 1995). ATP sulphurylase does not appear to be present in mitochondria of higher plants, but does occur in that organelle in *Euglena gracilis* (Li et al. 1991).

Sulphate deprivation, or growth on poorly assimilatable sulphur sources such as L-djenkolate or cysteic acid, leads to an increase in ATP sulphurylase activity or transcript level in several plant species (Reuveny & Filner 1977; Reuveny et al. 1980; Zink 1984; Barney & Bush 1985; Haller et al. 1986; Logan et al. 1996). This activation, or derepression, is achieved only under non-limiting reduced nitrogen status of the plants (Reuveny & Filner 1977; Reuveny et al. 1980; Barney & Bush 1985). The sulphur-containing amino acids cysteine and methionine repressed ATP sulphurylase activity in cultured *Nicotiana tabacum* (tobacco) cells (Reuveny & Filner 1977), whereas asparagine and glutamine did not affect activity in *Lemna minor* (Suter et al. 1986). The effect of sulphate and amino acids were not observable *in vitro* (Ellis 1969), indicating that secondary mediators, which must be synthesised *de novo*, effect regulation (Reuveny & Filner 1977). ATP sulphurylase has been shown to be regulated by light (Ghisi et al. 1987; Ferretti et al. 1995) and to be under developmental control (Schmutz & Brunold 1982).

### **1.2.3 Reduction of activated sulphate to sulphide - carrier-bound or free intermediate pathway?**

In plants, the pathway mediating reduction of the activated sulphate (sulpho group) on APS is controversial (Schmidt & Jäger 1992; Schwenn 1994). Early studies with the unicellular green algae *Chlorella pyrenoidosa* (Schmidt 1972; Goldschmidt et al. 1975) and *Euglena gracilis* (Brunold & Schiff 1976), and then *Spinacia oleracea* (Schürmann & Brunold 1980),

indicated that APS was a preferred substrate for reduction, and that 3'-phosphoadenosine 5'-phosphosulphate (PAPS), resulting from phosphorylation of APS, was not an obligatory intermediate. Reduction of the sulpho group of APS to sulphide involving a protein carrier and transferase reactions, the bound intermediate pathway, was proposed for photosynthetic autotrophs such as algae and higher plants (Schmidt 1972; Schiff & Hodson 1973; Schmidt et al. 1974; Schiff & Saidha 1987; Schmidt & Jäger 1992). Another pathway, first elucidated in yeasts and bacteria, required phosphorylation of APS to PAPS before reduction of the sulpho group of PAPS via free intermediates to sulphide (Wilson et al. 1961; Kredich 1971). Enzymes for the free intermediate pathway have been assayed in higher plants (Brunold 1990; Section 1.2.3.2), so it has therefore been unclear whether the carrier-bound and/or free intermediate pathways operate in higher plants *in vivo*. In the next sections, a description of the two pathways will be given, followed by a critical assessment of each.

#### 1.2.3.1 Bound intermediate pathway

##### *Transfer of the sulpho group of APS to a carrier molecule by APS sulphotransferase*

The carrier-bound pathway was advanced after studies on *C. pyrenoidosa* cell-free extracts (Schmidt 1972, 1973; Schmidt et al. 1974). In the first step, the sulpho group of APS is transferred by the enzyme APS sulphotransferase to a carrier molecule to form a bound sulphite intermediate (Schmidt 1972, 1973; Urlaub & Jankowski 1982; Li & Schiff 1992; Schiff et al. 1993). APS sulphotransferase activity has been demonstrated in more than fifty plant families (Schmidt 1975, 1976; Fankhauser & Brunold 1978; Urlaub & Jankowski 1982; Brunold et al. 1987; Suter et al. 1992) and some photosynthetic algae (Schmidt 1972, 1973; Goldschmidt et al.

1975; Tsang & Schiff 1975; Brunold & Schiff 1976; Li & Schiff 1991, 1992). Two possible kinetic mechanisms of the APS sulphotransferase reaction have been postulated. In the sequential mechanism, the enzyme binds with the sulpho group of APS to form an enzyme-substrate complex, and then the sulpho group is transferred to the thiol of an active cysteine residue, forming an S-sulphocysteine residue (Schmidt 1972; Li & Schiff 1992; Schiff et al. 1993). The S-sulphocysteine residue reacts with a carrier molecule (or an enzyme cofactor), transferring the sulpho group to the carrier for subsequent reduction. An alternative parallel kinetic mechanism, whereby the sulpho group of APS is transferred directly to a carrier molecule, without formation of an S-sulphocysteine residue on APS sulphotransferase, has not been excluded (Schmidt 1973; Li & Schiff 1992; Schiff et al. 1993).

Properties of APS sulphotransferase from *C. pyrenoidosa* (Tsang & Schiff 1976a), *E. gracilis* (Li & Schiff 1991), *Nicotiana sylvestris* (Jenni et al. 1980), *Picea abies* (Suter et al. 1992) and *S. oleracea* (Schmidt 1976) have been examined. Studies suggest that the functional enzyme, which is very labile, is a homotetramer of 102 - 115 kD, with 25 kD subunits held together by covalent (probably disulphide) bonds that are dissociated by thiols such as dithiothreitol (DTT) (Li & Schiff 1991; Suter et al. 1992). Each subunit does not display APS sulphotransferase activity (Li & Schiff 1991). Apparent  $K_m$  for APS is 0.1 - 29  $\mu\text{M}$ , with concentrations higher than 50  $\mu\text{M}$  inhibitory (Schmidt 1976; Jenni et al. 1980; Suter et al. 1992). In *S. oleracea*, the enzyme has been located exclusively in the chloroplast stroma (Fankhauser & Brunold 1978). APS sulphotransferase activity has been found to be upregulated during sulphate deprivation (Barney & Bush 1985), and downregulation by the end-product of sulphate reduction, cysteine, has been observed in *N. sylvestris* cell cultures (Jenni et al. 1980). An APS sulphotransferase isolated from the marine macroalga *Porphyra*

*yezoensis* was reported to have different properties to those described for plants and green algae, including a higher native enzyme molecular mass of 350 kD that comprises eight subunits of 43 kD (Kanno et al. 1996).

Identity of the physiological carrier molecule to which APS sulphotransferase transfers a sulpho group has not been clarified, because most thiols will perform the task *in vitro* (Tsang & Schiff 1976a; Schmidt & Jäger 1992; Schiff et al. 1993). The carrier is reported to be glutathione in *C. pyrenoidosa* (Schmidt 1972; Tsang & Schiff 1978). A larger molecule has been implicated as the carrier in *S. oleracea* (Brunold 1990) and a phytochelatin-like molecule was found to be active in the sulphotransferase reaction in *Z. mays* (Schiff et al. 1993).

#### *Reduction of bound sulphite to bound sulphide by thiosulphonate reductase*

The carrier molecule for APS sulphotransferase may be a prosthetic group for both APS sulphotransferase and the second enzyme in the bound intermediate pathway, thiosulphonate reductase (also referred to as ferredoxin:glutathione thiosulphonate reductase or organic thiosulphate reductase) (Schiff & Hodson 1973; Schmidt 1973; Schmidt et al. 1974). In the presence of reduced ferredoxin, thiosulphonate reductase catalyses the reduction of bound sulphite to bound sulphide (Schmidt 1973; Schmidt et al. 1974; Peck & Lissolo 1988; Brunold 1990). Descriptions of thiosulphonate reductase in the literature are exiguous. Partial purification of the enzyme in *C. pyrenoidosa* suggests a molecular weight of 200 kD (Schmidt 1973).

Early studies in which free sulphite and sulphide could not be detected favoured the idea of a thiosulphonate reductase enzyme (Schmidt 1973; Brunold 1990). Furthermore, a *C. pyrenoidosa* mutant which lacked thiosulphonate reductase

activity, but which had sulphite reductase activity, could not grow on sulphate as sole sulphur source or reduce sulphate (Schmidt et al. 1974). It was concluded that the bound pathway was preferred for sulphate assimilation *in vivo*. However, the possibility of parallel pathways involving free sulphite and sulphide was not excluded (Schiff & Hodson 1973). Subsequent data which showed accumulation of free sulphite in *E. gracilis* mitochondria (Saidha et al. 1988), and re-evaluation of their own experiments by the groups of Schiff and Schmidt, further emphasised that there was uncertainty as to whether free or bound intermediate steps were involved after the APS sulphotransferase reaction (Schmidt & Jäger 1992; Schiff et al. 1993).

#### *Incorporation of sulphide into cysteine*

The final step of the proposed bound intermediate pathway is incorporation of carrier-bound sulphide (glutathione-persulphide in *C. pyrenoidosa*) into cysteine by the enzyme *O*-acetylserine (thiol)-lyase, using *O*-acetylserine and requiring two electrons from an unidentified donor (Schmidt et al. 1974; Anderson 1980; Giovanelli et al. 1980; Schiff et al. 1993; Section 1.2.4). Alternative pathways, whereby free sulphide could be formed from bound sulphide in a chemical reaction (sulphidolysis) and then incorporated into cysteine, were deemed feasible (Schiff & Hodson 1973).

#### *Conversion of APS to PAPS as a supply of activated sulphate*

Although APS is the preferred substrate for sulphate reduction in the carrier-bound pathway, it was proposed that cells initially accumulate PAPS, brought about by phosphorylation of APS by an APS kinase (Schmidt 1972; Goldschmidt et al. 1975; Section 1.2.3.2). APS could be formed again by hydrolysis of PAPS by a 3'(2'),5'-diphosphonucleoside 3'(2')-phosphohydrolase (DPNPase) (Goldschmidt et al. 1975; Schiff & Saidha 1987; Peck &

Lissolo 1988). PAPS therefore could act as a reservoir for use in sulphate esterification reactions or for APS formation, and DPNPase would be able to control levels of PAPS and APS (Schiff & Saidha 1987). APS kinase activity has been demonstrated in numerous plant and algal species (Mercer & Thomas 1969; Burnell & Anderson 1973; Stanley et al. 1975; Schwenn & Jender 1981; Urlaub & Jankowski 1982; Kanno et al. 1990; Section 1.2.3.2), while an *Oryza sativa* (rice) cDNA encoding a Mg<sup>2+</sup>-dependent, Ca<sup>2+</sup>-sensitive DPNPase, that catalyses the conversion of PAPS to APS and 3'(2')-phosphoadenosine 5'-phosphosulphate (PAP) to AMP, has been cloned recently (Peng & Verma 1995).

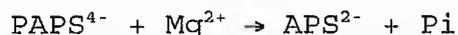
#### *Summary of chemical reactions*

Chemical reactions involved in the bound-intermediate pathway of reductive sulphate assimilation after formation of APS by ATP sulphurylase can be summarised from the above discussion:

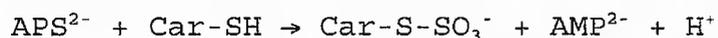
1. Phosphorylation of APS to a storage pool of PAPS; catalysed by APS kinase:



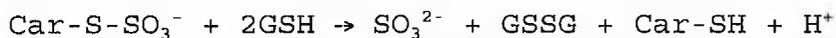
2. Formation of APS from the PAPS pool; catalysed by DPNPase:



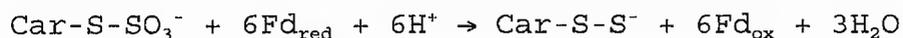
3. Transfer of the sulpho group on APS to a thiol carrier (Car-SH), probably an enzyme prosthetic group; catalysed by APS sulphotransferase:



4. Free sulphite can be formed non-enzymatically from bound sulphite by sulphitolysis in the presence of a thiol (such as glutathione [GSH]):

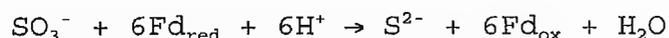


5. (a) Formation of carrier-bound sulphide from carrier-bound sulphite using reduced ferredoxin ( $\text{Fd}_{\text{red}}$ ) as reductant; catalysed by thiosulphonate reductase:

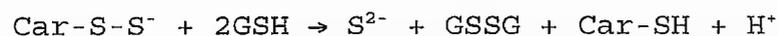


or

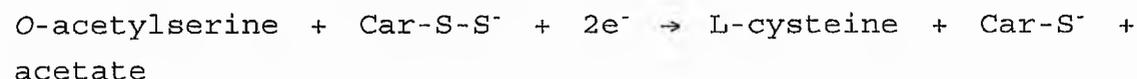
(b) Formation of free sulphide from free sulphite using reduced ferredoxin; catalysed by sulphite reductase:



6. Free sulphide can be formed non-enzymatically from carrier-bound sulphide by sulphidolysis in the presence of an active thiol such as GSH:

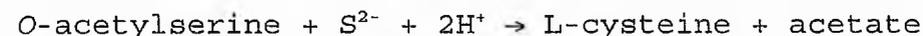


7. (a) Formation of cysteine from bound sulphide using two electrons from an unidentified donor; catalysed by O-acetylserine (thiol)-lyase:



or

(b) Cysteine formation from free sulphide; catalysed by O-acetylserine (thiol)-lyase:



In summary, the bound intermediate pathway involves the reduction of the sulpho group of APS to bound sulphite by APS sulphotransferase, further reduction of bound (or

spontaneously generated free) sulphite to bound or free sulphide, and incorporation of sulphide into cysteine (Schmidt et al. 1974; Anderson 1980; Schmidt & Jäger 1992). Proponents of the bound intermediate pathway proposed initially that carrier-bound sulphite reduction was favoured by plants *in vivo* (Schiff & Hodson 1973; Schmidt 1973; Schmidt et al. 1974; Goldschmidt et al. 1975), but subsequently were less certain of the biochemical evidence and concluded that genetic methods and downregulation of specific enzymes would be required to determine the physiological steps involved (Schmidt & Jäger 1992; Schiff et al. 1993).

#### 1.2.3.2 Free intermediate pathway

In heterotrophic and some photosynthetic bacteria, in yeasts and in fungi, the pathway of sulphate reduction has been established to involve phosphorylation of APS to PAPS followed by reduction of the sulpho group of PAPS to form free sulphite, which is then reduced to free sulphide (Bandurski et al. 1956; Robbins & Lipmann 1956; Asahi et al. 1961; Wilson et al. 1961; Kredich 1971, 1987; Peck & Lissolo 1988). This free intermediate pathway has been shown genetically to be the sole pathway for sulphate reduction in *Escherichia coli* and *Salmonella typhimurium* (Dreyfuss & Monty 1963; Jones-Mortimer 1968; Tei et al. 1990a; Hunt et al. 1987; Kredich 1987), and the genes encoding the sulphate reduction enzymes have been cloned in those enterobacteria (Li et al. 1987; Ostrowski et al. 1989a, b; Krone et al. 1990b, 1991). Using the *E. coli* PAPS reductase (*cysH*) gene as probe, Krone et al. (1990a) detected apparently homologous sequences in the genomic DNA of several higher plants. As the first step of the free intermediate pathway, phosphorylation of APS by APS kinase to form PAPS, had already been demonstrated in plants, Krone et al. (1990a) proposed that sulphate reduction in plants was analogous to the bacterial

pathway, and questioned the involvement of APS sulphotransferase and thiosulphonate reductase. Workers from the same group subsequently noted amino acid, nucleotide or functional homologies between higher plant, cyanobacterial and enterobacterial sulphate assimilation enzymes (Gisselmann et al. 1992; Arz et al. 1994), and therefore proposed that the free intermediate pathway of sulphate assimilation operated in plants exclusively, utilising reduced ferredoxin as reductant.

#### *Phosphorylation of APS to PAPS by APS kinase*

Phosphorylation of the sulpho group on APS to form PAPS by APS kinase, the first step of the free intermediate pathway, has been detected in several higher plants (Mercer & Thomas 1969; Burnell & Anderson 1973, Stanley et al. 1975; Urlaub & Jankowski 1982). An *Arabidopsis thaliana* gene encoding APS kinase has been isolated independently in two laboratories by: (i) screening an *A. thaliana* cDNA library with a degenerate nucleotide probe corresponding to a conserved region in polypeptide sequences of *E. coli* (Leyh et al. 1988; Satishchandran et al. 1992), *Rhizobium meliloti* (Schwedock & Long 1989) and *Saccharomyces cerevisiae* (Cherest et al. 1987; Korch et al. 1991) APS kinases (Arz et al. 1994); and (ii) functional complementation of a *S. cerevisiae* Met14 mutant, deficient in APS kinase, with an *A. thaliana* cDNA library (Jain & Leustek 1994). As expected, extensive similarities were observed between the plant, bacterial and yeast enzymes (Arz et al. 1994).

Mature APS kinase from *A. thaliana* has a molecular weight of 23 kD, is targeted to the chloroplast, and contains an ATP-binding site and PAPS binding motif (Arz et al. 1994; Jain & Leustek 1994). Although detailed studies of plant APS kinase have not been undertaken, the enzyme from plants, algae and bacteria appears to have similar kinetic properties (Krone et al. 1990a; Schwenn 1994). APS kinase from *E. coli*

(Satishchandran & Markham 1989; Satishchandran et al. 1992), *Penicillium chrysogenum* (Renosto et al. 1984, 1985) *Porphyra yezoensis* (Kanno et al. 1990) and *Chlamydomonas reinhardtii* (Jender & Schwenn 1984) is very efficient, with  $K_m$  (APS) of 0.25 - 2  $\mu$ M. At low substrate concentrations, the *E. coli* enzyme is probably diffusion limited (Satishchandran et al. 1992). Substrate inhibition by APS as a result of dead-end complex formation has been described for both prokaryotic (Satishchandran et al. 1992) and eukaryotic (Schwenn & Jender 1981; Renosto et al. 1984, Kanno et al. 1990) APS kinases, but the enzyme from *Spinacia oleracea* may be affected less (Burnell & Whatley 1975).

The high affinity of APS kinase for APS indicates that coupling of ATP sulphurylase and APS kinase activities would allow greater sulphate activation and overcome the unfavourable reaction equilibrium for APS formation (Schiff & Saidha 1987; Segel et al. 1987; Schwenn 1994). Association of ATP sulphurylase and APS kinase to form a reaction complex is not necessary for PAPS accumulation at physiological substrate levels (Schwenn 1994), and these enzymes were found not to associate in *Penicillium chrysogenum* (Renosto et al. 1989). A single multifunctional protein, PAPS synthase, which has very high identity (approximately 60%) with both *A. thaliana* ATP sulphurylase (Murillo & Leustek 1995) and *A. thaliana* APS kinase (Jain & Leustek 1994), has been cloned in the marine worm *Urechis caupo* (Rosenthal & Leustek 1995), but an analogous enzyme has not been reported for plants.

#### *Reduction of PAPS by PAPS reductase to form sulphite*

The second step of the free intermediate pathway was contended to involve reduction of the sulpho group of PAPS to form sulphite using reduced ferredoxin as reductant, catalysed by PAPS reductase (EC number unassigned) (Schwenn 1994). Enzyme activity has been shown in *S. oleracea*, where bacterial thioredoxin was used as reductant, but reported to

be detectable also in *Nicotiana tabacum* cell cultures and *Brassica pekinensis* (cabbage) (Schwenn 1989). Partial purification of the *S. oleracea* enzyme suggests a molecular weight of 68 - 72 kD (Schwenn 1989). Further evidence for existence of a plant PAPS reductase comes from hybridisation studies which infer sequence homology between *E. coli cysH* and genomic DNA in yeasts, cyanobacteria and plants (Krone et al. 1990a). Genes encoding PAPS reductases have been isolated from *E. coli* (Li et al. 1987; Krone et al. 1991), *Saccharomyces cerevisiae* (Thomas et al. 1990), *Thiocapsa roseopersicina* (Haverkamp, T. Gisselmann, G., Schwenn, J.D. 1993, unpublished GenBank entry Z23169) and the cyanobacterium *Synechococcus* PCC7942 (Niehaus et al. 1992). The kinetic mechanism of PAPS reductase in plants is proposed by Schwenn (1994) to be homologous to that in bacteria (Berendt et al. 1995) and *Saccharomyces cerevisiae* (Schwenn et al. 1988), ie. a two-step reaction in which two electrons from the dithiol group of reduced thioredoxin are transferred to the enzyme, followed by reduction of PAPS by the enzyme.

There has been some disagreement over whether reduction of the sulpho group of PAPS involves formation of free or bound sulphite. In *E. coli*, it has been proposed that the enzyme acts as a PAPS-specific sulphotransferase, with thioredoxin a thiol acceptor that forms a bound sulphite derivative which eventually yields free sulphite (Tsang & Schiff 1976b, 1978; Kredich 1987). A similar PAPS sulphotransferase has been proposed for higher plants (Brunold 1990). However, detailed kinetic studies in *S. cerevisiae* (Schwenn et al. 1988) and *E. coli* (Tsang 1981; Berendt et al. 1995) show that thioredoxin reduces the enzyme first, rather than acting as a carrier, and that the enzyme is therefore a PAPS reductase. In *S. cerevisiae*, it was found that formation of a bound sulphite followed rather than preceded free sulphite formation (Schwenn et al. 1988).

## *Reduction of sulphite to sulphide by sulphite reductase*

In the next step of the free intermediate pathway, free sulphite is reduced to sulphide by a ferredoxin-dependent sulphite reductase (EC 1.8.1.2) (Anderson 1980; Brunold 1990; Schmidt & Jäger 1992; Schwenn 1994). Presence of sulphite reductase is ubiquitous within higher plant genera (Asada 1967; Asada et al. 1968; Hennies 1975; Sawhney & Nicholas 1975; Aketagawa & Tamura 1980; von Arb & Brunold 1983; Schwenn & Kemena 1984; Schmutz & Brunold 1985; von Arb & Brunold 1985; Koguchi et al. 1988; Dittrich et al. 1992). An early study with a partially purified enzyme from *Spinacia oleracea* leaves showed that six electrons were required for the formation of one molecule of sulphide (Asada 1967). Whereas sulphite reductase from *E. coli* (EC 1.8.1.2) was found to be dependent on reduced nicotinamide-adenine dinucleotide phosphate (NADPH) (Kredich 1987), plant sulphite reductase could not be reduced by NADPH (Asada 1967) and was found rather to require reduced thioredoxin as cofactor (Hennies 1975; Aketagawa & Tamura 1980).

Sulphite reductase from *E. coli* and *Salmonella typhimurium* is a large (670 kD), complex enzyme with two different polypeptides: a flavoprotein ( $\alpha$ ) subunit which contains flavin-adenine dinucleotide (FAD) and flavin mononucleotide (FMN) prosthetic groups, and a haemoprotein ( $\beta$ ) subunit which contains an  $\text{Fe}_4\text{S}_4$  cluster and a novel sirohaem, with the holoenzyme having an  $\alpha_8\beta_4$  structure (Siegel et al. 1973; Siegel & Davis 1974; Jackson et al. 1981; Ostrowski et al. 1989a, b; Young & Siegel 1990; Kaufman et al. 1993). Plant sulphite reductase, however, was found to be a homodimer or homotetramer with a single subunit each having one mole of sirohaem and one  $\text{Fe}_4\text{S}_4$  cluster (Aketagawa & Tamura 1980; Brunold 1990). NADPH-sulphite reductase and ferredoxin-sulphite reductase are thought to have similar catalytic mechanisms: sulphite binds to iron in the  $\text{Fe}_4\text{S}_4$  cluster through the sulphur atom, and is then reduced to sulphide by

a series of two electron cleavages of the sulphur-oxygen bond without the release of intermediates (Kredich 1987; Tan & Cowan 1991; Schwenn 1994).

Recently, an *A. thaliana* gene encoding sulphite reductase has been isolated (Bruehl, A., Haverkamp, T., Gisselmann, G., Schwenn, J.D. 1995, unpublished GenBank accession Z49217). The deduced polypeptide has an amino-terminal transit peptide for import into the chloroplast, while the mature protein of 576 amino acids has a predicted size of 65.3 kD, a similar size to the ferredoxin-sulphite reductase from *Synechococcus* PCC7942 (Gisselmann et al. 1993). Sequence comparison reveals that the plant and cyanobacterial ferredoxin-sulphite reductases have greatest similarity (57.6 % and 58.7 %, respectively) with the haemoprotein  $\beta$ -subunit of *E. coli* sulphite reductase, encoded by the *cysI* gene (Li et al. 1987; Gisselmann et al. 1993).

#### *Incorporation of sulphide into cysteine*

The final step of the free intermediate pathway is formation of cysteine from sulphide and *O*-acetylserine, catalysed by the enzyme *O*-acetylserine (thiol)-lyase (Anderson 1980; Giovanelli et al. 1980; Schwenn 1994; Section 1.2.4), as has been suggested for the bound-intermediate pathway.

#### *Summary of chemical reactions*

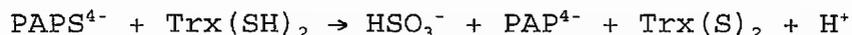
Reactions involved in the free intermediate pathway for cysteine formation proposed for higher plants following sulphate activation can be summarised as:

1. Phosphorylation of APS to PAPS; catalysed by APS kinase:

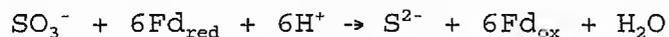


2. Reduction of PAPS to free sulphite using reduced

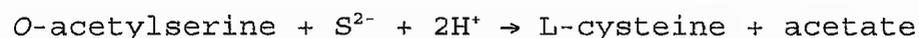
thioredoxin [Trx(SH)<sub>2</sub>] as reductant; catalysed by PAPS reductase:



3. Formation of free sulphide from free sulphite using reduced ferredoxin; catalysed by sulphite reductase:



4. Cysteine formation from free sulphide; catalysed by O-acetylserine (thiol)-lyase:



Proponents of the free intermediate pathway, predominantly Schwenn and co-workers (Krone et al. 1990a; Gisselmann et al. 1992; Arz et al. 1994; Schwenn 1994), argue therefore that reductive sulphate assimilation in higher plants is analogous to the pathway in various prokaryotes and lower eukaryotes. In this free intermediate pathway, APS is phosphorylated to form PAPS, free sulphite is produced from reduction of the sulpho group of PAPS, and then free sulphite is reduced to free sulphide and incorporated into cysteine.

#### 1.2.3.3 Comparison of the free and bound intermediate pathways of sulphate assimilation

##### *Reactivity of sulphur-containing compounds*

The spontaneous reaction of different sulphur-containing compounds with each other has made interpretation of many biochemical results ambiguous, as has been noted (Schiff & Hodson 1973; Schmidt & Jäger 1992; Schiff et al. 1993). For example, when sulphite and sulphide are both present in a

solution, several sulphur compounds of intermediate oxidation state can arise non-enzymatically (Schiff et al. 1993). Thus some workers (Schmidt 1973; Schiff & Hodson 1973) proposed that central intermediates in sulphate reduction were bound to carriers, and that these may give rise to free sulphite, sulphide and thiosulphate via reactions that are independent from the main pathway. The antithetical argument, that incidental bound sulphur compounds arise non-enzymatically from free intermediates of sulphate reduction, has also been advanced (Schwenn 1994). However, a critical evaluation of most experimental data cannot differentiate between free or bound-intermediate formation as principal steps (Schiff & Hodson 1973; Schmidt & Jäger 1992).

#### *Evaluation of sulphate assimilation enzyme assays*

Further difficulty in interpreting results arises from the non-specific and non-physiological nature of several of the enzyme assays routinely employed. Primary examples are:

(i) The established APS sulphotransferase assay involves measuring acid-volatile radioactivity produced from  $^{35}\text{S}$ -APS as substrate, in the presence of a thiol or dithiol and non-radioactive carrier sulphite (Schiff & Levinthal 1968; Brunold 1990). However, a number of acid-volatile compounds are possible, depending on the product formed from the reaction. Acidification of sulphite forms water and volatile sulphur dioxide, acidified thiosulphate forms sulphur, water and volatile sulphur dioxide, and acidification of sulphide forms volatile hydrogen sulphide (Schiff & Hodson 1973; Tsang & Schiff 1976a). The specific assay products formed from  $^{35}\text{S}$ -APS therefore can be identified only if separated experimentally, for example by chromatography (Levinthal & Schiff 1968).

Furthermore, most thiol or dithiol molecules tested were found to substitute for the hypothetical carrier molecule of

the APS sulphotransferase reaction (Schiff et al. 1993). At the concentrations used, the dithiols DTT and dithioerythritol (DTE) yielded optimal levels of free radioactive sulphite in the reaction, and therefore were employed most frequently (Schmidt 1972, 1975; Tsang & Schiff 1975; Fankhauser & Brunold 1978; Brunold et al. 1987; Li & Schiff 1991; Suter et al. 1992). While the *in vitro* APS sulphotransferase assay measured free sulphite formation, the *in vivo* reaction was assumed to involve a carrier molecule which accepted the sulpho group of APS (Brunold 1990; Schiff et al. 1993). The lack of specificity of APS sulphotransferase for different thiols proved inconvenient for those attempting to identify the hypothetical endogenous carrier thiol (Schiff et al. 1993).

Addition of nonradioactive sulphite as carrier molecule for volatilisation allowed for exchange of the radioactive label from the reaction product with sulphite and release of volatile sulphur dioxide as an assay artifact (Levinthal & Schiff 1968; Saidha & Schiff 1989). It is feasible to obtain the preferred product for the assay by selecting the appropriate carrier molecule (Schiff & Hodson 1973; Tsang & Schiff 1976a).

The APS sulphotransferase assay, although convenient and sensitive (Schiff & Hodson 1973), lacks specificity and cannot provide evidence for the mechanism of APS reduction. Importantly, the assay cannot differentiate between a sulphotransferase reaction, whereby the sulpho group is transferred to a carrier thiol, and a reductase reaction, whereby APS is reduced to release free sulphite;

(ii) PAPS reductase activity has been quantified using a system analogous to the APS sulphotransferase assay, but with <sup>35</sup>S-PAPS as substrate, and measuring acid-volatile radioactivity in the presence of reduced thioredoxin, DTT (as auxiliary reductant to ensure regeneration of reduced

thioredoxin) and carrier sulphite (Schwenn et al. 1988; Schwenn 1989). An identical assay has been employed to measure PAPS sulphotransferase activity (Imhoff 1982; Brunold 1990), so it is apparent that the assay cannot differentiate between reductase and sulphotransferase reactions. As for the APS sulphotransferase assay, addition of non-radioactive sulphite as a carrier molecule prevents identification of the original enzyme product. In view of the notable effects of thiols in sulphur reducing reactions (Schiff et al. 1993), clarification of the actual role of DTT in the assay is required (for example, by using the thioredoxin-NADP-reductase system as an alternative for maintaining reduced thioredoxin [Schmidt 1973]). This PAPS reductase/sulphotransferase assay can therefore be used justifiably only to demonstrate that the sulpho group on PAPS is reduced to a compound which can form, or exchange with, acid-volatile product(s); and

(iii) The prototype thiosulphonate reductase assay employed radioactive sulphogluthione ( $\text{GS-}^{35}\text{SO}_3^-$ ) as a carrier-bound sulphite, and measured ferredoxin-dependent acid-volatile radioactive sulphide ( $^{35}\text{S}^{2-}$ ) liberated from the presumed reaction product  $\text{GS-}^{35}\text{S}$  by exchange with nonradioactive carrier sulphide (Schmidt 1973; Schmidt et al. 1974). Similar to the APS and PAPS reductase/sulphotransferase assays, actual products of thiosulphonate reductase activity, and hence the enzyme reaction, cannot be ascertained from the assay. A simplified thiosulphonate reductase assay, which used methylviologen oxidation to methyl blue by dithionite (Schmidt 1973), also measures sulphite reductase activity (Brunold 1990).

*Is PAPS an obligatory intermediate in plant sulphate assimilation?*

Given that sulphur compounds can be highly reactive with each other, making identification of primary rather than

incidental intermediates difficult, and also that many of the crucial enzyme assays which have been used to further arguments for the existence of both free and bound intermediates cannot differentiate between the two pathways, it can be proposed that the most significant discrepancy in this field is whether APS or PAPS serve as the primary sulpho donor for sulphate reduction in plants. Biochemical studies of assimilatory sulphate reducers in which the ability to reduce  $^{35}\text{S}$ -APS and/or  $^{35}\text{S}$ -PAPS to acid-volatile radioactivity was determined led to suggestions that these organisms should be divided into two broad groups: most photosynthetic eukaryotes (including higher plants and green algae) would use APS as the immediate sulphate donor for sulphate reduction, while heterotrophic prokaryotes (for example *Escherichia coli* and *Salmonella typhimurium*) and eukaryotes (including certain fungi such as yeasts) would use PAPS (Schiff & Hodson 1973; Tsang & Schiff 1975; Anderson 1980; Schmidt & Jäger 1992; Schiff et al. 1993). Although some families of phototrophic bacteria and cyanobacteria contain species that can utilise either sulphonucleotide (Imhoff 1982; Schmidt & Jäger 1992), a generalisation that APS utilisation is a characteristic of chloroplasts and chloroplast-containing organisms has been proposed (Tsang & Schiff 1975). In attempting to resolve the controversy over the free or bound intermediate pathways, it is therefore useful to examine the evidence regarding the primary sulphate donor for reduction in plants.

Studies with *Chlorella pyrenoidosa* and *Euglena gracilis* showed that both  $^{35}\text{S}$ -APS and  $^{35}\text{S}$ -PAPS could be converted by cell-free extracts to acid-volatile radioactivity (Schmidt 1972; Goldschmidt et al. 1975; Brunold & Schiff 1976). When non-radioactive PAPS was included in the assay with  $^{35}\text{S}$ -APS, slight changes in acid-volatile activity were observed, whereas non-radioactive APS severely reduced the amount of acid-volatile radioactivity obtained from  $^{35}\text{S}$ -PAPS (Brunold & Schiff 1976). Similar results were obtained subsequently

using extracts from *Spinacia oleracea* leaf chloroplasts (Schürmann & Brunold 1980). Interpretation of those results was that APS is the substrate for sulphate reduction in green algae and higher plants, and that PAPS requires prior conversion to APS before reduction (Brunold & Schiff 1976; Schürmann & Brunold 1980). Those experiments could not rule out the possibility that APS was preferred if both substrates were present together but that PAPS could be reduced in the absence of APS. In *C. pyrenoidosa*, however, it was found that PAPS could only serve as substrate for sulphate reduction if a crude enzyme fraction containing 3'-nucleotidase activity was present (Hodson & Schiff 1971; Schmidt 1972). Partially purified enzyme fraction "A" subsequently was characterised further as a 3'-nucleotidase that converted PAPS to APS, while fraction "S" contained APS-reducing activity (Goldschmidt et al. 1975). That work indicated that in *C. pyrenoidosa* sulphate reduction proceeded from APS, and that PAPS could serve as substrate by conversion to APS (Schmidt 1972; Goldschmidt et al. 1975). However, Brunold (1990) points out that the *in vitro* assay conditions used in these studies may not have been suitable for detecting PAPS sulphotransferase (or PAPS reductase) activity. More definitive analyses on the preferred substrate for sulphate assimilation in plants have not been conducted.

Descriptions of APS sulphotransferase activity in higher plants are legion (Schmidt 1975, 1976; Fankhauser & Brunold 1978; Urlaub & Jankowski 1982; Brunold et al. 1987; Suter et al. 1992). Recently, the relevance of this APS sulphotransferase activity has been challenged. When an *A. thaliana* cDNA encoding APS kinase was expressed in *Escherichia coli*, it was found to confer both APS kinase and APS sulphotransferase activity (Arz et al. 1994). *A. thaliana* APS kinase in its monomeric form of about 28 kD shows APS kinase activity, but a homotetrameric form exhibiting APS sulphotransferase activity can be isolated in the absence of reducing thiols (Schiffmann & Schwenn 1994). Rates of APS

sulphotransferase activity of the plant APS kinase tetramer (Schiffmann & Schwenn 1994) correspond well with reported values for APS sulphotransferase from *Spinacia oleracea* (Schmidt 1976) and *Euglena gracilis* (Li & Schiff 1991). Other similarities between the APS kinase tetramer and APS sulphotransferase functional enzyme include lability in the presence of thiols, molecular weight of monomer subunit and inability of the subunits to display APS sulphotransferase activity (Li & Schiff 1991; Suter et al. 1992; Schiffmann & Schwenn 1994). If experiments from Schwenn's group are verified, there will be little doubt that APS sulphotransferase activity in plants can be conferred by APS kinase. While Schiffmann and Schwenn (1994) conclude that APS sulphotransferase activity has no physiological significance for plants and is merely an artifactual byproduct of APS kinase, it needs to be demonstrated that the tetrameric form of APS kinase does not occur *in vivo* and function as a sulphotransferase. It is difficult to explain at present why *A. thaliana* APS kinase can yield APS sulphotransferase activity but *Escherichia coli* APS kinase cannot (Schiffmann & Schwenn 1994), or why radioactivity from <sup>35</sup>S-APS becomes bound to the APS sulphotransferase native enzyme (the APS kinase tetramer?) in *Euglena gracilis* (Li & Schiff 1992). Even if APS sulphotransferase activity can be a byproduct of APS kinase, that APS might be a preferred substrate for sulphate assimilation in plants is not proved or disproved. The existence of a different form of APS sulphotransferase in the macroalga *Porphyra yezoensis*, which has a subunit size of 43 kD and aggregates as an homo-octomer (Kanno et al. 1996), raises the possibility that enzymes as yet uncharacterised might be responsible for observed APS sulphotransferase activity in higher plant extracts.

Two possibilities for the mechanism of APS reduction can be inferred from experimental data: either the sulpho group on APS is transferred to an acceptor carrier molecule by APS sulphotransferase (Schmidt 1972; Abrams & Schiff 1973;

Schmidt 1973; Tsang & Schiff 1978; Li et al. 1991), or APS sulphotransferase is a thiol-dependent reductase (Schmidt & Jäger 1992). Considerable evidence relating to a sulpho group carrier in *C. pyrenoidosa* has been accumulated (Abrams & Schiff 1973; Schiff et al. 1993), but a carrier in plants has not been found (Schmidt & Jäger 1992). Transferase and reductase reactions cannot be differentiated in the enzyme assays used, and it is again difficult to determine whether bound or free intermediates are formed originally (Schmidt & Jäger 1992; Schwenn 1994).

The presence of a thiosulphonate reductase, which would reduce the sulpho group attached to a carrier molecule to carrier-bound sulphite, was suggested from experiments in *C. pyrenoidosa* in which radioactive GS-<sup>35</sup>S<sub>0</sub> was incorporated into protein at rates similar to radioactive sulphur (which could exchange with H<sub>2</sub>S) (Schwenn 1972). A *C. pyrenoidosa* mutant Sat<sub>2</sub><sup>-</sup> which could not grow on sulphate as a sole sulphur source and that lacked thiosulphonate reductase activity but had sulphite reductase activity was presented as further evidence in favour of an obligatory role for thiosulphonate reductase in sulphate assimilation (Schmidt et al. 1974). However, thiosulphonate reductase activity in the Sat<sub>2</sub><sup>-</sup> mutant was lacking only when β-mercaptoethanol was used as reductant, but comparable to wild-type levels in the presence of thiols like DTT (Schmidt et al. 1974), so the nature of the enzyme lesion in the mutant is not clear. In *Escherichia coli*, sulphite reductase can exhibit both thiosulphonate reductase and sulphite reductase activities (Tsang & Schiff 1976b). Furthermore, thiosulphonate reductase has not been characterised further in plants, and its existence is uncertain (Schwenn 1994). If a persulphide is formed by thiosulphonate reductase activity, two electrons would be necessary to release sulphide for incorporation into cysteine by O-acetylserine (thiol)-lyase (Schiff & Hodson 1973; Tsang & Schiff 1975). This requirement has not been demonstrated in plants or algae (Schwenn 1994).

## Conclusions

The ready cross-reaction of sulphur compounds has made it difficult to identify the intermediates of sulphate assimilation in plants. Design of specific, physiological and simple enzyme assays has not been achieved, and more detailed assays to elucidate kinetic mechanisms are required. While several steps of the proposed bound-intermediate pathway are not understood, there is substantial evidence that designates APS as the immediate sulphate donor for reduction. However, all of the enzymes of the free intermediate pathway which uses PAPS as sulphate donor have been identified in plants, and this pathway has been well described in other assimilatory sulphate reducers such as *E. coli*. Is it possible that more than one pathway operates in plants? Genetic methods for the isolation of genes involved and then specific downregulation of various steps in the sulphate assimilation pathway, especially the reduction of APS or PAPS, will probably be necessary to assist in resolution of the controversies (Schmidt & Jäger 1992).

### 1.2.4 Biosynthesis of cysteine

Plants have several mechanisms for incorporating sulphide into sulphur-containing amino acids, but the formation of cysteine by the enzyme *O*-acetylserine (thiol)-lyase (EC 4.2.99.8) using sulphide and *O*-acetyl-L-serine is the most important one (Anderson 1980; Giovanelli et al. 1980). *De novo* cysteine biosynthesis can be regarded as a fundamental reaction in biology due to the requirement of its derivative methionine in animal diets (Giovanelli 1990). Cysteine biosynthesis is a reaction requiring products from both sulphate reduction (sulphide) and nitrate assimilation (*O*-acetylserine), and *O*-acetylserine has been implicated in cross-regulation of the converging pathways (Giovanelli 1990; Neuenschwander et al. 1991; Brunold 1993; Chapter 3.1).

Furthermore, O-acetylserine may have a role in regulating flux through the sulphate assimilation pathway (Buwalda et al. 1992; Rennenberg 1983; Chapter 3.1).

#### *Formation of O-acetylserine by serine acetyltransferase*

Formation of O-acetylserine is catalysed by serine acetyltransferase (EC 2.3.1.30) from the substrates L-serine and acetyl coenzyme A (acetyl CoA). Serine acetyltransferase activity has been detected in several plants, including *Allium tuberosum* (leek) (Nakamura et al. 1987; Nakamura & Tamura 1990), *Brassica* sp. (Ngo & Shargool 1974; Nakamura et al. 1987, 1988), *Phaseolus vulgaris* (kidney bean) (Smith & Thompson 1971; Smith 1972), *Pisum sativum* (pea) (Ruffet et al. 1995), *Raphanus sativus* (radish) (Nakamura et al. 1987), *Spinacia oleracea* (Brunold & Suter 1982; Nakamura et al. 1987) and *Triticum aestivum* (Ascaño & Nicholas 1977). Isoforms of the enzyme have been located in the cytosol, mitochondria and chloroplasts (Smith 1972; Brunold & Suter 1982; Ruffet et al. 1994, 1995). Recently, genes encoding different isoforms of serine acetyltransferase have been described for *Arabidopsis thaliana* (Bogdonova et al. 1995; Ruffet et al. 1995; Roberts & Wray 1996; Howarth et al. 1997) and a gene encoding a cytosolic form of the enzyme has been cloned in *Citrullus vulgaris* (watermelon) (Saito et al. 1995).

Purified serine acetyltransferase has a subunit molecular mass of 31 - 33 kD separated on sodium dodecyl sulphate (SDS)-polyacrylamide gels (Nakamura & Tamura 1990; Ruffet et al. 1994), in agreement with a calculated mass of 31.5 - 34.3 kD (Ruffet et al. 1995; Saito et al. 1995). Purified or partially-purified plant serine acetyltransferase has reported  $K_m$  values for acetyl CoA of 90  $\mu$ M - 0.35 mM and for L-serine of 30  $\mu$ M - 2.3 mM (Smith & Thompson 1971; Ngo & Shargool 1974; Brunold & Suter 1982; Ruffet et al. 1994). Experimental data indicate that the enzyme follows a double

displacement Ping Pong Bi Bi kinetic mechanism (Ruffet et al. 1994). Whereas early studies showing cysteine inhibition of partially purified plant enzyme (Smith & Thompson 1971; Brunold & Suter 1982) were criticised for using nonphysiological levels of the amino acid (Giovanelli 1990), recombinant enzyme from *C. vulgaris* is inhibited non-competitively by low levels of cysteine, showing 50 % inhibition at 2.9  $\mu\text{M}$  (Saito et al. 1995). Plant serine acetyltransferase has similar properties to the bacterial enzyme which catalyses the same reaction (Kredich & Tomkins 1966; Jones-Mortimer et al. 1968; Leu & Cook 1994).

#### *Cysteine formation by O-acetylserine (thiol)-lyase*

O-acetylserine (thiol)-lyase activity was demonstrated first in *S. oleracea* (Giovanelli & Mudd 1967) and then *Brassica napus* (turnip) (Thompson & Moore 1968). Subsequently, activity has been documented widely in plants (Smith & Thompson 1971; Smith 1972; Ngo & Shargool 1973, 1974; Masada et al. 1975; Ascaño & Nicholas 1977; Ng & Anderson 1978; Burnell 1984). Multiple forms of O-acetylserine (thiol)-lyase localised in cytosol, mitochondria and chloroplasts have been reported (Fankhauser & Brunold 1978; Giovanelli et al. 1980; Burnell 1984; Schiff et al. 1993; Kuske et al. 1996), and several of these isozymes have been isolated from algae and plants (Diessner & Schmidt 1981; Murakoshi et al. 1986; Nakamura & Tamura 1989; Lunn et al. 1990; Leon & Vega 1991; Rolland et al. 1992; Römer et al. 1992; Ikegami et al. 1993; Yamaguchi & Masada 1995). Lunn et al. (1990) suggested that the enzyme is present in all cellular compartments where protein synthesis is required because cysteine may not be readily transportable. Isoforms of O-acetylserine (thiol)-lyase have been cloned in *A. thaliana* (Hell et al. 1994; Barroso et al. 1995; Hesse & Altmann 1995), *Capsicum anuum* (bell pepper) (Römer et al. 1992), *Citrullus vulgaris* (Noji et al. 1994), *T. aestivum* (Youssefian et al. 1993) and *Zea mays* (Brander et al. 1995), and in *S. oleracea* the cytosolic

(Saito et al. 1992), mitochondrial (Saito et al. 1994b) and chloroplastic (Hell et al. 1993; Rolland et al. 1993a; Saito et al. 1993b) members of a small gene family have been described.

O-acetylserine (thiol)-lyase isoforms from different plants have reported molecular masses of 26 - 40 kD (Murakoshi et al. 1985; Ikegami et al. 1988; Droux et al. 1992; Römer et al. 1992; Saito et al. 1992; Rolland et al. 1993a; Ikegami et al. 1993). Enzyme activity is dependent on pyridoxal 5'-phosphate (Masada et al. 1975), and the functional lysine residue that acts as binding site has been located by site-directed mutagenesis in *S. oleracea* (Saito et al. 1993a). Kinetic studies on the enzyme suggests low affinity for O-acetylserine, with  $K_m$  values of 0.24 - 6.7 mM reported, but a higher affinity for sulphide ( $K_m = 22 \mu\text{M} - 2.5 \text{mM}$ ) (Ascaño & Nicholas 1977; Ng & Anderson 1978; Murakoshi et al. 1985, 1986; Ikegami et al. 1988; Droux et al. 1992; Ikegami et al. 1993; Yamaguchi & Masada 1995). Variation of  $K_m$  values between different isoforms may explain the wide range published. Although all isozymes are capable of synthesising cysteine from O-acetylserine and sulphide, some have ability to synthesise other compounds such as the non-protein amino acid L-quisqualic acid (Murakoshi et al. 1985, 1986; Ikegami et al. 1988, 1990, 1992, 1993). In addition to cysteine biosynthesis, isoforms of O-acetylserine (thiol)-lyase may have a role in protecting plants against endogenous or exogenous toxins such as hydrogen sulphide gas or cyanide (Ikegami et al. 1988; Youssefian et al. 1993).

#### *A multifunctional complex for cysteine biosynthesis*

Most preparations of serine acetyltransferase were found to have O-acetylserine (thiol)-lyase activity (Smith & Thompson 1971; Nakamura et al. 1987, 1988; Nakamura & Tamura 1990; Droux et al. 1992). In *E. coli* and *Salmonella typhimurium*, a multifunctional enzyme (previously referred to as "cysteine

synthetase") is formed as a supramolecular complex of serine acetyltransferase and O-acetylserine (thiol)-lyase, and kinetic properties of the individual component enzymes are modified in the complex such that substrate affinity is enhanced (Kredich et al. 1969). Indirect evidence in favour of such a complex in plants is that serine acetyltransferase and O-acetylserine (thiol)-lyase have been found to be associated during purification, with a size comparable to the bacterial complex (Kredich et al. 1969; Droux et al. 1992; Nakamura et al. 1988; Nakamura & Tamura 1990). This association was confirmed using enzyme-specific antibodies which showed aggregation of recombinant forms of *Spinacia oleracea* serine acetyltransferase and O-acetylserine (thiol)-lyase (Saito et al. 1995). In addition, comparative analysis of plant and bacterial enzymes shows kinetic and structural similarities, with at least plant O-acetylserine (thiol)-lyase probably originating from a prokaryotic ancestor (Rolland et al. 1993b). The discrepancy between mM ranges for substrate affinity of each individual plant enzyme and lower measured *in vivo* substrate concentrations (Schmidt & Jäger 1992) could be resolved as the complex is expected to be more efficient than the component enzymes due to substrate channelling (Droux et al. 1992; Saito et al. 1995), as has been found for the bacterial complex (Kredich et al. 1969).

### 1.3 Functional complementation - the cloning method of choice?

Due to difficulties encountered using biochemical techniques in elucidating the steps involved in sulphate assimilation, particularly at the stage of APS or PAPS reduction, it has been suggested that the solution would be found in genetic methods and down-regulation of specific enzymes (Schmidt & Jäger 1992; Schiff et al. 1993). At the inception of this study, genes encoding only one enzyme of the sulphate assimilation pathway, *O*-acetylserine (thiol)-lyase, had been isolated (Römer et al. 1992; Saito et al. 1992). Although several techniques to isolate novel genes have been developed (reviewed in Frommer & Ninnemann 1995), functional complementation of a genetically well described organism such as *E. coli* was a potentially effective choice here because:

- (i) similarities in sulphate assimilation pathways in *E. coli* and plants were apparent (Kredich 1987; Gisselmann et al. 1992); and
- (ii) several cysteine auxotrophic *E. coli* mutants deficient in specific sulphate assimilation enzymes had been described (Wheldrake & Pasternak 1965; Jones-Mortimer 1968, 1973).

Functional expression of eukaryotic genes in *E. coli* was first observed by Struhl et al. (1976), who were able to revert an *E. coli* histidine auxotroph by integration of a segment of *Saccharomyces cerevisiae* DNA putatively encoding imidazole glycerol phosphate dehydratase. The first plant gene isolated by functional complementation of an *E. coli* auxotroph was *meso*-diaminopimelate dehydrogenase from *Glycine max* (soybean) (Wenko et al. 1985). In that experiment, an *E. coli* mutant was infected with *G. max* genomic DNA fragments inserted into a bacteriophage  $\lambda$  vector. Colonies containing an expressed gene that complemented or bypassed the bacterial lesion were able to grow on the selection medium, and the growing cells were then lysed by the bacteriophage to form plaques. Inserts within the bacteriophage DNA from the

plaques were recloned into plasmid vectors for further analysis. A simpler method was developed by Izui et al. (1986), who infected an *E. coli* mutant defective in phosphoenolpyruvate carboxylase activity with a *Zea mays* leaf cDNA expression vector. It was demonstrated in the same year that an expression vector containing a cloned glutamine synthetase gene was able to complement the respective *E. coli* mutant (DasSarma et al. 1986). DasSarma et al. (1986) have been credited as the first researchers to identify a plant gene by complementation (Frommer & Ninnemann 1995), whereas Wenko et al. (1985) warrant this acknowledgement. Furthermore, DasSarma et al. (1986) did not use functional complementation to isolate the plant glutamine synthetase gene, but to confirm its identity.

Functional complementation of *E. coli* mutants subsequently has been utilised extensively to isolate plant genes (eg. Snustad et al. 1988; Delauney & Verma 1990; Van Camp et al. 1990; Frisch et al. 1991; Lal et al. 1991; Senecoff & Meagher 1993; Schnorr et al. 1994; Stallmeyer et al. 1995) or to confirm the identity of a cloned plant gene (eg. Smith et al. 1989; Saito et al. 1993b; Arz et al. 1994; Hell et al. 1994; Murillo & Leustek 1995). Functional complementation of *S. cerevisiae* mutants has also been used widely as a method for isolation of plant genes (eg. Minet et al. 1992; Hsu et al. 1993; Jain & Leustek 1994; Klonus et al. 1994; Smith et al. 1995). Selection of plant genes in *S. cerevisiae* mutants can be preferential to bacterial systems due to the lack of eukaryotic mRNA and protein processing mechanisms, toxicity of membrane proteins and aggregation of insoluble eukaryotic polypeptides in *E. coli* (Frommer & Ninnemann 1995).

In this study, the basic requirements for functional complementation, ie. a plant cDNA expression library, appropriate *E. coli* mutants and a suitable screen for selection of phenotypic rescue, were satisfied by the availability of an *A. thaliana* cDNA expression library in

$\lambda$ YES (Elledge et al. 1991), *E. coli* cysteine auxotrophs (Wheldrake & Pasternak 1965; Jones-Mortimer 1968, 1973) and the ability to detect complemented cells by growth on sulphate as a sole sulphur source, respectively.

#### 1.4 Aims of the investigation

Previous studies have failed to resolve the primary pathways of sulphate assimilation in higher plants. A greater understanding of these pathways will be achieved by characterisation of sulphate assimilation enzymes at the molecular level. This work describes research to isolate, by functional complementation of *E. coli* cysteine auxotrophs, and to analyse *A. thaliana* cDNAs encoding: (i) serine acetyltransferase, which catalyses an important step in cysteine biosynthesis; (ii) APS kinase, catalysing the phosphorylation of APS to PAPS; and (iii) PAPS reductase, which reduces the sulpho group of PAPS to sulphide. An application of this work was envisaged to involve specific downregulation of each of these enzymes in order to resolve key questions about the biochemical steps and regulation of plant sulphate assimilation.

## CHAPTER 2: MATERIALS AND METHODS

### 2.1 Reagents

Analytical grade biochemical and molecular biology reagents were obtained from Sigma (UK) or BDH (UK), unless indicated otherwise in the text. Restriction enzymes were purchased from Promega Corporation (UK) or NBL (UK), while DNA modifying enzymes were obtained from Promega Corporation (UK) or Boehringer Mannheim (Germany).

## 2.2 Bacterial strains and their maintenance

*Escherichia coli* cysteine auxotrophic strains utilised for functional complementation experiments and obtained from Dr B. Bachmann, *E. coli* Genetic Stock Center, Yale University, USA, were:

- (i) JM15 (*cysE50*, *tfr-8*), deficient in serine acetyltransferase (Jones-Mortimer 1968);
- (ii) JM81A (*cysC92*, *tfr-8*), deficient in APS kinase (Wheldrake & Pasternak 1965); and
- (iii) JM96 (*thr-1*, *leuB6*, *fhuA2*, *lacY1*, *supE44*, *gal-6*,  $\lambda^-$ , *trp-1*, *hisG1*, *cysH56*, *galP63*, *gltB31*, *rpsL9*, *malT1*( $\lambda^R$ ), *xyl-7*, *mtl-2*, *argH1*, *thi-1*), deficient in PAPS reductase (Jones-Mortimer 1968; 1973).

Other *E. coli* strains used were:

- (i) DH5 $\alpha$  ( $F^-$ ,  $\phi 80$ dlacZ $\Delta$ M15, *recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17* [ $r_K^-$ ,  $m_K^+$ ], *supE44*, *relA1*, *deoR*,  $\Delta$ [lacZYA-argF]U169; Hanahan 1983; Bethesda Research Laboratories 1986), purchased from Gibco-BRL, UK;
- (ii) LE392 ( $F^-$ , *hsdR574*, [ $r_K^-$ ,  $m_K^+$ ], *supE44*, *supF58*, *lacY1* or  $\Delta$ [lacIZY]6, *galk2*, *galT22*, *metB1*, *trpR55*; Borck et al. 1976; Murray et al. 1977), obtained from Dr A. Sherman, University of St Andrews, UK;
- (iii) TB1 (*ara*,  $\Delta$ [lac, proAB], *rpsL*, [ $\phi 80$  lacZ $\Delta$ M15], *hsdR*; Johnston et al. 1986), obtained from Dr J. Sommerville, University of St Andrews, UK;
- (iv) BNN132 ( $\lambda$ KC, *endA1*, *thi*, *gyrA96*, *hsdR17* [ $r_K^-$ ,  $m_K^+$ ], *relA1*, *supE44*,  $\Delta$ [lac-proAB], [ $F'$ , *traD36*, *proAB^+*, *lacI<sup>q</sup>Z $\Delta$ M15*]; Elledge et al. 1991), acquired from Dr J. Mulligan, Stanford University School of Medicine, USA; and
- (v) XL1-blue (*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, *relA1*, lac[ $F'$ , proAB, *lacI<sup>q</sup>Z $\Delta$ M15*Tn10(*tet<sup>r</sup>*)]); Bullock et al. 1987), purchased from Stratagene, UK.

Frequently used microbiological media (Sambrook et al. 1989), sterilised by autoclaving at 120°C for 15 - 20 min except where indicated, were:

- (i) Luria-Bertani (LB) broth: 10 % (w/v) tryptone, 5 % (w/v) yeast extract, 10 % (w/v) NaCl, pH 7.4;
- (ii) LB plates: LB broth with 1.5 % (w/v) agar;
- (iii) M9 minimal medium: 100 ml M9 minimal medium was prepared by mixing 20 ml 5 x M9 salts (5 x M9 salts consists of 211 mM Na<sub>2</sub>HPO<sub>4</sub>, 110 mM KH<sub>2</sub>PO<sub>4</sub>, 43 mM NaCl and 93 mM NH<sub>4</sub>Cl, pH 7.4), 1 ml 20 % (w/v) glucose (filter-sterilised) or 2 ml 10 % (w/v) mannitol (filter-sterilised), 0.2 ml 1 M MgSO<sub>4</sub>, 0.01 ml 1 M CaCl<sub>2</sub>, and appropriate amino acids and supplements (filter-sterilised), taken to 100 ml with sterile water; and
- (iv) M9 minimal medium plates: M9 minimal medium with 1.5 % (w/v) agar.

Bacterial strains were stored in 10 - 15 % (v/v) glycerol/LB at -80°C. When required, glycerol stock cultures were streaked onto appropriate plates and incubated at 37°C overnight. For growth on M9 minimal medium, the *E. coli* sulphate auxotrophic strains required the following amino acid supplements (final concentration): JM81A, 0.05 - 1 mM cysteine; JM15, 0.05 - 1 mM cystine; and JM96, 0.05 - 1 mM cysteine, 40 mg/ml threonine, 40 mg/ml leucine, 40 mg/ml tryptophan, 40 mg/ml histidine, 40 mg/ml arginine and 4 mg/ml thiamine.

### 2.3 Plant material

*Arabidopsis thaliana* (Columbia ecotype) seeds were obtained from the Nottingham *Arabidopsis* Stock Centre, Nottingham University, UK. Seeds were germinated on wet compost in trays and grown in greenhouse conditions with natural lighting supplemented with a 16/8 light/dark regime. Germinating seedlings were sensitive to waterlogged compost, yet desiccation intolerant, and were therefore watered with a fine mist spray. Plant tissue was usually harvested after bolting, and seeds were collected when siliques were dried completely. Seeds were stored in sealed tubes at 4°C - under these conditions the seeds were found to remain viable for at least two years.

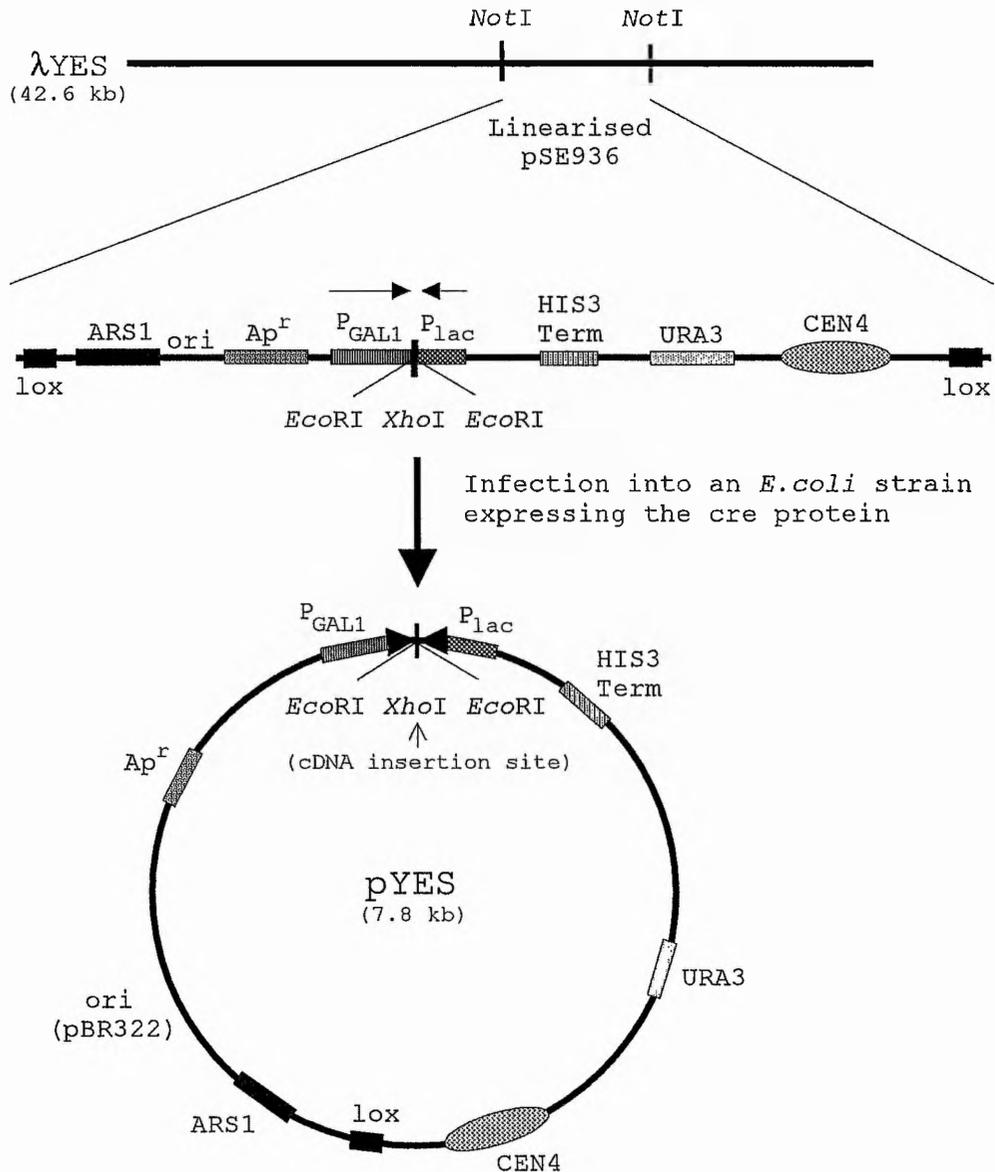
Spinach (*Spinacia oleracea* var. Medania) and oilseed rape (*Brassica napus*, cultivars COMET [Spring Double low], FALCON [Winter Double low] and PASHA [Winter Single low]) plants were obtained from Dr H. Ross, Scottish Crop Research Institute, Invergowrie, UK. The plants were kept on compost in greenhouse conditions until leaf tissue was harvested.

## 2.4 The $\lambda$ YES *Arabidopsis thaliana* cDNA expression vector for functional complementation

### 2.4.1 Description of the $\lambda$ YES cDNA library

An aliquot of the *A. thaliana* (Columbia ecotype) cDNA library constructed in the  $\lambda$ YES yeast-*E. coli* expression vector system (Elledge et al. 1991) was obtained from Dr J. Mulligan, Stanford University School of Medicine, USA. The cDNAs, prepared from polyadenylated [poly(A)] RNA isolated from above-ground parts of *A. thaliana* plants which varied in growth stage from those with newly opened primary leaves to those which were flowering, were inserted non-directionally into  $\lambda$ YES at a *Xho*I site flanked by two *Eco*RI sites. The original library of  $10^7$  independent recombinants was amplified once, and inserts are contained within 90 - 95 % of the resultant clones (Elledge et al. 1991). Aliquots of the  $\lambda$ YES library, maintained in LB with 7 % (v/v) dimethyl sulfoxide (DMSO), were stored at 4°C for frequent use, while long-term storage was at -80°C.

The  $\lambda$ YES expression vector was designed to facilitate cloning of *A. thaliana* genes by functional complementation of *E. coli* or *Saccharomyces cerevisiae* mutants (Elledge et al. 1991). When  $\lambda$ YES is infected into host bacterial cells containing  $\lambda$ KC, which expresses the *cre* gene, the phage form of the vector is converted into a plasmid by *cre-lox* site-specific recombination (Figure 2.1) (Sauer & Henderson 1988), at a reported efficiency rate of 67 % (Schnorr et al. 1994). The plasmid portion (pSE936) of the  $\lambda$ YES vector includes an *A. thaliana* cDNA insert and a  $\beta$ -lactamase gene which confers resistance to ampicillin, and is flanked by direct repeats of *lox* sites (Figure 2.1). In *E. coli*, the cDNA inserts are expressed from an *E. coli lac* promoter, and expression is induced by isopropyl  $\beta$ -D-thiogalactoside (IPTG). The  $\lambda$ YES expression system provides an efficient method for isolating *A. thaliana* genes by complementation because it combines the



**Figure 2.1** The λYES system for cre-lox-mediated automatic subcloning of phage λ to plasmid clones

The λYES phage (top) contains a linearised plasmid pSE936 inserted at NotI sites (middle). The pSE936 insert can be converted automatically upon infection into *Escherichia coli* strains lysogenic for λKC (expressing the cre protein) into its plasmidic form, pYES, by site-specific recombination (bottom). The pYES plasmid can replicate in both *E. coli* and *Saccharomyces cerevisiae*. A λYES library with *Arabidopsis thaliana* cDNAs inserted non-directionally into a XhoI site and flanked by two EcoRI sites in the pYES portion of the vector was employed in this study. In *E. coli*, the cDNA inserts are expressed from the *E. coli lac* promoter (P<sub>lac</sub>), which contains a ribosome binding site and ATG start codon before the cDNA insertion site. The β-lactamase gene (Ap<sup>r</sup>) of the pYES plasmid confers ampicillin resistance to its bacterial host. This figure was adapted from Fig. 2 in Elledge et al. (1991).

effectiveness of phage  $\lambda$  infection of *E. coli* with the ease of manipulation of ampicillin-resistant plasmid clones that are formed automatically by site-specific recombination.

#### 2.4.2 Titrating $\lambda$ phages

The expression library  $\lambda$ YES and helper phage  $\lambda$ KC were titred according to Davis et al. (1986).

##### *Preparation of plating cells*

A fresh colony of *E. coli* strain LE392 was inoculated into 100 ml of LB medium containing 2 ml of maltose (maltose increases the number of  $\lambda$  phage receptors, the lamB protein, which are required for maltose transport) and grown overnight at 37°C with shaking. Cells were harvested by centrifugation at 2500 x g and resuspended in 50 ml of sterile 10 mM MgSO<sub>4</sub>. Cells were stored at 4°C for up to two weeks.

##### *Plating phage $\lambda$*

After appropriate dilution in SM buffer (50 mM Tris-HCl, pH 7.4, 10 mM MgSO<sub>4</sub>, 0.01 % [w/v] gelatin and 100 mM NaCl), to yield an anticipated 500 plaques per 90 mm culture dish when approximate titres were known, 1  $\mu$ l and 10  $\mu$ l of diluted phage were dispensed into heat-sterilised 12 x 75 mm glass tubes. 200  $\mu$ l of prepared strain LE392 bacteria was added to each tube and incubated for 20 min at room temperature to allow  $\lambda$  phage adsorption. Then 2.5 ml of sterile LB top agar (LB broth with 0.7 % [w/v] agar, supplemented after autoclaving and cooling to 48°C with MgSO<sub>4</sub> to a final concentration of 10 mM) at 48°C was added to each tube, mixed, poured immediately onto a 90 mm LB agar plate prewarmed to 37°C, and swirled to distribute top agar evenly. Top agar was allowed to set at room temperature for 10 min, and the plates left inverted at 37°C until plaques were

visible (usually within 16 h). The titre of each phage was calculated as plaque-forming units (pfu)/ml. Control plates lacking phage were used to determine background plaque formation.

#### 2.4.3 Induction of $\lambda$ KC in *Escherichia coli* strain BNN132

Attempts to induce the *E. coli*  $\lambda$ KC lysogenic strain BNN132 using mitomycin C treatment (Otsuji et al. 1959; Korn & Weissbach 1962; Castellazzi et al. 1972b; Borek & Ryan 1973) were unsuccessful. Briefly, 5 ml of a 37°C overnight culture of strain BNN132 in LB supplemented with 40  $\mu$ g/ml kanamycin and 0.4 % (w/v) maltose was subcultured into 100 ml of fresh medium and grown to logarithmic phase (OD<sub>660</sub> of 0.6). To 20 ml aliquots of the cells, 0, 1, 5 and 10  $\mu$ g/ml (final concentration) of mitomycin C was added, and the cultures incubated further at 37°C in darkness. Decrease in turbidity, an indication of lytic induction, was monitored by visual inspection for up to 18 hours. Confirmation of the absence of lysis was obtained by titring of phage contained in the supernatant (Section 2.4.2).

An ultraviolet (UV) irradiation technique based on Castellazzi et al. (1972b) proved effective, and the optimal method (results in Chapter 3.2.1) is described. 4 ml of a 37°C overnight culture of *E. coli* strain BNN132 in LB supplemented with 40  $\mu$ g/ml kanamycin was subcultured into 100 ml of fresh medium and grown to an OD<sub>600</sub> of 1.0. The bacteria were pelleted at 2500 x g and resuspended in 50 ml of 10 mM MgSO<sub>4</sub>. Cells (6 ml/90 mm culture dish) on a horizontal shaker were exposed to UV from a Hanovia Bactericidal UV Unit at 60 cm for 1 min. After addition of 667  $\mu$ l of 10 x LB (10 % [w/v] tryptone, 5 % [w/v] yeast extract, 10 % [w/v] NaCl, pH 7.4) to the culture dish, the cells were incubated for 3 h without shaking at 37°C to allow for lysis. Debris was pelleted at 3500 x g for 10 min, 4°C. Supernatants were stored at 4°C

after the addition of a few drops of chloroform to stabilise the phage (Arber et al. 1983). Phage was titred as described in Section 2.4.2.

#### 2.4.4 Creating $\lambda$ KC lysogens of *Escherichia coli* cysteine auxotrophs

Prior to infection of the *E. coli* cysteine auxotrophic strains JM15 (*cysE*), JM81A (*cysC*) and JM96 (*cysH*) with the  $\lambda$ YES library, these host strains were lysogenised with the *cre*-expressing helper phage  $\lambda$ KC (induced from strain BNN132; Section 2.4.3) using a method derived from Kourilsky (1973) and Lech and Brent (1988). A 37°C shaking overnight culture grown in 100 ml lambda broth (1 % [w/v] tryptone, 0.25 % [w/v] NaCl, supplemented after autoclaving and cooling below 50°C with a final concentration of 0.4 % [w/v] maltose and 10 mM MgSO<sub>4</sub>) was centrifuged at 2500 x g, and the pellet resuspended in 25 ml 10 mM MgSO<sub>4</sub>. Cells were starved by incubation in MgSO<sub>4</sub> for a further 40 min (approximately 1 growth cycle) at 37°C with shaking to stimulate lysogenisation (Fry 1963). 0.3 - 0.4 ml aliquots of the cells were dispensed into heat-sterilised 12 x 75 mm glass tubes, mixed with 100  $\mu$ l of  $\lambda$ KC (diluted serially in SM medium to provide a phage input range of approximately 10<sup>4</sup> - 10<sup>8</sup> pfu/ml), and incubated at room temperature for 20 min to allow phage adsorption. The tubes were transferred to a 37°C water bath for 10 min to allow phage to inject their DNA into the host cells, and then 2.5 ml of lambda top agar (1 % [w/v] tryptone, 0.25 % [w/v] NaCl, 0.7 % [w/v] agar, supplemented with 40  $\mu$ g/ml kanamycin and maintained at 50°C) was added to each tube, mixed and poured onto lambda agar plates (1 % [w/v] tryptone, 0.25 % [w/v] NaCl, 1.0 % [w/v] agar, supplemented with 40  $\mu$ g/ml kanamycin). The plates were swirled to distribute the lambda top agar evenly, allowed to set at room temperature for 10 min, and placed inverted in a 37°C incubator. A selection of colonies that grew overnight

(putative lysogens) were replated onto LB plates supplemented with 40  $\mu\text{g/ml}$  kanamycin to confirm kanamycin resistance. These lysogens were plated onto M9 minimal medium containing sulphate as sole sulphur source to verify that they were still auxotrophic for cysteine.

#### 2.4.5 Functional complementation of *Escherichia coli* cysteine auxotrophic strains with the $\lambda\text{YES}$ cDNA library

The *E. coli* cysteine auxotrophic lysogens JM15/ $\lambda\text{KC}$  (*cysE*), JM81A/ $\lambda\text{KC}$  (*cysC*) and JM96/ $\lambda\text{KC}$  (*cysH*) were infected with the  $\lambda\text{YES}$  phage library essentially as described (Elledge et al. 1991). Host cells were grown overnight in a shaker incubator at 37°C in 50 ml LB supplemented with final concentrations of 0.4 % ( $\text{w/v}$ ) maltose and 10 mM IPTG. After centrifugation at 2500 x g, cells were resuspended in 25 ml 10 mM  $\text{MgSO}_4$ . Aliquots of 1 ml (approximately  $10^9$  cells) of the mutant host were incubated with  $6 \times 10^7$  to  $1 \times 10^8$   $\lambda\text{YES}$  phage for 30 min at 30°C without shaking to allow phage adsorption. Control experiments with cultures lacking  $\lambda\text{YES}$  were performed in parallel. 1.5 ml of M9 minimal medium containing 0.2 % ( $\text{w/v}$ ) mannitol, 1 mM IPTG and 0.5 - 1.0 mM of each required amino acid (including cysteine or cystine) was added and the culture was incubated for 2 h at 30°C with shaking to allow the cyclisation of pYES and expression of the ampicillin-resistance gene  $\beta$ -lactamase. Each culture was divided between two microcentrifuge tubes, cells were collected by microcentrifugation (6500 rpm, 10 min, room temperature), washed once with 1 ml sterile water to remove complete medium, resuspended in 0.49 ml 10 mM  $\text{MgSO}_4$ , and combined. Cells were plated with a glass spreader (100  $\mu\text{l}$  of cells/90 mm plate) on selective medium comprising M9 minimal medium, 0.2 % ( $\text{w/v}$ ) mannitol, 1 mM IPTG, 50  $\mu\text{g/ml}$  ampicillin and required amino acids except cysteine or cystine. Plates were covered in plastic wrap to prevent desiccation, inverted and incubated at 37°C for 3 - 4 days. Individual colonies that

were able to grow after being replated onto fresh selection medium were designated as complemented clones. For storage, individual complemented clones were grown overnight in screw-cap microcentrifuge tubes shaking at 37°C in LB broth supplemented with 10 % (v/v) glycerol and then frozen at -80°C.

#### 2.4.6 Production of empty vector pYES from clone pSAT1

An empty pYES plasmid without a cDNA insert, required as a negative control vector, was constructed from the plasmid clone pSAT1 (Chapter 3.2.1). Purified pSAT1 (4µg) was digested with 12 units of *EcoRI* at 37°C for 1 h, then fractionated by electrophoresis through a 0.8 % (w/v) agarose gel. An 8 kb gel fragment, the expected size of linearised pYES (Elledge et al. 1991), was excised and the DNA purified by QIAEX extraction (Section 2.6.3). Linearised pYES DNA (100 ng) was religated using T4 DNA ligase (0.1 Weiss unit) at 15°C for 17.5 h. Religated pYES was transformed into *E. coli* strain DH5α (Section 2.5.3), and ampicillin-resistant transformants were checked for the presence of a plasmid using the "cracking procedure" (Section 2.5.4). Two strain DH5α colonies containing the empty pYES plasmid were chosen for miniprep isolation of the plasmid (Section 2.5.2), and identity of the empty pYES plasmids was verified (results not shown) by digestion with the restriction enzymes *KpnI*, *SalI*, *SmaI* and *XbaI*, which have unique restriction sites within pYES, and *NotI*, which does not cut the *cre-lox* automatically subcloned pYES (Elledge et al. 1991).

## 2.5 Plasmid DNA

### 2.5.1 Plasmids employed

In addition to pYES plasmids derived from  $\lambda$ YES, the following plasmids were utilised:

(i) the cloning vector pGEM-7Zf(+) (Promega Corporation, UK) was used for subcloning DNA fragments. The plasmid was transformed into *E. coli* strain XL1-Blue for amplification and maintenance;

(ii) the pGEX-4T Fusion Vectors (Pharmacia Biotech, UK; obtained from Dr J. Sommerville, University of St Andrews, UK) were used to express and purify glutathione S-transferase (GST) fusion proteins. Plasmids pGEX-4T-1, pGEX-4T-2 and pGEX-4T-3, which have the multiple cloning site in three forward reading frames, were maintained in *E. coli* strain TB1; and

(iii) plasmid pZL1 containing the *A. thaliana* Expressed Sequence Tag (EST) clone 89K20T7, in the *E. coli* host strain DH10B, was obtained from the Arabidopsis Biological Resource Center, Ohio State University, USA.

### 2.5.2 Preparation of plasmid DNA

Plasmid DNA was isolated from *E. coli* host cells using five methods which differed in cost, speed of isolation, and purity and quantity of plasmid obtained. Plasmid DNA was stored at  $-20^{\circ}\text{C}$ .

(i) QIAprep-spin plasmid miniprep kits (Hybaid, UK) were used to extract rapidly up to 20  $\mu\text{g}$  plasmid of sequencing purity. An overnight host *E. coli* culture (5 ml) grown in LB broth with appropriate antibiotics at  $37^{\circ}\text{C}$  was pelleted by centrifugation at 6000 x g for 15 min at  $4^{\circ}\text{C}$ . The pellet was resuspended in 250  $\mu\text{l}$  buffer P1 (containing RNaseA), lysed in

250  $\mu$ l sodium dodecyl sulphate (SDS)-alkaline buffer P2 for 5 min at room temperature and then cellular debris and genomic DNA were precipitated by adding 350  $\mu$ l chilled buffer N3 and incubating on ice for 5 min. The solution was microcentrifuged for 10 min at 13000 rpm, 4°C, and the supernatant applied to a QIAprep-spin column. The column was microcentrifuged at room temperature for 30 s, washed with buffer PB to remove trace nucleases, and then washed with buffer PE. Plasmid DNA was eluted with 100  $\mu$ l 10 mM Tris-HCl.

(ii) Midi plasmid purification kits (Hybaid, UK) enabled the rapid extraction of up to 100  $\mu$ g plasmid of sequencing purity. An overnight culture of host *E. coli* (30 ml for hosts containing high copy number plasmids such as pBluescript, pGEM and pUC19; 150 ml for cells with low copy number plasmids such as pBR322, pYES) grown at 37°C in LB broth with appropriate antibiotics was pelleted by centrifugation at 6000 x g for 15 min, 4°C. The pellet was resuspended thoroughly in 4 ml buffer P1, and cells lysed with the addition of 4 ml buffer P2 and incubated at room temperature for 5 min. Cellular debris and genomic DNA were precipitated by adding chilled buffer P3 and incubating for 15 min on ice. The solution was centrifuged at 30000 x g for 30 min, 4°C. If the supernatant was not clear, a second centrifugation for 15 min as previously was carried out. The cleared supernatant was applied to a QIAGEN-tip 100 column, previously equilibrated with 4 ml buffer QBT, and allowed to empty by gravity flow. The column was washed twice with 10 ml buffer QC to remove contaminants, and the DNA eluted with 5 ml buffer QF. DNA was precipitated with 0.7 volumes isopropanol and the solution centrifuged at 15000 x g for 35 min, 4°C. The DNA pellet was washed once with 5 ml cold 70 % ( $v/v$ ) ethanol, air-dried for 10 min, and redissolved in 200  $\mu$ l TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).

(iii) Magic Minipreps DNA purification system (Promega Corporation, UK) was used to isolate rapidly up to 10  $\mu$ g

plasmid DNA for restriction analysis. An overnight culture of *E. coli* host cells (1.3 ml) grown in LB broth with appropriate antibiotics at 37°C was microcentrifuged for 5 min at 6500 rpm, 4°C, and the pellet resuspended in 200 µl Cell Resuspension Solution. To this was added 200 µl Cell Lysis Solution and the mixture inverted until it became clear. The cleared suspension was mixed with 200 µl Neutralization Solution and the mixture microcentrifuged at 13000 rpm for 5 min, 4°C. 1 ml of DNA Purification Resin was mixed with the cleared supernatant, applied to a Magic Minicolumn, and forced into the column using a luer-lock syringe. The column was washed with 2 ml of Column Wash Solution, dried by microcentrifugation for 20 s at 13000 rpm, and the plasmid DNA eluted with 50 µl TE buffer.

(iv) A hot alkaline miniprep method (Musich & Chu 1993) was used to extract up to 20 µg crude plasmid DNA for restriction analysis. *E. coli* host cells were grown overnight at 37°C in 5 ml of Terrific Broth (1.2 % [w/v] tryptone, 2.4 % [w/v] yeast extract, 0.4 % [v/v] glycerol; 17 mM KH<sub>2</sub>PO<sub>4</sub> and 72 mM K<sub>2</sub>HPO<sub>4</sub> [final concentrations] were added after autoclaving and cooling to below 50°C) with appropriate antibiotics, then centrifuged at 1000 x g for 10 min, room temperature. The pellet was vortexed until creamy before addition of 1.3 ml SET buffer (20 % [w/v] sucrose, 50 mM EDTA, 50 mM Tris-HCl, pH 8.0), and mixed further by vortexing. The cells were collected by microcentrifugation at 13000 rpm for 20 s, the soft pellet vortexed until creamy, then resuspended in 250 µl SET buffer. 0.5 ml of freshly made 0.2 N NaOH, 1 % (w/v) SDS was added, the solution mixed by gentle inversion and incubated at 65°C for 30 min to lyse cells, denature chromosomal DNA, solubilise cellular material and degrade RNA. At room temperature, 375 µl of 5 M potassium acetate (3 M in potassium, 5 M in acetate [using glacial acetic acid], pH 4.8) was added and mixed, and the solution placed on ice for 15 min. Debris was collected by microcentrifugation at 13000 rpm for 15 min, room temperature, and the supernatant

mixed with 300  $\mu$ l of 27 % (w/v) polyethylene glycol (PEG) in 3.3 M NaCl and placed on ice for 0.5 - 2 h to selectively precipitate DNA. DNA was collected by microcentrifugation at 13000 rpm for 15 min, room temperature, and washed with 70 % (v/v) ethanol to remove excess PEG and salt. The DNA pellet was air-dried for 5 min, resuspended in 400  $\mu$ l 10 mM Tris-HCl, pH 8.4, 1 mM EDTA, and reprecipitated with 400  $\mu$ l 5 M ammonium acetate and 800  $\mu$ l isopropanol. The DNA pellet was dissolved finally in 50  $\mu$ l TE buffer.

(v) The *E. coli* "cracking" procedure (Barnes 1977) was used to estimate plasmid size rapidly from individual host colonies. Single *E. coli* colonies grown overnight at 37°C on LB plates with appropriate antibiotics were picked using sterile toothpicks and smeared into 50  $\mu$ l 10 mM EDTA, pH 8.0, contained in 1.5 ml microcentrifuge tubes. 50  $\mu$ l of 2 x cracking buffer (0.2 M NaOH, 0.5 % [w/v] SDS, 20 % [w/v] sucrose) was added, the cells resuspended by vortexing, and the solution incubated at 70°C for 5 min. Samples were allowed to cool to room temperature for 10 min, then 1.5  $\mu$ l 4 M KCl and 1  $\mu$ l 0.4 % (w/v) bromophenol blue were added and mixed by vortexing. The samples were placed on ice for 5 min, then microcentrifuged for 3 min at 13000 rpm, 4°C. Up to 50  $\mu$ l of the supernatant was run on 0.7 - 0.8 % (w/v) agarose gels (Section 2.5.6). Control plasmids without DNA inserts were prepared if colonies were being analysed for presence of plasmids with ligated DNA inserts.

### 2.5.3 Transformation of *Escherichia coli* with plasmid DNA

Three protocols were employed to prepare and transform competent *E. coli* cells: (i) a basic calcium chloride method (Seidman et al. 1989), (ii) a higher efficiency calcium chloride method (Sambrook et al. 1989), and (iii) a one-step preparation and transformation method (Chung et al. 1989).

(i) Basic calcium chloride method (Seidman et al. 1989): An overnight culture of *E. coli* cells grown at 37°C in LB medium was inoculated 1:100 into fresh medium and the cells grown at 37°C to an OD<sub>590</sub> of 0.375 (mid-logarithmic phase). The culture was decanted into appropriate centrifuge tubes, left on ice for 5 min, then centrifuged at 1600 x g for 7 min, 4°C, and allowed to decelerate without brake. Pellets were resuspended carefully in 5 ml cold CaCl<sub>2</sub> solution (60 mM CaCl<sub>2</sub>, 15 % [ $\frac{v}{v}$ ] glycerol, 10 mM piperazine-N,N'-bis[2-ethanesulfonic acid], pH 7.0) per 25 ml of original culture, and centrifuged at 1100 x g for 5 min, 4°C. Cells were resuspended in 5 ml cold CaCl<sub>2</sub> per 25 ml of original culture, kept on ice for 30 min, then centrifuged at 1100 x g for 5 min, 4°C. The pellets were resuspended finally in 1 ml CaCl<sub>2</sub> solution per 25 ml of original culture, and the competent cells used immediately for transformation or stored in 250  $\mu$ l aliquots at -80°C. For transformation, a 10-25  $\mu$ l aliquot of plasmid DNA (and a control sample containing no DNA) was mixed with 100  $\mu$ l competent cells and the mixture placed on ice for 10 min to allow DNA binding. The cells were heat-shocked for 2 min at 42°C to permit entry of plasmid DNA, 0.9 ml of LB medium was added to the cells and they were shaken at 225 rpm for 1 h, 37°C, to allow expression of the plasmid  $\beta$ -lactamase gene for resistance to ampicillin. Aliquots of cells (usually 100  $\mu$ l) were then spread onto fresh LB plates containing 50 - 100  $\mu$ g/ml ampicillin using a glass spreader and left overnight at 37°C to select transformed colonies.

(ii) High efficiency calcium chloride method (Sambrook et al. 1989): *E. coli* cells were grown to mid-logarithmic phase as described for the basic calcium chloride method, incubated on ice for 10 min, then centrifuged at 2700 x g for 10 min, 4°C. Pellets were drained, resuspended in 5 ml of cold 0.1 M CaCl<sub>2</sub> per 25 ml of original culture and stored on ice for 15 min. Cells were recovered by centrifugation at 2700 x g for 10 min, 4°C, and drained pellets resuspended in 1 ml cold 0.1 M CaCl<sub>2</sub> per 25 ml of original culture. These competent cells

were either used fresh or stored at  $-80^{\circ}\text{C}$  in  $250\ \mu\text{l}$  aliquots. For transformation, DNA in a volume of up to  $10\ \mu\text{l}$  was mixed with  $200\ \mu\text{l}$  competent cells and the suspension stored on ice for 30 min. After a heat-shock treatment at  $42^{\circ}\text{C}$  for 90 s,  $800\ \mu\text{l}$  of SOC medium (prepared by dissolving 2 % [w/v] tryptone, 0.5 % [w/v] yeast extract, 0.05 % [w/v] NaCl and 2.5 mM KCl, adjusting pH to 7.0, autoclaving and when below  $50^{\circ}\text{C}$  adding sterile solutions of glucose [to 20 mM] and  $\text{MgCl}_2$  [to 10 mM]) was added to the cells and the cultures shaken at 225 rpm for 45 min,  $37^{\circ}\text{C}$ . Transformed cells were selected as described for the basic calcium chloride method.

(iii) One-step preparation and transformation method (Chung et al. 1989): *E. coli* cells were grown to early exponential phase ( $\text{OD}_{600}$  of 0.3 - 0.4) as described for the basic calcium chloride method, then diluted with 1 volume of cold 2 x transformation and storage solution (2 x TSS contains 2 % [w/v] tryptone, 1 % [w/v] yeast extract, 2 % [w/v] NaCl, 20 % [w/v] PEG 6000, 100 mM  $\text{MgCl}_2$ , 10 % [v/v] DMSO, pH 6.5). These competent cells were stored at  $-80^{\circ}\text{C}$  or used immediately for transformation.  $100\ \mu\text{l}$  cells were mixed with plasmid DNA in a volume of 1 -  $5\ \mu\text{l}$  and incubated for 5 - 60 min on ice (time of incubation on ice is not a critical parameter in this method). 0.9 ml of LB supplemented with either 20 mM glucose or 0.2 % (w/v) mannitol was added to the cells, and the culture was grown at  $37^{\circ}\text{C}$  for 1 h with shaking (225 rpm). Transformants were selected as described for the basic calcium chloride protocol.

#### 2.5.4 Restriction enzyme digestion of DNA

DNA samples were restricted with restriction enzymes in small volumes (10 -  $20\ \mu\text{l}$ ) containing appropriate 1 x restriction enzyme buffers. Star activity (cleavage of DNA at sequences other than recognition sites) was avoided by using restriction enzymes at approximately 0.5 - 2 units/ $\mu\text{g}$  DNA.

Digests were incubated at the recommended temperature (normally 37°C) for 1 - 16 h, and were terminated by the addition of DNA sample loading buffer (Section 2.5.5) or by freezing at -20°C.

#### 2.5.5 Agarose gel electrophoresis of DNA fragments

DNA fragments were fractionated on 0.7 - 2.0 % (w/v) agarose gels, depending on the separation range required (Sambrook et al. 1989). Agarose was melted in 1 x TAE buffer (a stock of 50 x TAE consists of 2 M Tris-HCl, pH 8.0, 1 M acetic acid and 50 mM EDTA), cooled to 50°C and poured into a gel casting tray. The solidified gel was placed into an electrophoresis tank and immersed in 1 x TAE running buffer. DNA samples and 0.5 µg of a Sigma λ/HindIII molecular weight marker were mixed with 0.1 volumes of 10 x DNA sample loading buffer (0.25 % [w/v] bromophenol blue, 1 x TAE, 50 % [v/v] glycerol), and loaded into the wells. Where separation of low molecular weight DNA fragments (50 - 500 bp) was required, 1 µg of pGEM DNA markers (Promega Corporation, UK) and 6 x orange/blue DNA loading buffer (Promega Corporation, UK) were used. Voltage applied and duration of electrophoresis was determined by gel size and agarose concentration. Typically, a 120 ml 1 % (w/v) agarose gel in a midi-apparatus was subjected to electrophoresis at 60 V for 2 h or 22 V overnight. After staining the gel in a 0.5 µg/ml solution of ethidium bromide (a 10 mg/ml stock was stored at 4°C) for 20 min and destaining in distilled water for 20 min, separated DNA fragments were visualised by ethidium bromide fluorescence on a Chromato-Vue UV Transilluminator (UVP, USA) and photographed with a Polaroid MP-4 Camera (Copal, Japan) using 667 Black and White Film (Polaroid, UK).

## 2.6 Isolation and analysis of plant genomic DNA

### 2.6.1 Extraction of plant genomic DNA

A small scale cetyltrimethylammonium bromide (CTAB) method modified by Dean et al. (1992) from J. Keller (DNAP, Oakland, California, USA) was used for the isolation of genomic DNA from *Arabidopsis thaliana*, *Brassica napus* and *Spinacia oleracea*. Plant tissue was harvested, surface-sterilised with 96 % (v/v) ethanol, rinsed with distilled water, weighed, frozen in liquid nitrogen and stored at -80°C. 1 - 2 g (fresh weight) of frozen material was ground to a fine powder in liquid nitrogen. The powder was transferred to a pre-cooled 30 ml glass Corex tube containing 10 ml of CTAB extraction buffer (140 mM sorbitol, 220 mM Tris-HCl, pH 8.0, 22 mM EDTA, 800 mM NaCl, 1 % [w/v] N-lauroylsarcosine, 0.8 % [w/v] CTAB) and mixed by gentle shaking. The tube was incubated at 65°C for 20 min, with occasional shaking, and then 4 ml of chloroform/isoamyl alcohol (IAA) (24:1 [v/v]) was added and the tube was shaken at room temperature for 20 min. After centrifugation at 2800 x g for 5 min (15°C) to resolve phases, the aqueous phase was transferred to a clean tube. 6.8 ml of ice-cold isopropanol was added, and the tube was placed on ice for 10 min or left overnight at 4°C to allow precipitation of nucleic acids. Following centrifugation at 2800 x g for 15 min (4°C) to collect precipitate, the tube was drained carefully. 1.6 ml of TE buffer was added to the tube, the pellet dissolved by gentle swirling, mixed with 1.6 ml of 4M LiCl and incubated on ice for 20 min to precipitate high molecular weight RNA. The tube was centrifuged at 2800 x g for 10 min (4°C), the supernatant removed and DNA precipitated by the addition of 2 volumes of ethanol (pre-cooled to -20°C) and incubation on ice for 20 min. Precipitated DNA was collected by centrifugation at 2000 x g for 15 min (4°C), and the DNA pellet resuspended in 0.6 ml TE buffer. Contaminating RNA was removed by an RNase treatment (15 µl of a 0.5 mg/ml stock, to give a working concentration

of 12.5  $\mu\text{g/ml}$ ) for 1 h at 37°C. Proteins were removed from the DNA solution by phenol/chloroform and chloroform extraction: 66  $\mu\text{l}$  3M sodium acetate (pH 5.2) and 0.5 ml phenol (equilibrated with Tris-HCl to pH 8.0)/chloroform/IAA (25:24:1 volume ratio) were mixed with the sample by several inversions, then subjected to centrifugation with a benchtop microcentrifuge at 13000 rpm for 15 s to separate phases. The aqueous phase was transferred to a fresh tube, mixed with 0.6 ml chloroform/IAA (24:1 [ $\text{v/v}$ ]), and microcentrifuged as before. In a final precipitation step, the aqueous phase was mixed with 2 volumes of ethanol (pre-cooled to -20°C), left on ice for 20 min, then centrifuged in a microcentrifuge for 15 min at 13000 rpm, 4°C. The pellet was washed once in 70 % ( $\text{v/v}$ ) ethanol, repelleted by microcentrifugation for 5 min, and air-dried for 10 - 15 min. The DNA pellet was redissolved in 200  $\mu\text{l}$  TE, and quantified by comparison with known amounts of  $\lambda$ /*Hind*III marker DNA fractionated on a 0.8 % ( $\text{w/v}$ ) agarose gel (Section 2.5.5). Yields of 50 - 80  $\mu\text{g}$  DNA/g fw tissue obtained from *A. thaliana* compared favourably with the expected yield of >10  $\mu\text{g/g}$  fw (Dean et al. 1992).

*Pachyrizus tuberosus* genomic DNA which had been extracted using a caesium chloride method was a gift from Dr A. Gillies, University of St Andrews, UK.

### 2.6.2 Restriction enzyme digestion of plant genomic DNA

Genomic DNA extracted from *A. thaliana* (up to 6  $\mu\text{g}$ ) was digested with restriction enzymes overnight at 37°C (or 60°C for *Bst*EII) in a total volume of 400  $\mu\text{l}$  containing 1 x restriction enzyme buffer, 60 units of restriction enzyme (ie. 10- to 12-fold excess) and 50  $\mu\text{g/ml}$  bovine serum albumin (BSA). Restricted DNA was precipitated with 0.3 M sodium acetate (pH 5.2) and 2 volumes ethanol (pre-cooled to -20°C), incubated at -20°C for at least 1 h, pelleted at 13000 rpm in a microcentrifuge for 10 min, 4°C, and resuspended in 20  $\mu\text{l}$

TE buffer. Genomic DNA (7.5  $\mu\text{g}$ ) extracted from *B. napus*, *S. oleracea* and *P. tuberosus* was restricted overnight at 37°C (or 60°C for *BstEII*) in a total volume of 61.2  $\mu\text{l}$  containing 1 x restriction enzyme buffer and 30 - 60 units of restriction enzyme (ie. 4- to 8-fold excess), and stored at -20°C until use. Restricted genomic DNA was fractionated on 0.8 % (w/v) agarose gels (Section 2.5.5) for Southern blotting (Southern 1975).

### 2.6.3 Southern analysis of DNA

#### *Agarose gel electrophoresis*

Digested genomic DNA was fractionated on 0.8 % (w/v) agarose gels, as described (Section 2.5.5). The lower left corner of the gel was nicked prior to photography for subsequent orientation and alignment of the Southern transfer membrane.

#### *Southern transfer*

The photographed gel was washed in denaturation buffer (1.5 M NaCl, 0.5 M NaOH) for 30 min to denature DNA fragments, rinsed briefly in distilled water, and then soaked in neutralisation buffer (1.5 M NaCl, 1 mM EDTA, 0.5 M Tris) for 30 min (Sambrook et al. 1989). A Southern blot was assembled in a tray containing 10 x SSC transfer buffer (a 20 x SSC stock contains 3 M NaCl, 0.3 M trisodium citrate) on a plastic platform overlaid with a wick of 3 MM filter paper (Southern 1975). The gel, placed inverted onto the 3 MM paper, was covered with Hybond-N nylon membrane that had been trimmed exactly to the size of the gel and prewet first in sterile water and then in 10 x SSC. Air pockets between transfer layers were removed using a glass rod lubricated with 10 x SSC. Two sheets of trimmed 3 MM filter paper soaked in 2 x SSC were placed carefully onto the nylon membrane, followed by a 10 cm stack of paper towels to draw up the

transfer buffer by capillary action. A window of plastic-wrap surrounding the gel and membrane was used to prevent the paper towels from absorbing transfer buffer via the platform wick directly. Pressure from a 1 kg weight placed on a glass or plastic plate was used to compress the stack uniformly. Transfer was allowed to proceed for 16 - 24 h. During dismantling of the blot, well positions were marked on the non-DNA side of the nylon membrane while *in situ*, and then the membrane was removed and washed briefly in 2 x SSC to remove agarose gel fragments and left to dry at room temperature for 1 h. Routinely, the agarose gel was restained with ethidium bromide and visualised on a UV transilluminator (Section 2.5.5) to verify that fractionated DNA had been transferred to the nylon membrane. DNA was cross-linked irreversibly to the dry nylon membrane on a UV transilluminator for 4.5 min, and the membrane was stored dry between 3 MM filter paper until use.

#### *Probe isolation and labelling*

Double-stranded DNA fragments for use as probes in Southern and northern hybridisation were prepared after restriction enzyme digestion of appropriate plasmid vectors (Section 2.5.4) and electrophoresis through agarose gels (Sections 2.5.5). Samples of digested DNA were fractionated to yield 100 - 500 ng of the appropriate DNA insert fragment, and the agarose gel was stained for only 3 - 4 min with 0.5  $\mu\text{g/ml}$  ethidium bromide to prevent extensive intercalation into the DNA double-strands which can affect downstream enzymatic applications. A gel slice containing the probe fragment was excised from the agarose gel on a UV transilluminator using a clean scalpel and transferred to a pre-weighed 1.5 ml microcentrifuge tube. Exposure of the DNA to UV light was minimised to avoid radiation damage such as DNA nicking, which reduces efficiency of labelling. Gel slices containing DNA fragments could be stored frozen at  $-20^{\circ}\text{C}$ . DNA fragments were extracted from agarose gel slices using a QIAEX DNA Gel

Extraction kit (Hybaid, UK) following the manufacturer's protocol. Solubilisation buffer QX1 (300  $\mu$ l/100 mg gel) and resuspended QIAEX particles (1  $\mu$ l/500 ng DNA) were added to the gel slice and the mixture incubated at 50°C for 10 min with a brief vortex every 2 min. During this incubation, DNA is released from the agarose slice and binds to the QIAEX particles. The QIAEX containing DNA was centrifuged in a benchtop microcentrifuge at 13000 rpm for 30 s, and then washed twice with 500  $\mu$ l of buffer QX2 to remove agarose contaminants. The pellet was washed twice with 500  $\mu$ l of buffer QX3 to remove salt contaminants. The QIAEX-DNA pellet was then air-dried for 10 - 15 min, and the DNA eluted in 20  $\mu$ l of TE buffer with an incubation in TE for 5 min at room temperature. After a final 30 s centrifugation to pellet the QIAEX particles, the supernatant containing DNA was collected carefully and transferred to a clean microcentrifuge tube. The concentration of the DNA probe fragment was determined by comparison with  $\lambda$ /HindIII molecular weight markers fractionated on an agarose gel.

Probe DNA was labelled using the Multiprime DNA Labelling System 1601Y (Amersham, UK) according to the principle of random primer extension (Feinberg & Fogelstein, 1983; 1984). 25 ng of DNA in a total volume of 28  $\mu$ l was boiled for 3 min to denature DNA strands, then placed on ice. The reaction mixture contained: 28  $\mu$ l denatured probe DNA, 10  $\mu$ l labelling buffer (including dATP, dGTP and dTTP), 5  $\mu$ l hexamer random primers/BSA mix, 5  $\mu$ l [ $\alpha$ -<sup>32</sup>P]-dCTP (specific activity 3000 Ci/mmol, ICN Flow, UK) and 2  $\mu$ l DNA Polymerase I Klenow fragment. The reaction was incubated at 37°C for 1 h. Unincorporated nucleotides were removed from the reaction mixture by gravity-flow passage through Nick Columns (Pharmacia, UK). After equilibration of the column with TE buffer, the labelling reaction products (50  $\mu$ l) mixed with 10  $\mu$ l of a tracer dye (a saturated solution of blue dextran) were added to the column and the radiolabelled DNA fragment eluted with TE buffer.

### *Prehybridisation and hybridisation*

Fixed nylon membranes were prewet in 5 x SSPE (20 x SSPE contains 3 M NaCl, 0.2 M NaH<sub>2</sub>PO<sub>4</sub>.2H<sub>2</sub>O and 20 mM EDTA, pH 7.4), then incubated (DNA-side facing inwards) at 65°C for 5 h in a Hybridizer HB-10 hybridisation oven (Techne, UK) with prehybridisation solution (5 x SSPE, 6% [w/v] PEG 6000, 0.5% [w/v] Marvel non-fat milk powder, 0.1% [w/v] tetrasodium pyrophosphate, 1% [w/v] SDS and 250 µg/ml heat-treated, sheared DNA from herring testes; a 10 mg/ml stock of DNA from herring testes was boiled for 10 min, passed through a 0.8 x 40 mm syringe needle 12 times, boiled for a further 10 min, aliquoted into 1 ml fractions and stored at -20°C, and denatured by boiling for 3 min immediately prior to use) to avoid non-specific hybridisation of radiolabelled DNA. The <sup>32</sup>P-labelled DNA probe was denatured by boiling for 6 - 7 min, added to the pre-hybridisation mixture, and hybridisation continued at 65°C for 16 - 24 h.

### *Washing and autoradiography*

Washes were preheated to 65°C. After discarding the hybridisation mixture, filters were cleared of non-specific hybridisation by washing at low stringency (2 x SSC, 0.1% [w/v] SDS, 65°C, 15 min, twice) and then subsequently at high stringency (0.2 x SSC, 0.1% [w/v] SDS, 65°C, 15 min, twice). Filters were inverted at least once to wash the underside. Washed filters (either after low or high stringency washes) were sealed wet in plastic bagging, and exposed to X-Omat AR (Kodak, UK) or RK (Fuji, UK) X-ray film at -80°C between intensifier screens in an autoradiography cassette for 2 - 28 days.

### *Gene copy number determination*

A reconstruction experiment to determine gene copy number by Southern hybridisation was calculated using a method

described in Sabelli and Shewry (1993). Given that an average bp is equivalent to  $1.036 \times 10^{-9}$  pg/mol DNA and that the haploid genome size of *A. thaliana* is approximately  $1 \times 10^8$  bp (Wilson et al. 1991), one genome will be 0.1036 pg/mol DNA. In one  $\mu\text{g}$  of DNA, there are thus:

$$1 \times 10^6 \text{ pg} \times 1 / (0.1036 \text{ pg/mol}) = 9.65 \times 10^6 \text{ mol.}$$

Using a DNA fragment insert of 1.4 kb (the Sat-1 clone; Chapter 3.2.3), the amount of insert DNA corresponding to 1 copy per haploid genome is:

$$9.65 \times 10^6 \text{ mol} \times 1.036 \times 10^{-9} \text{ pg/mol} \times 1.4 \times 10^3 \text{ bp} = 14 \text{ pg.}$$

With 5  $\mu\text{g}$  of *A. thaliana* DNA used for Southern hybridisation, the equivalents are thus:

1 copy	-	70 pg insert
2 copies	-	140 pg insert
5 copies	-	350 pg insert
10 copies	-	700 pg insert.

Comparison of another three published methods for calculating gene copy number revealed that one yielded identical results (Croy et al. 1993), while the other two produced values which were 13 % (Chee et al. 1991) and 27 % higher (Pasternak 1993).

## 2.7 Isolation and analysis of *Arabidopsis thaliana* RNA

### 2.7.1 General precautions against RNases

All equipment and reagents used for the isolation and analysis of intact RNA were protected against RNase contamination following guidelines in Sambrook et al. (1989). Distilled water and reagents were treated with 0.1 % (v/v) diethyl pyrocarbonate (DEPC), a strong inhibitor of RNases, except for buffers containing Tris or 3-[N-morpholino]propanesulphonic acid [MOPS] that have amine groups which react with DEPC. The DEPC-water and reagents were allowed to stand at room temperature for several hours and then autoclaved to denature the DEPC. Glassware and metalware (eg. spatulas) were baked overnight at 180°C. Disposable plasticware was used without pretreatment, whereas equipment that could not be heat-treated was decontaminated by soaking in 0.1 % (w/v) SDS and rinsed thoroughly with DEPC-water. Where possible, a separate stock of chemicals only for work with RNA was used. Clean disposable gloves were worn during all handling and experimental procedures.

### 2.7.2 Extraction of *Arabidopsis thaliana* total RNA

Total RNA was extracted from *A. thaliana* leaf or above-ground (leaf, stem, flower and silique) tissue using a method adapted from Logemann et al. (1987). 0.5 g frozen tissue (stored at -80°C) was ground to fine powder with mortar and pestle in liquid nitrogen. The powder was transferred to a pre-cooled 1.5 ml microcentrifuge tube and 2 volumes of guanidine buffer (8 M guanidine hydrochloride, 20 mM 4-morpholineethanesulfonic acid, 20 mM EDTA and 50 mM  $\beta$ -mercaptoethanol, pH 7.0), a strong protein denaturing solution for inhibition of RNases, was added and mixed by vortexing. The sample could be stored at room temperature after this step for up to 2 h. After microcentrifugation for

10 min at 13000 rpm (4°C), the supernatant was removed and filtered through a nylon mesh (100 µ pore size) to remove residual particulates. 0.2 - 1 volume of phenol/chloroform/IAA (25:24:1 volume ratio) was added, mixed by vortexing for 5 s, and centrifuged in a microcentrifuge for 15 min at 13000 rpm, room temperature. The RNA-containing aqueous phase was collected, mixed with 0.7 volumes of ethanol (pre-cooled to -20°C) and 0.2 volumes of 1 M acetic acid and incubated overnight at -20°C (or for 1 h at -80°C) to precipitate RNA selectively. RNA was collected by microcentrifugation at 13000 rpm for 10 min (4°C) and washed twice with 3 M sodium acetate (pH 5.2) to remove low molecular weight (degraded) RNA and contaminating polysaccharides. Salt was removed by final wash in 70 % (v/v) ethanol, the RNA pellet dried at room temperature for 20 min and redissolved in 50 µl DEPC-water. RNA was quantified spectrophotometrically, with an OD<sub>260</sub> of 1.0 corresponding to a concentration of RNA of 40 µg/ml (Sambrook et al. 1989), or by comparison of known concentrations of RNA markers fractionated on a formaldehyde agarose gel (Section 2.7.4). Yields of approximately 50 µg total RNA/g fw tissue obtained from *A. thaliana* were lower than that published for *Solanum tuberosum* tissue using the original method (Logemann et al. 1987), but the RNA quality was good (Chapter 3.2.4.2).

### 2.7.3 Extraction of *Arabidopsis thaliana* mRNA

The Dynabeads mRNA DIRECT kit (Dynal, UK) was used to isolate *A. thaliana* mRNA from aerial tissue. The kit contains Dynabeads, magnetic polystyrene beads with covalently bound oligo (dT)<sub>25</sub> sequences that bind poly(A) residues contained in most mRNA species. Frozen tissue (0.15 - 0.3 g) was ground to a fine powder in liquid nitrogen, mixed with 1 ml Lysis/binding buffer, transferred to a cold heat-treated glass tube and homogenised on ice for 1 - 2 min with a S25 N-10 dispersing tool (Janke & Kunkel, Germany) attached to an

Ultraturrax mechanical grinder. Debris was removed by microcentrifugation for 30 s at 13000 rpm, 4°C, and the supernatant incubated for 5 min at room temperature with 2 mg conditioned Dynabeads (Dynabeads were conditioned by washing once in Lysis/binding buffer). Microcentrifuge tubes were placed in a Dynal MPC magnetic separator for 2 min, and the supernatant removed. The beads were washed twice with Washing buffer, and the mRNAs eluted at 65°C for 2 min into 10 µl Elution solution. Quality of mRNA was estimated by autoradiography after fractionation of first-strand cDNA synthesised in the presence of [ $\alpha$ -<sup>32</sup>P]-dCTP (Section 2.7.5).

#### 2.7.4 Northern analysis of RNA

##### *Agarose gel electrophoresis*

RNA was fractionated on a denaturing formaldehyde agarose gel as described in Davis et al. (1986). The gel was prepared by melting 1.5 % (w/v) agarose in 1 x MOPS buffer (a 10 x MOPS stock solution consists of 200 mM MOPS, pH 6.0, 50 mM sodium acetate and 10 mM EDTA), cooling to 50°C and then adding a final concentration of 2 % (v/v) formaldehyde. The agarose solution was poured into the gel casting plate in a fume cupboard to avoid inhalation of formaldehyde fumes. When set, the gel was placed into an electrophoresis tank and immersed in 1 x MOPS buffer. RNA samples (up to 20 µg) were vacuum-dried, mixed with 15 µl RNA loading buffer (72 % [v/v] deionised formamide, 1.6 x MOPS buffer, 9.6 % [v/v] formaldehyde, 8 % [v/v] glycerol and 8 % [v/v] of a saturated bromophenol blue solution) and 0.6 µl of a 1 mg/ml ethidium bromide solution (final concentration 40 µg/ml), and then denatured for 10 min at 65°C. 3 µl (3 µg) of an RNA ladder (Gibco BRL) mixed with 12 µl of RNA loading buffer and 0.6 µl of 1 mg/ml ethidium bromide was denatured similarly. Prior to loading the RNA samples, the gel was run at 70 V for 5 min and the wells were flushed with buffer. After loading, the

gel was run overnight at 26 V, for 3.5 h at 60 V, or for 2 h at 120 V. Running buffer was circulated by means of a peristaltic pump. The lower left corner of the gel was nicked for orientation and alignment of the northern transfer filter, and then the fractionated RNA was visualised on a UV transilluminator using ethidium bromide fluorescence and photographed with black and white Polaroid film (Section 2.5.5).

#### *Northern transfer*

The formaldehyde agarose gel was rinsed with two changes of 10 x SSC (made up in DEPC-water) to remove the formaldehyde. The gel was then set up for transfer of RNAs to an Amersham Hybond N nylon filter (prewet in 10 x SSC buffer for 3 min) using capillary action, as described for Southern transfer (Section 2.6.3). After 16 - 24 h, the transfer stack was dismantled partially, sample wells marked on the filter with a pencil, and the filter washed in 6 x SSC to remove residual agarose. The filter was air-dried for 30 - 60 min, and the RNA fixed to the filter by exposure to UV on a transilluminator for 4.5 min. Dried, fixed filters were stored between two layers of Whatman 3MM paper at room temperature.

#### *Prehybridisation and hybridisation*

Following Sambrook et al. (1989), fixed filters were prewet in 5 x SSPE buffer for 5 min before being incubated with prehybridisation solution (40 % [v/v] deionised formamide, 5 x SSPE, 0.5 % [w/v] SDS, 5 x Denhardt's solution and 250 µg/ml boiled sheared DNA from herring testes; 50 x Denhardt's stock solution is comprised of 1 % [w/v] of each of Ficoll 400, polyvinylpyrrolidone 360 and BSA, filter-sterilised and stored at -20°C) at 42°C in a hybridisation oven. Filters were prehybridised for 3 - 6 h before being hybridised to boiled radioactively-labelled probe DNA (Section 2.6.3) at 42°C for

16 - 24 h.

### *Washing and autoradiography*

Filters were washed twice each for 15 min with 2 x SSC, 0.1 % SDS (w/v) and then 0.2 x SSC, 0.1 % SDS (w/v) at 42°C. If high background radioactivity could be detected with a hand-held Mini-monitor Type 5.10 Geiger counter (Mini Instruments, UK), the last wash was repeated at 55°C and, if necessary, at 65°C. Washed filters were sealed wet in plastic bagging, and exposed to X-ray film at -80°C between two intensifier screens in an autoradiography cassette for 2 - 21 days.

### **2.7.5 Synthesis of radiolabelled first-strand cDNA to assess poly(A) RNA quality**

Radiolabelled first-strand cDNA was synthesised from poly(A) RNA by reverse transcriptase in the presence of [ $\alpha$ -<sup>32</sup>P]-dCTP (Sambrook et al. 1989). 12  $\mu$ l of poly(A) RNA isolated from *A. thaliana* (Section 2.7.3) was incubated at 70°C for 3 min, briefly spun down in a microcentrifuge and placed on ice. To the tube was added 1  $\mu$ l 20  $\mu$ M oligo(dT) 17-mer primer (final concentration: 1  $\mu$ M), 4  $\mu$ l 5 x reaction mix (final concentration: 50 mM Tris-HCl, pH 8.3, 75 mM KCl and 3 mM MgCl<sub>2</sub>), 0.5  $\mu$ l dNTP mix (final concentration: 0.25 mM each), 0.5  $\mu$ l [ $\alpha$ -<sup>32</sup>P]-dCTP (specific activity 3000 Ci/mmol, ICN Flow, UK), 1  $\mu$ l 40 units/ $\mu$ l RNAGuard RNase inhibitor (Pharmacia Biotech, UK) (final concentration: 2 units/ $\mu$ l), and 1  $\mu$ l 200 U/ $\mu$ l recombinant Moloney-murine leukemia virus reverse transcriptase (Clontech Laboratories, USA) (final concentration: 1 U/ $\mu$ l). The mixture was incubated at 42°C for 1 h, and the reaction terminated by incubation at 94°C for 5 min. The sample was fractionated on a 1 % (w/v) agarose gel (Section 2.5.5), the marker lane trimmed from the gel and photographed, and the rest of the gel wrapped in clingfilm and exposed to X-ray film between two metal plates.

## 2.8 DNA sequence determination

### 2.8.1 Introduction

Novel DNA sequence was determined using the enzymatic dideoxynucleotide chain-termination method (Sanger et al. 1977). The chain-termination method typically included the following steps: generation of single-stranded DNA template using alkaline denaturation of double-stranded plasmid DNA; annealing of a specific primer to template DNA; a labelling step, in which a DNA polymerase extended the primer in the presence of limiting amounts of deoxynucleoside triphosphates and radioactively labelled dATP; and an extension/termination step, where the deoxynucleoside triphosphate concentration was increased and a dideoxynucleoside triphosphate added to each of four separate reactions so that the DNA polymerase extended chains until they were randomly terminated by a dideoxynucleotide. Resulting radioactively labelled DNA fragments with the same origin but terminated at different nucleotides were separated according to size by high resolution polyacrylamide gel electrophoresis, and the DNA sequence could be determined following exposure of the dried gel to X-ray film.

Determining the sequence of large fragment of DNA requires dividing the fragment into smaller sections or reactions and then accumulating the information (Wu 1994). Three general strategies are used for this purpose (Sambrook et al. 1989; Slatko et al. 1993): (i) "primer walking" involves first obtaining DNA sequence information of a fragment cloned into a sequencing vector with vector-derived oligonucleotide primers. Using new oligonucleotides obtained from the 3' end of known sequence, further sequence information is then gained in subsequent reactions; (ii) another strategy requires generation of nested sets of deletions of the complete cloned fragment, followed by sequencing of the progressive deletion fragments. Several protocols are used to

create nested deletions enzymatically, the most common of which uses *E. coli* exonuclease III (Henikoff 1984, 1987); and (iii) "shotgun cloning" involves shearing or digesting by nucleases or restriction endonucleases the complete DNA sequence into smaller fragments which are then subcloned into appropriate vectors and sequenced. The complete sequence is obtained by computer-assisted assembly of the random individual overlapping fragment sequences, with gaps in the sequence filled by primer walking. Of these three strategies, primer walking is the most efficient as little redundant sequence information is generated and subcloning of fragments from the initial DNA fragment is not required: however, synthesis of new primer oligonucleotides is potentially a major time constraint (Kotler et al. 1993). Primer walking was the strategy employed here for obtaining the DNA sequence of large clones.

#### 2.8.2 Selection, manufacture and preparation of oligonucleotide primers

Guidelines followed for the manual selection of suitable sequencing oligonucleotide primers (Sambrook et al. 1989, and Mr A. Grierson, University of St Andrews, UK, pers. comm.) were: (i) ensure oligonucleotide is complementary to the correct strand in a sequence determined unambiguously; (ii) check that the primer has a balanced base composition, with a 40 - 55 % ratio of G + C nucleotides; (iii) check that there are no regions of dyad symmetry to avoid hairpin or stem-loop structures; (iv) check that the sequence is not complementary to vector sequence DNA and investigate identity with sequences in the GenBank database; (v) ensure primer is 17 - 25 bp in length; and (vi) attempt to have an A nucleotide at the penultimate 5' position (weaker 2 hydrogen bond A-T pairing), and a G nucleotide at the ultimate 3' position (stronger 3 hydrogen bond G-C pairing to anchor primer).

Custom oligonucleotides were purchased from Cruachem (UK) or synthesised in-house by Mr I. Armit, University of St Andrews, UK, using a 381A DNA Synthesiser (Applied Biosystems, UK). Oligonucleotides synthesised in-house required ethanol purification before use. Primers were resuspended in 100  $\mu$ l sterile water for 10 - 15 min at room temperature, then precipitated with 0.1 volumes of 3 M sodium acetate (pH 5.2) and 3 volumes ethanal (precooled to -20°C). The solution was incubated at -80°C for 15 - 20 min, then microcentrifuged for 15 min at 13000 rpm, 4°C, and the DNA pellet air-dried for 10 min. This purification step was repeated once more, and the pellet finally resuspended in 10  $\mu$ l sterile water. Primer concentration was determined by measuring OD<sub>260</sub>, then using the formula: primer concentration (in pmol/ $\mu$ l) = (OD<sub>260</sub>/{0.01 x N}) x Dilution factor, where N is the number of bases. Sequencing primers were diluted to a working concentration of 1 pmol/ $\mu$ l and stored at -20°C.

Vector-based oligonucleotides flanking the pYES vector cDNA insertion site and used for double-stranded DNA sequencing of 5' and 3' ends of *A. thaliana* cDNA inserts were 5'-TGT GGA ATT GTG AGC GG (P2) and 5'-ACT TTA ACG TCA AGG AG (P1), respectively (Elledge et al. 1991).

### 2.8.3 DNA sequencing with the Sequenase Version 2.0 DNA Sequencing kit

Sequenase Version 2.0 DNA polymerase used in the Sequenase Version 2.0 DNA Sequencing kit (United States Biochemicals, USA) is a genetically modified enzyme that synthesises DNA strands initiated at double-stranded sites where a specific oligonucleotide primer and the single-stranded DNA template are annealed. Modified Version 2.0 enzyme is devoid of 3'-5' exonuclease activity, unlike the wild-type T7 DNA polymerase (Tabor & Richardson 1987), and has high processivity and speed (Tabor & Richardson 1989). The Sequenase kit was used

most frequently for routine DNA sequencing.

#### *Denaturation of double-stranded DNA template*

3 - 10  $\mu\text{g}$  plasmid DNA in a volume of 50  $\mu\text{l}$  was alkaline-denatured in 0.1 volume 2 M NaOH, 2 mM EDTA at 37°C for 30 min. The solution was neutralised by adding 0.1 volume 3 M sodium acetate (pH 5.2), and DNA precipitated by adding 3 volumes -20°C ethanol and incubating at -80°C for 15 - 20 min. After microcentrifugation at 13000 rpm for 30 min, 4°C, the DNA pellet was washed in 70 % (v/v) ethanol, allowed to dry at room temperature, and kept on ice until required. Denatured template DNA was redissolved in 7  $\mu\text{l}$  sterile water immediately prior to use.

#### *Annealing template and primer*

Annealing mixture consisted of 7  $\mu\text{l}$  denatured template DNA (3 - 10  $\mu\text{g}$ ), 2  $\mu\text{l}$  Sequenase 5 x Reaction Buffer and 1  $\mu\text{l}$  sequencing primer (1 pmol/ $\mu\text{l}$ ). Template and primer were annealed by incubating in microcentrifuge tubes at 65°C for 2 min in a heating block, then cooling slowly to 35°C by leaving tubes in the heating block (turned off) for 10 min at room temperature and 15 min at 4°C. The tubes were kept on ice for up to 4 h before use.

#### *Labelling and extension/termination reactions*

Ice-cold annealed primer/DNA solution (7  $\mu\text{l}$ ) was mixed with: 1  $\mu\text{l}$  0.1 M DTT, 2  $\mu\text{l}$  diluted Labelling Mix (7.5  $\mu\text{M}$  of dGTP, dCTP and dTTP stock diluted 1:5 with sterile water), 0.5  $\mu\text{l}$  [ $\alpha$ -<sup>35</sup>S]-dATP (10  $\mu\text{Ci}/\mu\text{l}$  and 1000 Ci/mol stock, ICN Pharmaceuticals, UK) and 2  $\mu\text{l}$  diluted Sequenase Version 2.0 T7 DNA polymerase (13 units/ $\mu\text{l}$  stock diluted 1:8 with 6.5 volumes Enzyme Dilution Buffer and 0.5 volumes pyrophosphatase). The labelling reaction was incubated at room temperature for 3.5 - 4 min, then 3.5  $\mu\text{l}$  was withdrawn

and added to 2.5  $\mu$ l ddNTP Termination Mixture (prewarmed to 37°C) for each of the four termination reactions (G, A, T and C). Extension/termination reactions were continued for 4 - 5 min before being stopped by the addition of 4  $\mu$ l Stop Solution. Samples were stored at -20°C until fractionated on a denaturing acrylamide gel (Section 2.8.7).

#### 2.8.4 DNA sequencing with the TaqTrack Sequencing System

The TaqTrack Sequencing System (Promega Corporation, UK) employs the thermal stability of a modified Sequencing Grade *Taq* DNA Polymerase to perform chain-termination DNA sequencing reactions at elevated temperatures (70°C) where secondary structure of DNA templates is decreased and polymerisation through highly structured regions is theoretically possible (Innis et al. 1988; TaqTrack Sequencing Systems Technical Manual, Promega Corporation, USA, 1993). The TaqTrack system uses 7-deaza-dGTP rather than dGTP to alleviate problems of band compressions which often result from incomplete denaturation of GC rich sequences during electrophoresis.

#### *Alkaline denaturation of plasmid DNA*

4 - 10  $\mu$ g of plasmid DNA in a volume of 18  $\mu$ l was mixed with 2  $\mu$ l of 2 M NaOH, 2 mM EDTA, and the solution incubated at room temperature for 5 min. The reaction was neutralised by adding 8  $\mu$ l of 5 M ammonium acetate (pH 7.5), then DNA precipitated with 112  $\mu$ l ethanol and pelleted by microcentrifugation for 10 min at 13000 rpm, 4°C. The pellet was washed once with 70 % (v/v) ethanol, air-dried for 10 min, then resuspended in 16  $\mu$ l sterile water for primer/template annealing.

Although this denaturation protocol suggested by TaqTrack Sequencing System Technical Manual was used initially for *Taq*

DNA polymerase sequencing, better results were obtained using the Sequenase 2.0 denaturation protocol (Section 2.8.3) (data not shown). The latter method was thus used routinely.

#### *Annealing template and primer*

Template (4 - 10  $\mu\text{g}$  denatured plasmid DNA in a volume of 16  $\mu\text{l}$ ) and primer (2 pmol in a volume of 2  $\mu\text{l}$ ) were annealed in a solution containing 5  $\mu\text{l}$  5 x *Taq* DNA Polymerase Buffer and 2  $\mu\text{l}$  Extension/Labelling Mix at 37°C for 10 min.

#### *Labelling reaction*

0.5  $\mu\text{l}$  of [ $\alpha$ -<sup>35</sup>S]-dATP (10  $\mu\text{Ci}/\mu\text{l}$  and 1000 Ci/mol stock, ICN Pharmaceuticals, UK) and 1  $\mu\text{l}$  of Sequencing Grade *Taq* DNA Polymerase (5 units/ $\mu\text{l}$ ) were mixed with the annealed template/primer mixture and the reaction incubated at 37°C for 5 min.

#### *Extension/termination reaction*

6  $\mu\text{l}$  of extension/labelling reaction mixture was added to each of four tubes (G, A, T, C) containing 1  $\mu\text{l}$  deoxy/dideoxynucleotide mix, the reactions incubated at 70°C for 15 min, and terminated by the addition of 4  $\mu\text{l}$  Stop Solution. Samples were stored at -20°C until loaded onto denaturing acrylamide gels (Section 2.8.7).

#### *Chase step to eliminate false bands*

Reaction tubes were cooled to room temperature for 10 min after the 15 min termination reaction, then 1 unit (0.2  $\mu\text{l}$ ) of *Taq* DNA polymerase was added to each tube and the reactions incubated for a further 15 min at 70°C. The reaction was terminated by adding 4  $\mu\text{l}$  Stop Solution.

### 2.8.5 Direct DNA sequencing of PCR products

Double-stranded DNA products amplified from PCR reactions were fractionated by agarose gel electrophoresis (Section 2.5.5), purified from agarose, PCR primers and other contaminants using a QIAEX DNA Gel Extraction kit (Section 2.6.3), and finally resuspended in 7.5  $\mu$ l sterile water for 45 min at room temperature. Alkaline denaturation of PCR products yields poor results, so the QIAEX-purified DNA was denatured at 95°C for 2 min, followed by snap-cooling for 1 min in a -80°C isopropanol bath (Dorit et al. 1991). Primer (0.5  $\mu$ l, 1 pmol/ $\mu$ l stock) and 2  $\mu$ l 5 x Sequenase Buffer (Section 2.8.3) were added to the frozen DNA, the mixture thawed by microcentrifugation for 5 s, and the primer annealed to the template at room temperature for 30 min. Sequencing reactions using the Sequenase 2.0 DNA Sequencing System were then performed as described in Section 2.8.3.

### 2.8.6 End-labelled primer DNA sequencing

[ $\gamma$ -<sup>32</sup>P]-ATP end-labelled oligonucleotide primer prepared for primer extension experiments (Section 2.11) was used for sequencing DNA following a modified protocol of the Sequenase Version 2 DNA Sequencing System primer extension method (Section 2.8.3). Modifications were: end-labelled primer (1  $\mu$ l of 0.5 pmol/ $\mu$ l stock) was annealed to template DNA; 2  $\mu$ l of 1.5  $\mu$ M dATP stock was added to the labelling/extension step; and [ $\alpha$ -<sup>35</sup>S]-dATP was omitted from the labelling/extension reaction.

### 2.8.7 Preparation and running of polyacrylamide sequencing gels

The Sequi-Gen Nucleic Acid Sequencing System (Bio-Rad, UK) was used for casting and running polyacrylamide sequencing

gels. Glass plates were wiped with distilled water and then ethanol, the backplate (Integral Plate/Chamber) siliconised using dimethyldichlorosilane solution (2 % [w/v] in 1,1,1-trichloroethane), and the apparatus assembled with 0.5 mm spacers. A plug gel (quantities provided for 21 X 50 cm apparatus) was prepared by mixing 20 ml gel solution (6 % [w/v] acrylamide:N,N'-methylene-bisacrylamide [19:1] from an EASigel 40 % [w/v] acrylamide mix stock [Scotlab, UK], 7 M urea, 1 x TBE buffer; 10 X TBE buffer consists of 0.9 M Tris, pH 8.3, 0.9 M boric acid and 0.02 M EDTA), 100  $\mu$ l of 25 % (w/v) ammonium persulphate (stored for up to 2 weeks at 4°C) and 100  $\mu$ l N,N,N',N'-tetramethylethylenediamine (TEMED). The sealing gel was poured onto the casting tray, and the bottom end of the gel apparatus inserted to form a plug. The sequencing gel, prepared by mixing 60 ml gel solution, 60  $\mu$ l 25 % (w/v) ammonium persulphate and 60  $\mu$ l TEMED, was injected between the plates using a 60 ml syringe. An inverted sharkstooth comb was placed at the top end to form a horizontal loading surface, and the gel was left to polymerise for 2 - 16 h.

Prior to electrophoresis, the gel apparatus was set into the base, the upper chamber and base filled with 1 l of 1 x TBE running buffer (prewarmed to 50°C), and the sharkstooth comb inserted, teeth downwards, into the gel. The gel was heated to 50 - 55°C by applying a voltage of 2000 V, and 2 - 3.5  $\mu$ l sequencing reaction samples, incubated at 80°C for 2 - 2.5 min, were loaded. Electrophoresis was continued for 2 - 6 h at a voltage that maintained the gel at 50°C, with a second or third series of sample loadings separated by 2 h performed for longer runs.

#### 2.8.8 Gel drying and autoradiography

Following polyacrylamide gel electrophoresis, the sequencing gel apparatus was dismantled leaving the gel attached to the

upper (non-siliconised) glass plate. A sheet of Whatman No. 1 paper, trimmed slightly larger than the gel, was placed on top of the gel, a slight pressure applied evenly with a glass plate, and the gel peeled off attached to the paper. The gel was dried for 2 h at 80°C under vacuum pressure using a Model 583 Gel Dryer (Bio-Rad, UK) run on the sequencing cycle. Dried gels were exposed to X-ray film in autoradiography cassettes at room temperature for 24 h to 4 weeks.

## 2.9 Rapid amplification of cDNA ends (RACE)

The RACE technique, a PCR-based method for obtaining missing 5' or 3' ends of cDNAs, was used to amplify the 5' end of the *A. thaliana* clone Sat-1 (Chapter 3.2.3.2). The 5'-AmpliFINDER RACE Kit (Clontech Laboratories, USA), an improved modified strategy (Edwards et al. 1991) of the original RACE (Frohman et al. 1988, Belyavsky et al. 1989), was utilised. Positive control reactions performed prior to experimental work, using human placental poly (A) RNA and transferrin receptor primers included in the kit, confirmed that the 5'-AmpliFINDER RACE kit protocol was functional (Chapter 3.2.3.2).

### 2.9.1 Primers

The gene-specific primer PR1, a 30-mer primer for first-strand cDNA synthesis (5'-CAA CGA ACG CTG AGA AAC AAT CGA AGC GTG, 224 bp downstream from the 5' end of Sat-1), was chosen from a region where there was little identity between Sat-1 and other published *A. thaliana* serine acetyltransferase DNA sequences (data not shown). Gene-specific primers used for PCR amplification were a 28-mer primer PR2 (5'-GGC GAT ATC AGA TTT AGC CTC TTC TCG G, 169 bp from the 5' end of Sat-1), and a 17-mer primer P9 (5'-TAA GCG GAA ACA ATA GG, 203 bp from the 5' end of Sat-1). The AmpliFINDER Anchor primer sequence used in PCR amplification was 5'-CTG GTT CGG CCC ACC TCT GAA GGT TCC AGA ATC GAT AG.

### 2.9.2 cDNA synthesis

First-strand cDNA was synthesised from *A. thaliana* total RNA (Section 2.7.2) or mRNA (Section 2.7.3). The reaction was primed by preincubating 1  $\mu$ l primer PR1 (10  $\mu$ M stock) and 9  $\mu$ l mRNA ( $\pm$  1  $\mu$ g) or RNA ( $\pm$  10  $\mu$ g) at 65°C for 5 min. Reverse transcription was initiated by adding 20  $\mu$ l of a reverse

transcription master mix (final concentration in 30  $\mu$ l: 1 x reverse transcriptase buffer, 1.8 units RNase inhibitor, 1 mM dNTP ultrapure mix, 0.35 units avian myeloblastosis virus reverse transcriptase), the reaction incubated at 52°C for 30 min and terminated by adding 1  $\mu$ l of 0.5 M EDTA.

### 2.9.3 RNA hydrolysis, cDNA purification and precipitation

RNA was hydrolysed by adding 2  $\mu$ l of 6 N NaOH to 30  $\mu$ l of the cDNA synthesis reaction and incubating at 65°C for 30 min. The solution was neutralised by adding 2  $\mu$ l of 6 N acetic acid and 80  $\mu$ l 6 M NaI. To this was added 8  $\mu$ l of GENO-BIND suspension which selectively binds the cDNA, and the mixture was kept on ice for 10 min with occasional vortexing. After centrifugation in a microcentrifuge at 4°C for 10 s at 13000 rpm, the GENO-BIND pellet was washed twice with 80 % (v/v) ethanol and allowed to air-dry for 5 min. The pellet was resuspended in 50  $\mu$ l DEPC-water and incubated at 65°C for 5 min to release bound cDNA. After microcentrifugation for 2 min at room temperature, 13000 rpm, 45  $\mu$ l of the supernatant containing cDNA was removed. cDNA was precipitated by adding 2  $\mu$ l (20  $\mu$ g) glycogen as a carrier, 5  $\mu$ l 2 M sodium acetate and 100  $\mu$ l 96 % (v/v) ethanol and incubating at -20°C for 30 min. After microcentrifugation at 13000 rpm for 10 min, 4°C, the visible cDNA pellet was washed once with 80 % (v/v) ethanol, allowed to air-dry for 5 min and resuspended in 6  $\mu$ l DEPC-water. Purified first-strand cDNA was stored at -80°C.

### 2.9.4 Anchor ligation

The AmpliFINDER anchor (3'NH<sub>3</sub>-GGA GAC TTC CAA GGT CTT AGC TAT CAC TTA AGC AC-P 5') was ligated directly onto the 3' end of the first-strand cDNA using T4 RNA ligase in the presence of cobaltous chloride. In a volume of 10  $\mu$ l, the final reaction mixture contained 2.5  $\mu$ l purified first strand cDNA, 1 x

single-stranded ligation buffer, 4 pmol AmpliFINDER anchor and 10 units T4 RNA ligase. The ligation reaction was incubated at 22°C for 16 h, and stored at -80°C.

#### 2.9.5 PCR amplification

One  $\mu$ l of a 10-fold dilution of the anchor-ligated first-strand cDNA was used to amplify the 5' end of Sat-1 by PCR. The final reaction mixture in a volume of 50  $\mu$ l contained 1 x PCR reaction buffer (Promega Corporation, UK), 0.2 mM dNTP mix (Boehringer Mannheim, Germany), 2.5 units *Taq* DNA polymerase (Boehringer Mannheim, Germany, or Promega Corporation, UK), 0.2  $\mu$ M anchor primer, 0.2  $\mu$ M of primers PR2 or P9 and 1  $\mu$ l cDNA. The reaction mixture was heated to 82°C for 1 min prior to the addition of *Taq* DNA polymerase (a "hot start" to enhance PCR specificity; D'Aquila et al. 1991), and then cycled through differing denaturation, annealing and extension parameters (Chapter 3.2.3.2) using PHC-100 Thermal Cycler (Techne, UK) or Programmable Thermal Controller-100 (MJ Research, UK) PCR machines. Results were analysed by agarose gel electrophoresis (Section 2.5.5).

#### 2.9.6 Purification of PCR products

PCR DNA products were purified from primers, nucleotides and *Taq* DNA polymerase using a QIAquick-spin PCR Purification Kit (Hybaid, UK). Five volumes of buffer PB was added to the PCR reaction, mixed, and then applied to a QIAquick-spin column. The column was microcentrifuged for 1 min at 13000 rpm, washed with 0.75 ml buffer PE, and then eluted with 50  $\mu$ l 10 mM Tris-HCl (pH 8.5). Purified PCR products were used for direct sequencing (Section 2.8.5) or subcloned into appropriate vectors (Section 2.10).

## 2.10 Ligation of DNA into plasmid vectors

Plasmid vectors such as pBluescript II KS (Stratagene, UK) or pGEM-7Zf(+) (Promega Corporation, UK) and foreign DNA fragments were digested with appropriate restriction enzymes creating dissimilar cohesive termini, or cohesive and blunt end termini, for directional cloning. Promega Corporation (UK) restriction enzymes, quality control tested for ligation and transformation with the Blue/White Cloning Assay, were used wherever possible. Digested fragments and plasmid DNA were fractionated by agarose gel electrophoresis, desired bands excised from the gel and purified using a QIAEX DNA Gel Extraction Kit (Section 2.6.3). Ligation reactions were set up to contain a final concentration of: 1 x T4 DNA ligase reaction buffer (including 0.5 mM ATP), 0.3 - 1.5 Weiss units bacteriophage T4 DNA ligase (Boehringer Mannheim, Germany, or Promega Corporation, UK), and a range of molar ratios of vector to insert DNA. Negative control reactions without plasmid and/or insert were routinely performed. Ligation reactions were incubated at 15°C for 16 h, and then aliquots were used to transform competent cells (Section 2.5.3).

## 2.11 Determination of mRNA 5' ends by primer extension

The primer extension protocol (Arndt et al. 1989) was used to map the 5' end of the *A. thaliana* Sat-1a mRNA species (Chapter 3.2.3.2).

### 2.11.1 Primer labelling

An 18-mer primer, P172 (5'-AGG AAA GAG TGA TGA TTG, complementary to bp 91 - 108 of Sat-1a and derived initially from the *A. thaliana* EST clone 89K20T7), was end-labelled with [ $\gamma$ - $^{32}$ P]-ATP in a reaction containing: 0.5  $\mu$ l primer P172 (10 pmol/ $\mu$ l stock), 1  $\mu$ l 10 x polynucleotide kinase buffer, 8.1  $\mu$ l of [ $\gamma$ - $^{32}$ P]-ATP (2 mCi/ml) and 0.4  $\mu$ l polynucleotide kinase (10 units/ $\mu$ l, Promega Corporation, UK). The reaction was incubated at 37°C for 45 min and the kinase inactivated by heating to 90°C for 2 min. An aliquot of 5  $\mu$ l was removed and stored at -20°C for primer-labelled sequencing (Section 2.8.6), while the remaining 5  $\mu$ l was diluted with 20  $\mu$ l DEPC-water (final concentration of primer: 100 fmol/ $\mu$ l) for further use in the primer extension protocol.

### 2.11.2 Primer hybridisation and extension

*A. thaliana* total RNA (6.4  $\mu$ l, about 6  $\mu$ g) was hybridised with 1  $\mu$ l of 100 fmol/ $\mu$ l end-labelled primer at 65°C for 2 min, followed by a gradual cooling for 25 min to about 35°C. The hybridised solution was mixed with 4  $\mu$ l of 5 x reverse transcriptase buffer, 2  $\mu$ l 0.1 M dithiothreitol (DTT), 5  $\mu$ l 2 mM dNTP mix, 0.5  $\mu$ l 1 mg/ml actinomycin D and 0.1  $\mu$ l 40 units/ $\mu$ l RNAGuard RNase inhibitor (Pharmacia Biotech, UK). After a pre-incubation at 42°C for 2 min, 1  $\mu$ l Moloney murine leukemia virus reverse transcriptase (200 units/ $\mu$ l, Promega Corporation, UK) was added and the reaction continued for 30 min. DNA was precipitated by adding 2  $\mu$ l 3 M sodium acetate

(pH 5.2) and 50  $\mu$ l ethanol and incubating at  $-80^{\circ}\text{C}$  for 30 min. The solution was microcentrifuged at 13000 rpm for 20 min,  $4^{\circ}\text{C}$ , washed once with 70 % (v/v) ethanol, and resuspended in 3  $\mu$ l DEPC-water and 3  $\mu$ l sequencing stop solution (Section 2.8.3).

### 2.11.3 Electrophoresis

Primer extension product was analysed on a 6 % (w/v) acrylamide:N,N'-methylene-bisacrylamide (19:1)/ 7 M urea denaturing gel in parallel with sequencing reaction products from both  $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$  end-labelled primer (Section 2.8.6) and  $[\alpha\text{-}^{35}\text{S}]\text{-dATP}$  nucleotide sequencing (Section 2.8.3).

## 2.12 Analysis of DNA and protein sequences

The University of Wisconsin Genetics Computer Group (GCG) Software package (Devereux et al. 1984), Version 7.3.1 and 8.0.1, was used extensively for analysis of DNA and protein sequences. Programs most often employed were (in alphabetical order): BLAST (finds database sequences similar to a query sequence), CodonFrequency (tabulates codon frequency), DotPlot (makes dot-plot comparison between two sequences), Gap (finds optimal alignment between two sequences), LineUp (manual sequence editor for entering and manipulating multiple sequences), Map (shows enzyme restriction sites of nucleotide sequence), Pepdata (translates nucleotide sequence in all reading frames, forward and reverse), PileUp (shows optimal alignment of multiple sequences; see below), Pretty (arranges PileUp alignments), Publish (arranges sequences for publication), Seqed (sequence editor for entering and editing single sequences) and Translate (translates specified reading frame of a nucleotide sequence into corresponding amino acid sequence).

Sequence similarity searches were performed also using the BLAST electronic mail server (address: blast@ncbi.nlm.nih.gov), hosted by the National Center for Biotechnology Information (NCBI), Bethesda, USA (Henikoff 1993). This non-interactive server, accessing daily updates of the largest sequence databases including GenBank (Benson et al. 1994), European Molecular Biology Laboratory (EMBL) Data Library (Emmert et al. 1994), SWISS-PROT (Bairoch & Boeckmann 1994) and Protein Information Resource (PIR) (George et al. 1994), was found to be a convenient and effective method for similarity searches especially when in-house GCG database updates were not available. The BLAST (Basic Local Alignment Search Tool) heuristic algorithm rapidly finds ungapped, locally optimised sequence alignments between a query sequence and sequence databases (Altschul et al. 1990). Output is ranked according to an alignment identity

score for the highest-scoring ungapped maximal segment pair (MSP) that is found by pairwise comparison of the query sequence and database sequences. A corresponding Poisson probability value (P) gives the chance of an alignment identity score as high as that particular segment pair score being found in a random search of the same size. Low P values therefore can provide an indication of biologically significant sequence similarities. The BLAST algorithm is an order of magnitude faster than other sequence comparison tools such as FASTA and FASTB (Pearson & Lipman 1988).

Phylogenetic trees were constructed using GCG software from protein sequences aligned with the PileUp program (Feng & Doolittle 1990). In PileUp, an approximate phylogenetic order of input sequences is first constructed by a series of pairwise alignments according to Needleman & Wunsch (1970). The most closely related pair then serves as a starting point for a series of progressive pairwise alignments that incorporates increasingly dissimilar sequences until all sequences are aligned. Tree branching order and branch lengths are derived from the final PileUp alignment and the output is plotted as a dendrogram in which horizontal distances are proportional to evolutionary distance (Feng & Doolittle 1990; Doolittle 1994).

The Arabidopsis Research Companion database (Massachusetts General Hospital, Harvard, USA) was accessed by Internet Gopher server at the address "weeds.mgh.harvard.edu" and used for on-line searches to identify *A. thaliana* EST clones. Two Gopher servers ("weeds.mgh.harvard.edu" and "gopher.sunet.se") were used to access and search general molecular biology databases such as the PROSITE dictionary of protein sites and patterns (Bairoch 1993, Bairoch & Bucher 1994).

## 2.13 Expression and purification of GST-PRH-26 fusion proteins

The pGEX Fusion Vectors (Pharmacia Biotech, UK) are used to express and purify foreign polypeptides as fusions with *Schistosoma japonicum* glutathione S-transferase (GST) in bacterial systems (Smith & Johnson 1988). Soluble GST fusion proteins, expressed from a *tac* promoter and induced by IPTG, can be purified from bacterial lysates under non-denaturing conditions by affinity chromatography on glutathione-agarose beads and eluted using reduced glutathione. The pGEX system was used to express and purify a GST-PRH-26 fusion protein (Chapter 4.2.8).

### 2.13.1 Production of transformants

The pGEX-4T-2 plasmid (4950 bp) was restriction enzyme digested (Section 2.5.4) at the unique cloning restriction sites *Sma*I (creating a blunt end) and *Xho*I (creating a 5' overhang), fractionated on a 1 % (w/v) agarose gel (Section 2.5.5) and the 4.9 bp pGEX-4T-2/*Sma*I/*Xho*I fragment excised and purified using a QIAEX DNA Gel Extraction Kit (Section 2.6.3). *A. thaliana* cDNA clone Papsr-26 was digested with *Bgl*III (which cuts the Papsr-26 insert  $\pm$  120 bp downstream from the 5' terminus [68 bp downstream from the first ATG start codon], and the pYES plasmid at bp 4570) and the resultant 5' overhanging termini were filled to create blunt ends using the Klenow fragment of *E. coli* DNA polymerase (Sambrook et al. 1989). To the 20  $\mu$ l Papsr-26/*Bgl*III digestion mixture was added 1  $\mu$ l 1 mM dNTP mix and 1 unit Klenow (Boehringer Mannheim, Germany), and the reaction was incubated at room temperature for 15 min. After heat inactivation of Klenow for 10 min at 75°C (Note: *Bgl*III is not inactivated by this heat treatment, but will not digest the filled ends as the hexamer recognition sequence 5'-A↓GA TCT is repaired to 5'-AGA TC.), 2  $\mu$ l of the sample was removed

for electrophoretic analysis, and the remainder was digested with *Xho*I. The digested sample was fractionated on a 1 % (w/v) agarose gel (Section 2.5.5) and the 1.7 kb Papsr-26/*Bgl*III/*Klenow/Xho*I fragment excised and purified using a QIAEX DNA Gel Extraction Kit (Section 2.6.3). The purified Papsr-26/*Bgl*III/*Klenow/Xho*I fragment was ligated by directional blunt end/cohesive end cloning into purified pGEX-4T-2/*Sma*I/*Xho*I (Section 2.10) to form a pGST-PRH-26 plasmid, which was then transformed into competent *E. coli* strain TB1 cells (Section 2.5.3). A selection of TB1 colonies which grew overnight at 37°C on LB plates containing 50 µg/ml ampicillin were analysed further.

### 2.13.2 Analysis of GST recombinants for fusion protein expression

Two small-scale procedures for the analysis and optimisation of GST-PRH-26 fusion protein expression were used. In the first method, an adaptation of Smith and Corcoran (1994), individual *E. coli* strain TB1 colonies transformed with pGST-PRH-26 were grown at 37°C in 4 - 5 ml LB broth containing 50 µg/ml ampicillin for 3.5 - 5 h, then expression of the fusion protein induced by adding 0.4 mM IPTG and incubating at 30°C or 37°C for 1 - 3.5 h. Alternatively, an overnight culture of a transformed strain TB1 colony grown at 37°C in LB broth with 50 µg/ml ampicillin was diluted 1:10 into 4 - 5 ml fresh medium, grown for 1 - 4 h at 37°C and then induced with 0.2 - 0.4 mM IPTG at 30°C or 37°C for 1.5 - 6 h. Cells were centrifuged for 5 min at 7000 x g, room temperature, the pellet washed in cold 1 x phosphate-buffered saline (PBS; a 10 x stock consists of 8 % [w/v] NaCl, 0.2 % [w/v] KCl, 1.15 % [w/v] Na<sub>2</sub>HPO<sub>4</sub>.7H<sub>2</sub>O and 0.2 % [w/v] KH<sub>2</sub>PO<sub>4</sub>). Washed bacterial pellets could be stored at -20°C for up to 1 month (Frangioni & Neel 1993). Pellets were resuspended in 300 µl PBS, cells disrupted on ice in 3 - 6 bursts of 5 s each with an ethanol-sterilised Sonifer Cell Disrupter W185 probe sonicator (Heat

Systems-Ultrasonics, USA) and cellular debris removed by microcentrifugation at 4°C for 10 min, 13000 rpm. For purification of the GST-PRH-26 fusion protein, the supernatant was mixed for 5 min (room temperature) or 15 min (on ice) with 50 µl of a 50 % (w/v) glutathione-agarose slurry. (Lyophilised glutathione-agarose beads [Sigma, UK] were swelled in 1 x PBS or sterile water for 2 - 3 h, washed at least twice with 1 x PBS, and then resuspended to form a 50 % [w/v] slurry. Glutathione-agarose slurry was stored at 4°C for up to 1 month.) The glutathione-agarose beads were washed three times with cold 1 x PBS and resuspended in 1 volume of 2 x SDS sample loading buffer (2 x SDS sample buffer is: 125 mM Tris-HCl, pH 6.8, 4.1 % [w/v] SDS, 20 % [v/v] glycerol, 1 % [w/v] bromophenol blue and 0.57 mM β-mercaptoethanol) for analysis by SDS-polyacrylamide gel electrophoresis (SDS-PAGE; Section 2.15.1).

A second small-scale method, after Frangioni and Neel (1993), was used to enhance fusion protein solubility and binding to glutathione-agarose. Overnight cultures of a strain TB1/pGST-PRH-26 transformant grown at 37°C in LB broth with 50 µg/ml ampicillin were diluted 1:10 into 5 ml fresh medium, grown for 1.5 h at 37°C and then induced with 0.4 mM IPTG at 37°C for 4 - 6 h. Cells were centrifuged for 5 min at 7000 x g, room temperature, and washed in cold STE buffer (150 mM NaCl, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA). Washed pellets were resuspended in 300 µl STE buffer containing 100 mg/ml lysozyme and with or without the proteinase inhibitors phenylmethanesulfonyl fluoride (PMSF, 1 mM) and leupeptin (10 µM). After incubation on ice for 15 min, 33 µl of 50 mM DTT in STE (final concentration 5 mM) and 37 - 83 µl of 10 % (v/v) N-lauroylsarcosine (final concentrations 1 - 2 % [w/v], respectively) were added, the cells were sonicated on ice in 3 - 6 bursts of 5 s each and the sonicate clarified by microcentrifugation at 4°C for 10 min, 13000 rpm. The supernatant was mixed with 0.2 volumes of 10 % (v/v) Triton X-100 in STE (final concentration 2 % [v/v]) and then

incubated with 50  $\mu$ l of 50 % (w/v) glutathione-agarose slurry for 5 min at room temperature. The beads were washed 6 - 8 times in cold 1 x PBS, and resuspended either in 1 volume storage buffer (50 mM Hepes, pH 7.4, 150 mM NaCl, 5 mM DTT, 5 mM DTT, 10 % [w/v] glycerol) for storage at 4°C or -20°C, or in 1 volume of 2 x SDS sample loading buffer for analysis by SDS-PAGE (Section 2.15.1).

### 2.13.3 Large-scale purification of GST-PRH-26 fusion protein

Fusion protein from 500 ml of bacterial culture was expressed and purified as described for the small-scale procedure from Frangioni and Neel (1993), with appropriate corrections for increased volume. Cells was grown for 2 h at 37°C and induced with 0.4 mM IPTG for 5 h at 37°C. Final parameters were: 3 mg/ml lysozyme, 5 mM DTT, 1.5 % (v/v) N-lauroylsarcosine and 2 % (v/v) Triton X-100. The GST-PRH-26 fusion protein was eluted from 1.25 ml glutathione-agarose (bed volume) with a total volume of 0.94 ml glutathione elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM reduced glutathione, 5 mM DTT and 0.1 % [v/v] Triton X-100), glycerol added to a final concentration of 10 % (v/v), and the eluate stored at 4°C or -20°C.

## 2.14 Preparation of rabbit anti-GST-PRH-26 antibodies

Procedures described in this section were conducted with the assistance of Dr J. Somerville, University of St Andrews, UK. Prior to injection of antigen, 5 ml of blood was taken from a marginal ear vein of a New Zealand White rabbit to prepare a pre-immune serum. Blood collected into a sterilised glass test-tube was left at room temperature for at least 2 hours to allow clotting, the clot released from the tube wall with a heat-sterilised pasteur pipette, and the tube left at 4°C overnight. Serum was poured into a centrifuge tube, and contaminating clotted material cleared twice by centrifugation with a benchtop centrifuge at 4000 rpm for 10 min. The serum was aliquoted into 0.4 ml samples and kept at 4°C for short-term use or -40°C for storage.

Purified GST-PRH-26 fusion protein was prepared for injection as follows: a water-oil emulsion was made by mixing 1 volume oil (Drakeol 6VR [Pennsylvania Refining Company, USA] containing 0.1 volume Arlacel A emulsifier [Atlas, UK]) with the fusion protein solution (27.5 µg protein in a volume of 400 µl) by passing several times through a 0.8 x 40 mm syringe needle, and a water-oil-water emulsion was then made by mixing the oil-water emulsion with 1 volume of a 0.14 M NaCl solution containing a few drops of Tween 80 by passage through the syringe needle. The emulsion was injected subcutaneously in the back of the rabbit. After 52 days, 5 ml of blood was removed from the rabbit, and serum prepared as described for the pre-immune bleed. An antigen boost injection (21 µg purified GST-PRH-26 fusion protein) was administered on the same day, and a final bleed (10 ml) was taken two weeks later.

## 2.15 Western analysis of proteins

### 2.15.1 Protein fractionation by SDS-PAGE

Denaturing SDS-PAGE is used to separate proteins according to molecular weight (Laemmli 1970). A discontinuous gel system, with stacking and resolving gel layers as described in Sambrook et al. (1989), enables sharp banding of proteins. The Mini-PROTEAN II Electrophoresis Cell (Bio-Rad, UK) was used to fractionate proteins for analytical detection or for western blotting.

#### *Gel preparation and running*

Glass plates were cleaned with methanol and assembled on the gel casting tray according to the manufacturer's instructions. A 10 % ( $w/v$ ) acrylamide: $N,N'$ -methylene-bisacrylamide (37.5:1 [ $w/w$ ]), 0.1 % ( $w/v$ ) SDS resolving gel solution was prepared by mixing (per 5 ml) 1.9 ml distilled water, 1.7 ml 30 % ( $w/v$ ) EASigel acrylamide mix (Scotlab, UK), 1.3 ml 1.5 M Tris-HCl (pH 8.8), 0.05 ml 10 % ( $w/v$ ) SDS, 0.05 ml 10 % ( $w/v$ ) ammonium persulphate and 0.002 ml TEMED, injected between the glass plates using a 50 ml syringe with a 0.8 x 40 mm needle, and overlaid with isobutanol to prevent oxygen inhibition of polymerisation. When polymerised (after 30 - 45 min), the top of the gel was rinsed with distilled water to remove the isobutanol overlay and unpolymerised acrylamide, and then dried. A 5 % ( $w/v$ ) acrylamide: $N,N'$ -methylene-bisacrylamide (37.5:1 [ $w/w$ ]), 0.1 % ( $w/v$ ) SDS stacking gel solution was prepared by mixing (per 2 ml) 1.4 ml distilled water, 0.33 ml 30 % ( $w/v$ ) EASigel acrylamide mix, 0.25 ml 1.0 M Tris-HCl (pH 6.8), 0.02 ml 10 % ( $w/v$ ) SDS, 0.02 ml 10 % ( $w/v$ ) ammonium persulphate and 0.002 ml TEMED, and injected between the glass plates atop the resolving gel. A clean Teflon comb was immediately inserted into the stacking gel solution and the gel overlaid with isopropanol. After 30 - 45 min when polymerisation was complete, the comb was

removed and wells washed with distilled water. The gel was mounted in the electrophoresis apparatus, the lower and upper buffer reservoirs filled carefully with 1 x Tris-glycine electrophoresis buffer (5 x buffer consists of: 25 mM Tris, 250 mM glycine, 0.1 % [w/v] SDS) to avoid air bubbles, and the wells flushed with electrophoresis buffer. Protein samples were denatured by mixing with 1 volume 2 x SDS sample loading buffer and boiling for 3 min, then up to 20  $\mu$ l sample per well was loaded using microcapillary pipet tips fitted to a 20  $\mu$ l Pipetman (Gilson Medical Electronics, France). 5  $\mu$ l of Prestained SDS-PAGE Standard Solution (25 - 127 kDa; Sigma, UK) was used routinely for protein molecular weight markers. Wells without samples were filled with 20  $\mu$ l 1 x SDS sample loading buffer. Electrophoresis was conducted at 200 V for 40 - 45 min, until the bromophenol blue dye reached the end of the resolving gel. The apparatus was dismantled and gels carefully removed from the glass plates for Coomassie Brilliant Blue staining or western blotting.

#### *Coomassie Brilliant Blue staining*

SDS-polyacrylamide gels not destined for western transfer to nitrocellulose filters were stained with Coomassie Brilliant Blue and simultaneously fixed with methanol:glacial acetic acid (Sambrook et al. 1989). Gels were shaken gently in at least 5 volumes staining solution (prepared by dissolving 0.5 g Coomassie Brilliant Blue in 180 ml methanol:water 1:1 and 20 ml glacial acetic acid, then filtered through Whatman No. 1 paper to remove particulates) for 4 - 5 h at room temperature. Staining solution was decanted (to be reused), and the gel was destained for 16 - 24 h in destaining solution (prepared by mixing 360 ml distilled water, 360 ml methanol and 80 ml glacial acetic acid). Destaining solution was replaced four times. Gels were stored in distilled water, sealed in plastic bags at 4°C.

### 2.15.2 Western transfer

Following fractionation of protein samples by SDS-PAGE (Section 2.15.1), the electrophoresis apparatus was dismantled, and the polyacrylamide gel carefully removed and soaked for 30 min in cold transfer buffer (15.6 mM Tris, pH 8.3, 120 mM glycine, 20 % [v/v] methanol). A nitrocellulose filter (Hybond-C, Amersham, UK) was trimmed to the same size as the gel, and equilibrated in transfer buffer for 5 - 10 min. A gel sandwich consisting of the gel and the nitrocellulose filter flanked on each side by two pieces of 3MM paper (Whatman, UK) and a fibre pad was assembled in such a way that trapped air-bubbles were avoided, and the sandwich placed into a Mini Trans-Blot Cell (Biorad, UK) holder cassette. Protein was transferred from the gel to the nitrocellulose filter at 100 V for 1.5 h, the filter rinsed in distilled water and air-dried for 1 h. Filters could be orientated using transferred Prestained SDS-PAGE Standard Solution (Sigma, UK) molecular weight markers fractionated in the gel.

### 2.15.3 Staining of proteins immobilised on nitrocellulose filters

Proteins transferred to nitrocellulose were visualised using Ponceau S dye, a transient stain which is washed away during immunological detection of proteins (Sambrook et al. 1989). The dried filter was wetted in distilled water for 5 min, then transferred to Ponceau S stain solution (0.2 % [w/v] Ponceau S [3-hydroxy-4-(2-sulfo-4-[4-sulfophenylazo]-phenylazo)-2,7-naphthalenedisulfonic acid], 3 % [w/v] trichloroacetic acid, 3 % [w/v] sulfosalicylic acid) and shaken gently for 10 - 15 min. The filter was washed in distilled water to remove background staining and allow visualisation of transferred proteins.

#### 2.15.4 Immunodetection of proteins immobilised on nitrocellulose filters

Western blotted nitrocellulose filters were prewet in distilled water for 5 - 10 min. Nonspecific binding of immunological reagents was prevented by an initial blocking step in blocking solution (4 % [w/v] Marvel nonfat milk powder in TBST; TBST consists of 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.05 % [v/v] Tween 80) for 1 h at room temperature. Filters were removed from the blocking solution and incubated on a horizontal shaker for 2 h at room temperature or overnight at 4°C with the primary polyclonal rabbit anti-GST-PRH-26 fusion protein antibody (diluted 1/200 to 1/5000 in TBST containing 1 % [w/v] Marvel nonfat milk powder). After washing four times in TBST for 10 min each, filters were incubated with secondary enzyme-coupled antibody (alkaline phosphatase-linked goat anti-rabbit IgG [Sigma, UK] diluted 1/3000 in TBST with 1 % [w/v] Marvel nonfat milk powder) for 2 h at room temperature. The filters were then washed four times for 10 min each in TBST, and secondary antibody detected as follows using a chromogenic substrate solution which reacts with linked alkaline phosphatase: 66 µl of nitro blue tetrazolium (NBT) stock (5 % [w/v] NBT dissolved in 70 % [v/v] dimethylformamide) was mixed with 10 ml of alkaline phosphatase buffer (100 mM Tris-HCl, pH 9.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>), and to this was added 33 µl of a 5 % (w/v) 5-bromo-4-chloro-3-indolyl phosphate (BCIP) stock; the filters were incubated with the chromogenic substrate solution until dense blue precipitating bands of the desired intensity were observed, and then washed repeatedly in distilled water to stop the reaction (Sambrook et al. 1989). Filters were air-dried and stored at room temperature for photographic recording.

## 2.16 Protein quantification

Protein amounts were estimated using the protein-binding dye Coomassie Blue G-250 (Bradford 1976). 5 ml of dye reagent (0.01 % [w/v] Coomassie Brilliant Blue G-250, 9.35 % [v/v] orthophosphoric acid and 5.28 % [v/v] ethanol, allowed to stir for at least 2 h before filtration through two layers of Whatman No. 1 filter paper to remove particulates) was added to a range of BSA solutions (up to 100  $\mu$ g) for calibration of a standard curve and 100  $\mu$ l of the unknown protein solution (routinely diluted 20-fold). After colour development at room temperature for 15 min, absorbance at 595 nm was read using either a CamSpec UV-Visible Spectrophotometer or a Cecil CE272 Linear Readout Ultraviolet Spectrophotometer. Due to experimental variability, it was essential to perform the assay in triplicate for each sample. Dye reagent was stored at 4°C for up to two weeks. Linear equations from standard curves were calculated using Cricket Graph III (Computer Associates International, USA), and correlation coefficients (r) ranging from 0.99 - 1.00 were obtained.

## 2.17 Enzyme assays

### 2.17.1 $\beta$ -galactosidase

#### *Preparation of enzyme extract*

2.5 ml of an initial 37°C overnight shaking *E. coli* strain JM15 culture in M9 minimal medium, supplemented with 0.5 mM cystine, 0.2 % mannitol (w/v) and 1 mM IPTG, was inoculated into 100 ml of fresh M9 minimal medium with supplements and the cells grown to an OD<sub>420</sub> of 0.8 - 1.0. Cells were harvested with a 10 min centrifugation at 2500 x g, washed in 50 mM sodium phosphate buffer (pH 7.2), and recentrifuged as above. Pellets were stored at -80°C in pre-weighed microcentrifuge tubes for up to three months.

Crude strain JM15 extracts were obtained by grinding frozen pellets with mortar and pestle on ice with alumina (1:2 [w/w]) until cell disruption was evidenced by a claylike consistency (Hughes et al. 1971). After extraction with 2.5 ml 50 mM sodium phosphate buffer (pH 7.2) per g fresh weight pellet, a cell-free supernatant was obtained by microcentrifuge centrifugation for 20 min at 4°C, 13000 rpm. Crude extracts were kept on ice until use.

#### *Enzyme assay*

$\beta$ -galactosidase activity was assayed spectrophotometrically on a CamSpec UV-Visible Spectrophotometer at room temperature by following the increase in absorbance that occurs at 420 nm upon cleavage of the chromogenic substrate *o*-nitrophenyl- $\beta$ -D-galactose (ONPG, which is colourless) to *o*-nitrophenyl (yellow) and galactose (Dobrogosz 1981). The final reaction mixture contained 0.043 mM sodium phosphate buffer (pH 7.5), 1.3 mM reduced glutathione, 1 mM ONPG and 5 or 10  $\mu$ l enzyme extract in a final volume of 2.5 ml. Results were calculated as change in OD<sub>420</sub>/mg protein over time (in s).

## 2.17.2 Serine acetyltransferase

### *Preparation of enzyme extract*

Crude bacterial extracts from up to 500 ml cultures of *E. coli* strains JM15, SAT1, SAT2 and LE392 were prepared as for the  $\beta$ -galactosidase assay (Section 2.17.1). Crude extracts were prepared from *A. thaliana* above-ground tissue stored at  $-80^{\circ}\text{C}$ . About 1 g (fresh weight) of tissue was ground in liquid nitrogen, then extracted in 1 ml of 200 mM potassium phosphate buffer (pH 8.0), 10 mM  $\beta$ -mercaptoethanol, 1 mM PMSF, 1 mM EDTA and 1 mM DTT (Nakamura et al. 1987; Droux et al. 1992). A cleared supernatant, obtained after microcentrifugation at 13000 rpm,  $4^{\circ}\text{C}$ , for 20 min, was kept on ice prior to assaying for enzyme activity.

### *Enzyme assay*

Serine acetyltransferase activity was determined by following the L-serine-dependent decrease in  $\text{OD}_{232}$  that occurs upon cleavage of the thioester bond of acetyl CoA (Kredich & Tomkins 1966; Denk & Böck 1987). The assay was conducted at room temperature using a CamSpec UV-Visible Spectrophotometer. The final reaction mixture consisted of 50 mM Tris-HCl (pH 7.5), 0 - 0.15 mM acetyl CoA, 0 - 0.4 mM L-serine and appropriate amounts of enzyme extract in a total volume of 1 ml. Absorbance was followed at 232 nm for 2 min prior to the addition of L-serine, with which the L-serine dependent reaction was initiated. Results were expressed as nmol acetyl CoA utilised/mg protein/min, and a molar extinction coefficient of  $6.5 \times 10^3$  l/mol/cm for acetyl CoA (Denk & Böck 1987) was used. The minimum detection level of activity for the assay was 150 pmol/min.

### 2.17.3 PAPS reductase

#### *Preparation of enzyme extract*

1 ml of overnight *E. coli* cultures shaken at 37°C in LB medium containing 100 µg/ml ampicillin and 1 mM IPTG where appropriate was subcultured into 100 ml of fresh medium and grown to an OD<sub>600</sub> of 0.3 - 0.6. Cells were collected by centrifugation at 5000 x g for 10 min at 4°C, washed in 50 mM Tris-HCl (pH 8.0), 1 mM EDTA, and pelleted as previously. Bacterial pellets were stored at -20°C for up to 3 months.

For crude bacterial protein extracts, cells were resuspended in 0.005 volumes of 50 mM Tris-HCl (pH 8.0), 1 mM EDTA, disrupted on ice with 8 - 12 bursts of 5 s using an ethanol-sterilised probe sonicator, and particulate matter removed by microcentrifugation at 4°C for 10 min, 13000 rpm. A preliminary qualitative test using different cell resuspension volumes and sonication times established the conditions used here as optimal for disrupting cells and shearing DNA while avoiding overheating (results not shown).

Crude protein extracts were prepared from *A. thaliana* and *Spinacia oleracea* leaves following a method of Schwenn (1989). 0.3 - 0.85 g leaf material (deribbed for *S. oleracea*) was ground on ice with mortar and pestle in 1 ml/g fresh weight tissue of freshly prepared extraction buffer (50 mM potassium phosphate buffer, pH 7.7, 1 mM EDTA, 1 mM PMSF, 10 mM β-mercaptoethanol and 5 % [w/v] polyvinylpyrrolidone [Polyclar AT]). Cellular debris was removed by microcentrifugation at 13000 rpm and 4°C for 30 min. All extracts were kept on ice until use.

#### *Enzyme assay*

PAPS reductase activity was determined by measuring the formation of acid-volatile radioactivity from <sup>35</sup>S-PAPS

according to Schwenn and Schriek (1987).  $^{35}\text{S}$ -PAPS (specific activity  $4.8 - 10 \times 10^{10}$  Bq/mmol, du Pont, NEN Division, Germany) was obtained from Dr M. Coughtrie, Ninewells Hospital, Dundee, UK. The final complete assay mixture contained 50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 25 mM sodium fluoride, 5 mM DTT, 20 mM disodium sulphite,  $4.5 \mu\text{g}$  recombinant *E. coli* thioredoxin (Promega Corporation),  $2.25 - 5.0 \mu\text{M}$   $^{35}\text{S}$ -PAPS and crude enzyme extract in a total volume of  $100 \mu\text{l}$ . The reaction was performed in  $1.5 \text{ ml}$  microcentrifuge tubes at  $25^\circ\text{C}$ , initiated by the addition of crude extract and terminated after 5 min by the addition of  $100 \mu\text{l}$  acetone. Decapped microcentrifuge tubes were placed into  $15 \text{ ml}$  glass scintillation vials containing  $0.5 \text{ ml}$  trioctylamine. The distillation of volatile sulphur compounds was initiated by the addition of  $0.3 \text{ ml}$   $3 \text{ M}$  sulphuric acid to the microcentrifuge tube. Scintillation vials were then sealed and heated to  $60^\circ\text{C}$  for 45 min to release volatile radioactivity quantitatively, cooled to  $0^\circ\text{C}$  on ice, the microcentrifuge tube removed, and  $10 \text{ ml}$  of Ecoscint A scintillation fluid added to the trioctylamine. Samples were mixed thoroughly, left overnight to reduce quenching, and the radioactivity quantified as counts per minute (cpm) using a Beckman LS 6000SC Liquid Scintillation System. As a counting efficiency curve was not available for  $^{35}\text{S}$  detection, disintegrations per minute (dpm) from  $^{35}\text{S}$  were estimated using a  $^{14}\text{C}$  counting efficiency correction curve - this is feasible due to the similar energy emission range of  $^{14}\text{C}$  and  $^{35}\text{S}$  (Dr R. Griffiths, University of St Andrews, UK, pers. comm.).

Results were calculated as  $\text{pmol } ^{35}\text{S}\text{-PAPS utilised/mg protein/min}$  using the conversion factor  $1 \mu\text{Ci} = 2.22 \times 10^6 \text{ dpm}$  and correcting for radioactive decay of  $^{35}\text{S}$  (du Pont Safe Handling Guide for NEN Radionucleotides Manual).

## 2.18 Photography

Bacterial plates, certain x-ray films and some SDS-polyacrylamide gels were photographed using black and white film by Mr D. Roach, University of St Andrews, UK.

Black and white and Polaroid photographs and x-ray films were scanned with a ScanJet IICx/T scanner (Hewlett Packard, UK), edited with Photoshop Version 3.0 software (Adobe Systems, UK) and printed using a Pictrography 3000 processor (Fujix, UK).

CHAPTER 3: CHARACTERISATION OF ARABIDOPSIS THALIANA cDNAs  
ENCODING SERINE ACETYLTRANSFERASE CLONED BY FUNCTIONAL  
COMPLEMENTATION OF ESCHERICHIA COLI CYSE<sup>-</sup> STRAIN JM15

3.1 Introduction

Serine acetyltransferase catalyses the formation of *O*-acetylserine, which is incorporated with sulphide into cysteine by *O*-acetylserine (thiol)-lyase (Chapter 1.1.2.4). Experiments which demonstrated that feeding of exogenous *O*-acetylserine to *Cucurbita pepo* cells increased incorporation of <sup>35</sup>S-sulphate into cysteine, and that inhibition of endogenous *O*-acetylserine formation decreased sulphate incorporation into cysteine, led to the conclusion that supply of *O*-acetylserine is a dominant factor limiting cysteine formation (Rennenberg 1983). Studies with transgenic *Nicotiana tabacum* plants corroborate those findings, as higher rates of cysteine formation in isolated chloroplasts with elevated levels of *O*-acetylserine (thiol)-lyase were observed only when exogenous *O*-acetylserine was provided (Saito et al. 1994a). Further evidence that serine acetyltransferase activity may regulate cysteine biosynthesis has been provided by analysis of relative compartmental quantities of the cysteine biosynthetic enzymes. Although both serine acetyltransferase and *O*-acetylserine (thiol)-lyase can be found in cytosolic, chloroplastic and mitochondrial compartments of plant cells (Smith 1972; Brunold & Suter 1982; Saito et al. 1992; Rolland et al. 1993a; Ruffet et al. 1994; Saito et al. 1994b), their ratio within these compartments differs (Lunn et al. 1990; Ruffet et al. 1994, 1995). The amount of *O*-acetylserine (thiol)-lyase is 300-fold higher than that of serine acetyltransferase in *Pisum sativum* chloroplasts but only 3-fold higher in mitochondria (Ruffet et al. 1994, 1995). *In vitro*, the optimal ratio of *O*-acetylserine (thiol)-lyase to serine acetyltransferase for cysteine biosynthesis is 400:1

(Ruffet et al. 1994), which implies that, at least for some subcellular compartments, serine acetyltransferase levels may be critical for cysteine formation *in vivo*.

Serine acetyltransferase's product, *O*-acetylserine, has been implicated in cross-regulation of the converging pathways of sulphate and nitrate assimilation (Rennenberg 1983; Giovanelli 1990; Neuenschwander et al. 1991; Buwalda et al. 1992; Brunold 1993). Sulphate and nitrate assimilation pathways are thought to be co-regulated to meet protein synthesis requirements at a ratio of 1:25 to 1:32, the relative proportion of sulphur and nitrogen in proteins (Dijkshoorn & Van Wyk 1967; Reuveny & Filner 1977; Reuveny et al. 1980; Cacco et al. 1983). Numerous studies demonstrate the influence of nitrogen nutrition on sulphate assimilation enzymes (eg. Reuveny et al. 1980; Smith 1980; Brunold & Suter 1984; Saccomani et al. 1984; Barney & Bush 1985; Haller et al. 1986; Suter et al. 1986), and the effect of sulphur nutrition on nitrate assimilation enzymes (eg. Adams & Sheard 1966; Friedrich & Schrader 1978; Reuveny et al. 1980; DeBoer & Duke 1982; Haller et al. 1986; Clarkson et al. 1989; Neuenschwander et al. 1991). The nitrogenous compound *O*-acetylserine is thought to effect a positive regulatory role in sulphate assimilation, upregulating either ATP sulphurylase, and hence APS production (Giovanelli 1990), or APS sulphotransferase (Neuenschwander et al. 1991). In bacteria, *O*-acetylserine derepresses ATP sulphurylase, thereby allowing regulatory coupling of sulphate and nitrate assimilation (Kredich 1971; Umbarger 1978).

However, conflicting results on the role of *O*-acetylserine in regulatory coupling of sulphate and nitrate assimilation in plants have been obtained (Giovanelli 1990; Neuenschwander et al. 1991; Stulen & De Kok 1993), and control of sulphate assimilation enzymes are poorly understood in comparison with nitrate assimilation enzymes (Brunold 1993). In addition, a study has shown that extractable serine acetyltransferase

activity is not affected significantly by sulphate supply in *Phaseolus vulgaris* (Smith 1972), in contrast to sulphate permease (Chapter 1.2.1) and ATP sulphurylase (Chapter 1.2.2) activities which are regulated by sulphate supply. A regulatory role for serine acetyltransferase in cysteine biosynthesis has been excluded by Giovanelli (1990). Clearly, current evidence which supports the view that serine acetyltransferase catalyses a rate-limiting step in cysteine biosynthesis, and the potential importance of the enzyme's product, *O*-acetylserine, warrants further characterisation of the enzyme. As noted previously (Brunold 1990; Schmidt & Jäger 1992), understanding the role of sulphate assimilation enzymes would be facilitated greatly by analysis of the genes that encode them. This chapter describes work undertaken to obtain and characterise cDNAs encoding *Arabidopsis thaliana* serine acetyltransferase.

## 3.2 Results and Discussion

### **3.2.1 Complementation of the *Escherichia coli* *cysE* mutant JM15 to prototrophy with the $\lambda$ YES *Arabidopsis thaliana* cDNA library**

#### *Growth characteristics of Escherichia coli strain JM15*

The *Escherichia coli* strain JM15 (*cysE50*, *tfr-8*) is deficient in serine acetyltransferase and therefore cannot grow on sulphate, sulphite or sulphide as sole sulphur source (Jones-Mortimer 1968). An aliquot of strain JM15 was obtained from Dr B. Bachmann, *E. coli* Genetic Stock Center, Yale University, USA. Strain JM15 is sensitive to cysteine, therefore cystine was utilised as reduced sulphur source. Preliminary experiments were conducted to determine the growth characteristics of strain JM15 on M9 minimal medium (Sambrook et al. 1989) supplemented with 0.2 % (w/v) mannitol and with or without 0.5 mM cystine. Although *E. coli* grows more vigorously on carbon sources such as glucose, mannitol was required for subsequent complementation experiments to avoid repression of the  $\lambda$ YES *lac* promoter (Mandelstam 1962; Contesse et al. 1970; Magasanik 1970; Elledge et al. 1991). Strain JM15 was capable of growth on M9 minimal medium supplemented with cystine, but could not grow on basic M9 minimal medium in which sulphate as 2 mM MgSO<sub>4</sub> is the sole sulphur source (not shown). It was noted that growth of strain JM15 on M9 minimal medium was slower than on complete media such as LB (Sambrook et al. 1989), with typical incubation periods of 3 days at 37°C required on M9 minimal medium, compared with overnight incubations on LB medium (not shown). Spontaneous revertants of strain JM15 were not detected: the absence of spontaneous reversion would facilitate phenotypic selection of functionally complemented strains. Glycerol stocks of strain JM15 were prepared to ensure preservation of chromosomal markers, and in subsequent platings from the glycerol stocks controls were included

routinely to confirm cystine-dependence of the strain.

*Induction of helper phage  $\lambda$ KC and preparation of  $\lambda$ KC lysogens of Escherichia coli strain JM15*

The plasmid fragment, pYES, of the  $\lambda$ YES expression vector can be excised automatically by cre-lox site-specific recombination upon infection of *E. coli* cells expressing the cre protein (Sauer & Henderson 1988; Elledge et al. 1991). The helper phage  $\lambda$ KC (*kan-cre*) containing cre and kanamycin resistance genes (Elledge et al. 1991) was obtained from Dr J.T. Mulligan, Stanford University School of Medicine, USA, as a lysogen of *E. coli* strain BNN132 (strain JM107 lysogenised with  $\lambda$ KC). To induce  $\lambda$ KC from strain BNN132, protocols employing mitomycin C were employed (Otsugi et al. 1959; Korn & Weissbach 1962; Castellazzi et al. 1972b; Borek & Ryan 1973). However, in two separate experiments, lysis of strain BNN132 was not observed for up to 18 h in the presence of 1, 5 or 10  $\mu$ g/ml (final concentration) mitomycin C. Although responsiveness to mitomycin C has been found to be highly variable, most strains are induced after 10 - 90 min with 0.5 - 10  $\mu$ g/ml of the antibiotic (Otsuji et al. 1959; Reich et al. 1961; Korn & Weissbach 1962; Borek & Ryan 1973). Mitomycin C was unable to induce  $\lambda$ KC from strain BNN132 at the concentrations used and under these experimental conditions, so UV induction was attempted (Borek & Ryan 1958; Castellazzi et al. 1992a, b). As the incident dose of UV for maximal induction of lysogens is variable (Borek & Ryan 1973), varying exposure times with a Hanovia Bactericidal UV Unit at two different distances from the cells were examined. Optimal incident dose was found at 60 cm, with irradiation for 1 min: these conditions yielded  $\lambda$ KC from strain BNN132 at a concentration of  $1 \times 10^9$  pfu/ml (Table 3.1).

$\lambda$ KC lysogens of strain JM15 were prepared (Chapter 2.4.4), and kanamycin-resistant lysogens were selected initially on lambda agar medium supplemented with 40  $\mu$ g/ml kanamycin. When

---

**Table 3.1** UV induction of  $\lambda$ KC from *Escherichia coli* strain BNN132

<u>Exposure Time (min)</u>	<u>Bacteriophage Titre (pfu/ml)</u>	
	Lamp at 15 cm <sup>(1)</sup>	Lamp at 60 cm <sup>(2)</sup>
0	0	0
0.5	-	$7.5 \times 10^7$
1	$3.0 \times 10^6$	$1.0 \times 10^9$ <sup>(3)</sup>
2	-	$4.5 \times 10^8$
3	$3.0 \times 10^5$	$8.0 \times 10^7$
5	$3.0 \times 10^5$	$3.5 \times 10^7$
10	$6.0 \times 10^5$	0
20	-	0

Bacteriophage induced from *E. coli* strain BNN132 when exposed to UV radiation from a Hanovia Bactericidal UV Unit for different exposure times were quantified according to Davis et al. (1986). Volumes of logarithmic phase cell cultures contained within 9 cm culture dishes were 7 ml <sup>(1)</sup> or 6 ml <sup>(2)</sup>. Each experiment at different lamp distances was repeated once. Results are shown as the mean of two replicates, except for <sup>(3)</sup> where n=8.

---

infected with a range of  $10^4$  -  $10^6$  pfu/0.3 ml of cells, only the lowest dilution of  $\lambda$ KC yielded kanamycin-resistant colonies at a rate of 8.5 lysogens/ $10^6$  pfu (n=2). Fifteen strain JM15/ $\lambda$ KC kanamycin-resistant colonies were maintained as glycerol stocks, and subsequently were confirmed to retain both cystine-dependence and kanamycin-resistance when plated onto appropriate minimal medium (not shown).

*Complementation of Escherichia coli strain JM15/ $\lambda$ KC with the  $\lambda$ YES cDNA library*

Two aliquots of the *A. thaliana* cDNA expression vector  $\lambda$ YES (Elledge et al. 1991) were obtained from Dr J.T. Mulligan, Stanford University School of Medicine, USA. The  $\lambda$  library was titred according to Davis et al. (1986) at  $9.7 \times 10^9$  pfu/ml (aliquot 1) and  $1.04 \times 10^{10}$  pfu/ml (aliquot 2) (n=3 - 5; 2 separate experiments). The slightly more concentrated vector sample (aliquot 2) was used for all subsequent complementation experiments.

Infection of the *E. coli* *cysE* mutant lysogen JM15/ $\lambda$ KC with the  $\lambda$ YES cDNA library at a rate of  $6 \times 10^6$  pfu/ $10^9$  cells (recommended in Elledge et al. 1991) resulted in the appearance after 3 days of an average of 3 transformed colonies that were restored to prototrophy (two separate experiments), ie. which were ampicillin resistant (conferred by the  $\beta$ -lactamase gene of  $\lambda$ YES), and grew on minimal medium containing sulphate as the sole sulphur source. At a higher rate of infection ( $6 \times 10^7$  pfu of  $\lambda$ YES/ $10^9$  strain JM15/ $\lambda$ KC cells), 105 transformed and putatively complemented colonies were obtained. After replating onto fresh medium, a combined total of 110 strains that retained both the ability to grow with sulphate as sole sulphur source and which were ampicillin resistant were obtained. Glycerol stocks for long-term storage were prepared of each of the 110 JM15/ $\lambda$ KC/ $\lambda$ YES strains, designated as SAT1-110.

The  $\lambda$ YES phage library is converted to pYES plasmid clones by site-specific recombination upon infection into  $\lambda$ KC lysogens with approximately 67 % efficiency (Schnorr et al. 1994) - an efficiency of 100 % is theoretically feasible only in *recBC* strains where *ExoV* activity is absent (Sauer & Henderson 1988). It can be estimated that approximately  $4 \times 10^7$  ( $0.67 \times 6 \times 10^7$ ) strain JM15/ $\lambda$ KC transformants containing a pYES plasmid were obtained in the higher infection rate experiment here. Therefore the efficiency of complementation, ie. the number of cells restored to prototrophy in proportion to the total number of transformants, was approximately  $2.6 \times 10^{-6}$ . A wide range of cloning efficiencies has been obtained using functional complementation of *E. coli* mutants with plant cDNAs (Table 3.2). Cloning efficiency will be dependent on factors such as the quality and expression level of the cDNA library, the relative abundance of the specific plant gene(s), and whether a plant gene(s) catalyses the appropriate reaction to bypass or alleviate the mutation. Inability of plant cDNA expression libraries to complement bacterial lesions (eg. Stallmeyer et al. 1995) is most likely to be more frequent than published data suggest. It is interesting to note that subsequent to completion of *E. coli* *cysE* complementation experiments here, two other groups obtained clones for plant serine acetyltransferase with comparable cloning efficiency using functional complementation (Table 3.2; Ruffet et al. 1995; Saito et al. 1995).

#### *Retransformation of pSAT1 into Escherichia coli strain JM15*

Confirmation that restoration of strain JM15/ $\lambda$ KC to prototrophy was due to polypeptides encoded by pYES clones was obtained from retransformation experiments. The plasmid pSAT1 isolated from the complemented strain SAT1 and the empty plasmid pYES were each retransformed into the original mutant strain JM15. The strains JM15, JM15/pYES and JM15/pSAT1 were plated onto three media types: (A) consisted

**Table 3.2** Efficiency of cloning plant genes by functional complementation of *Escherichia coli* mutant strains

<u>Gene cloned</u>	<u>Complementation efficiency</u> (Complemented/Transformed)	<u>Vector type</u>	<u>Plant cDNA</u>	<u>Ref.</u>
PEP carboxylase	1.2x10 <sup>-4</sup>	pSI4001	<i>Zea mays</i>	1
OAS (thiol)-lyase	8.3x10 <sup>-5</sup>	pBluescript	<i>Citrullus vulgaris</i>	2
Enolase	5.0x10 <sup>-5</sup>	pUC8	<i>Z. mays</i>	3
Glutamine synthetase	1.0x10 <sup>-5</sup>	pUC13	<i>Z. mays</i>	4
SAICAR synthetase	1.0x10 <sup>-5</sup>	pcDNAII	<i>Arabidopsis thaliana</i>	5
Serine acetyltransferase	6.7x10 <sup>-6</sup>	pBluescript	<i>C. vulgaris</i>	6
AIR synthetase	3.7x10 <sup>-6</sup>	pcDNAII	<i>A. thaliana</i>	5
Adenylosuccinate lyase	3.0x10 <sup>-6</sup>	pcDNAII	<i>A. thaliana</i>	5
Serine acetyltransferase	2.6x10 <sup>-6</sup>	pYES	<i>A. thaliana</i>	7
Serine acetyltransferase	2.3x10 <sup>-6</sup>	pYES	<i>A. thaliana</i>	8
AICART IMP cyclohydrolase	1.0x10 <sup>-6</sup>	pcDNAII	<i>A. thaliana</i>	5
GAR synthetase	1.0x10 <sup>-6</sup>	pYES	<i>A. thaliana</i>	9
P5C reductase	6.7x10 <sup>-7</sup>	pBluescript	<i>Glycine max</i>	10
Superoxide dismutase	5.0x10 <sup>-7</sup>	pUC18	<i>Nicotiana glauca</i>	11
AIR synthetase	1.0x10 <sup>-7</sup>	pYES	<i>A. thaliana</i>	9
DHP synthase	5.0x10 <sup>-6 to -8</sup>	pUC13	<i>Z. mays</i>	12
MPT synthase (subunit)	0	pYES	<i>A. thaliana</i>	13

Data presented show a range of complementation efficiencies, the number of complemented *E. coli* cells as a proportion of total number of transformed cells, obtained with different plant cDNA libraries contained within different plasmid vectors.

Abbreviations:

AICART - 5'-phosphoribosyl-4-carboxamide-5-aminoimidazole transformylase  
 AIR - aminoimidazole ribonucleotide  
 DHP - dihydrodipicolinate  
 GAR - glycinamide ribonucleotide  
 IMP - inosine 5'-monophosphate  
 MPT - molybdopterin  
 OAS - O-acetylserine  
 P5C - Δ<sup>1</sup>-pyrroline-5-carboxylate  
 PEP - phosphoenolpyruvate  
 SAICAR - 5'-phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole

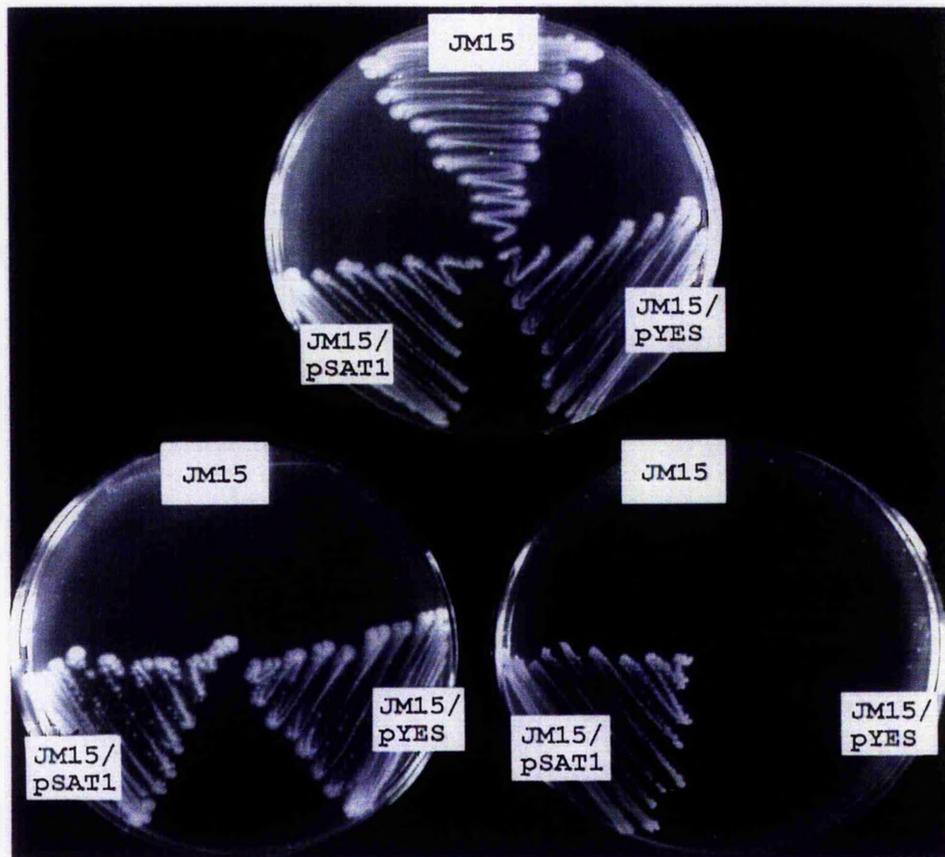
References:

- Izui et al. (1986)
- Noji et al. (1994)
- Lal et al. (1991)
- Snustad et al. (1988)
- Senecoff & Meagher (1993)
- Saito et al. (1995)
- This study
- Ruffet et al. (1995)
- Schnorr et al. (1994)
- Delauney & Verma (1990)
- Van Camp et al. (1990)
- Frisch et al. (1991)
- Stallmeyer et al. (1995)

of M9 minimal medium supplemented with 0.2 % ( $^*/_v$ ) mannitol and 0.5 mM cystine; (B) was the same as medium A but also contained 0.5 mM IPTG and 100  $\mu$ g/ml ampicillin; and (C) was the same as (B) but lacked cystine. All three strains were able to grow on medium A, but only the transformed strains JM15/pYES and JM15/pSAT1 grew in the presence of ampicillin (medium B) (Figure 3.1). Strain JM15/pSAT1 was restored to prototrophy and could grow on medium C containing ampicillin and lacking cystine, whereas strains JM15 and JM15/pYES could not. The possibility that strain JM15 is unable to grow on medium C solely due to the presence of ampicillin can be excluded by the inability of JM15/pYES to grow on this medium.

When different amounts of pSAT1 were used to retransform strain JM15, transformation and complementation efficiency were found to be essentially the same (Table 3.3). The pYES-derived plasmid pAPSK14 containing PAPS reductase homologue cDNA (Chapter 4.2.2.1) was unable to complement the *cysE* mutation when transformed into strain JM15 (Table 3.3). Collectively, these transformation experiments provide evidence that restoration to prototrophy of the mutant strain JM15 was conferred by pSAT1 rather than by reversion of the mutant, and also that complementation was achieved in the presence of the pSAT1 insert but not in plasmid pAPSK14 expressing a heterologous polypeptide. Retransformation results similar to those obtained for pSAT1 were obtained also for pSAT2, the plasmid isolated from another complemented JM15 strain, clone SAT2 (not shown).

Transformation efficiencies using the one-step method of Chung et al. (1989) for *E. coli* strain JM15 were found to be low (Table 3.3). Whereas those authors report routine yields of  $10^7 - 10^8$  transformants per  $\mu$ g of plasmid DNA, the highest yield obtained here was  $10^5$  transformants per  $\mu$ g plasmid. As expected (Chung et al. 1989), transformation efficiency was higher when lower quantities of plasmid were used.



**Figure 3.1** Growth characteristics of *Escherichia coli* cyst(e)ine auxotroph JM15 and the transformed strains JM15/pYES and JM15/pSAT1

*E. coli* strain JM15, deficient in serine acetyltransferase, strain JM15 transformed with the empty vector pYES (JM15/pYES) and strain JM15 transformed with pSAT1 (JM15/pSAT1) were plated onto M9 minimal medium containing sulphate as a sole sulphur source and 0.2 % (\*/.) mannitol, and supplemented with: 0.5 mM cystine (top); 0.5 mM cystine and 100 µg/ml ampicillin (bottom left); or 100 µg/ml ampicillin (bottom right). Plates were incubated at 37°C for three days.

---

**Table 3.3** Transformation and complementation of *Escherichia coli* strain JM15 by pSAT1

<u>Plasmid</u>	<u>Amount</u> (ng)	<u>Transformation</u> (colonies/ $\mu$ g DNA)	<u>Complementation</u> (colonies/ $\mu$ g DNA)
pSAT1	1000	1.25 ( $\pm$ 0.85) $\times 10^3$	1.16 ( $\pm$ 0.60) $\times 10^3$
	10	3.85 ( $\pm$ 0.68) $\times 10^4$	4.92 ( $\pm$ 0.96) $\times 10^4$
	1	9.81 ( $\pm$ 0.43) $\times 10^4$	1.10 ( $\pm$ 0.22) $\times 10^5$
	0	0	0
pAPSK14	10	2.39 ( $\pm$ 0.26) $\times 10^4$	0

*E. coli* strain JM15 was transformed with different amounts of plasmid pSAT1 extracted from the complemented strain SAT1 and plated onto M9 minimal medium supplemented with 0.2 % ( $w/v$ ) mannitol, 0.5 mM IPTG, 100  $\mu$ g/ml ampicillin and either with 0.5 mM cystine, to select transformed strain JM15/pSAT1 cells, or without cystine, which would allow growth of functionally complemented strain JM15/pSAT1 cells only. Cells plated onto M9 minimal medium supplemented with 0.2 % ( $w/v$ ) mannitol and 0.5 mM cystine but lacking ampicillin and IPTG to confirm viability always yielded confluent layers of growing cells (results not shown). Plasmid pAPSK14 was derived from complementation of *E. coli* strain JM81A defective in APS kinase (Chapter 4.2.1) and was used as a negative control for complementation. Results are from two separate experiments, with three replicates per treatment. Values shown are mean ( $\pm$  standard error).

---

Transformation probability has been found to decrease with increasing plasmid size (Hanahan 1983). Using the linear equation  $P_p[n \text{ kb}] = P_p[4 \text{ kb}] \times (4 \text{ kb}) / (n \text{ kb})$ , where  $P_p$  is the probability of transformation, and  $n$  represents an indefinite plasmid size between 4 and 70 kb (Hanahan 1983), it can be estimated that, relative to a 4 kb plasmid, the 8 kb pYES plasmid would have a 50 % reduction, and a 9.5 kb plasmid (pYES + 1.5 kb insert) a 42 % reduction, in transformation probability. The effect of plasmid size, therefore, does not explain the low transformation efficiencies obtained here. Different strains of *E. coli* are known to have dissimilar transformation efficiencies (Hanahan 1983, Chung et al. 1989), but data for strain JM15 are not available in the literature. A factor which may have affected transformation is that strain JM15, derived from *E. coli* K12 strain 703 (Jones-Mortimer 1968), has a wild-type restriction system that could cleave exogenous plasmid DNA (Sambrook et al. 1989). It was concluded that the one-step transformation protocol of Chung et al. (1989) was convenient and rapid, but not highly efficient. Reproducibility and reliability of the one-step method remains to be determined by other users (Seidman et al. 1989). In subsequent transformation experiments, more established protocols employing  $\text{CaCl}_2$  (Sambrook et al. 1989; Seidman et al. 1989) were utilised.

### **3.2.2 Enzymatic analysis of JM15, wild-type and complemented SAT strains**

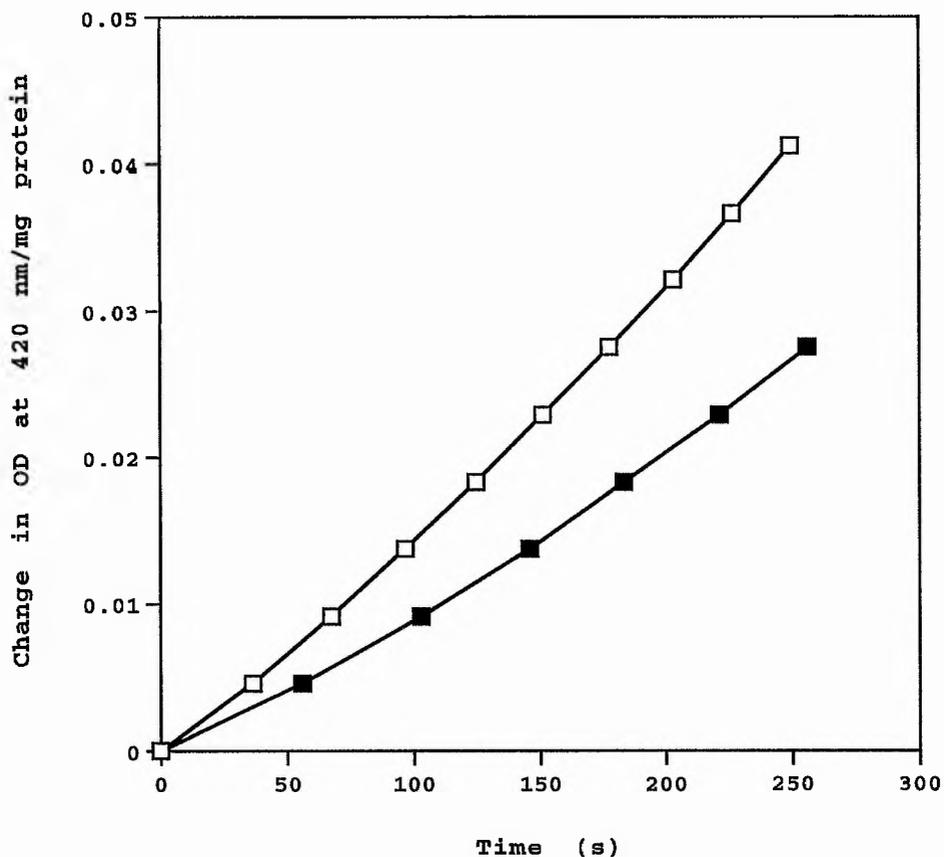
*$\beta$ -galactosidase activity of crude extracts from Escherichia coli strain JM15*

Prior to assaying serine acetyltransferase activity in strain JM15 and complemented strains, it was necessary to establish a simple and effective method for obtaining crude cell-free extracts exhibiting enzyme activity.  $\beta$ -galactosidase (EC 3.2.1.23) was used as a marker enzyme because activity can be

easily measured (Miller 1972; Dobrogosz 1981). Crude cell-free extracts from strain JM15 were obtained by grinding frozen cells with alumina in a mortar and pestle (Chapter 2.17.1), and  $\beta$ -galactosidase activity was measured by following the increase in absorbance at 420 nm that occurs when the artificial substrate o-nitrophenol- $\beta$ -D-galactose, a colourless compound, is converted by the enzyme to a yellow product, o-nitrophenol, and galactose. Crude extracts from strain JM15 were found to have considerable  $\beta$ -galactosidase activity (Figure 3.2). At the lowest amount of extract assayed (24.9  $\mu$ g total protein in 5  $\mu$ l extract), activity was in excess of levels recommended for obtaining linearity (Miller 1972). Figure 3.2 demonstrates that  $\beta$ -galactosidase activity was achievable when strain JM15 cells were disrupted by manual grinding.

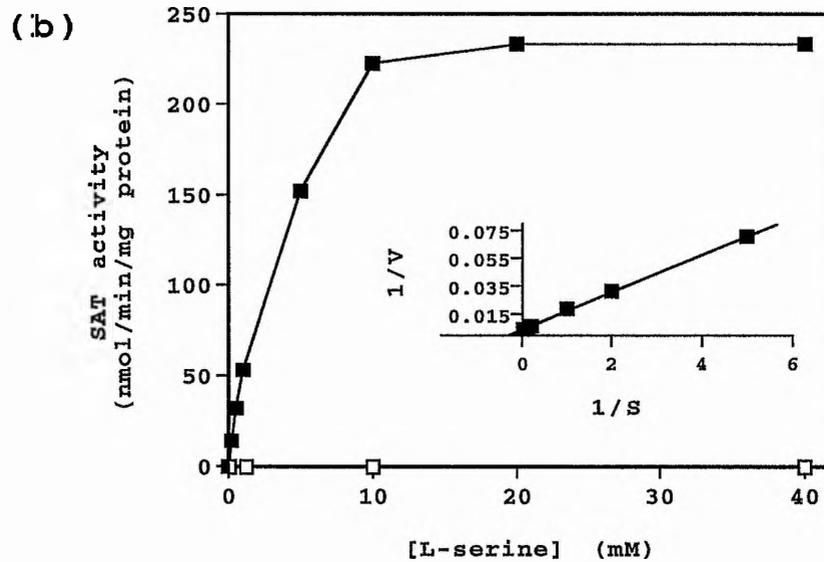
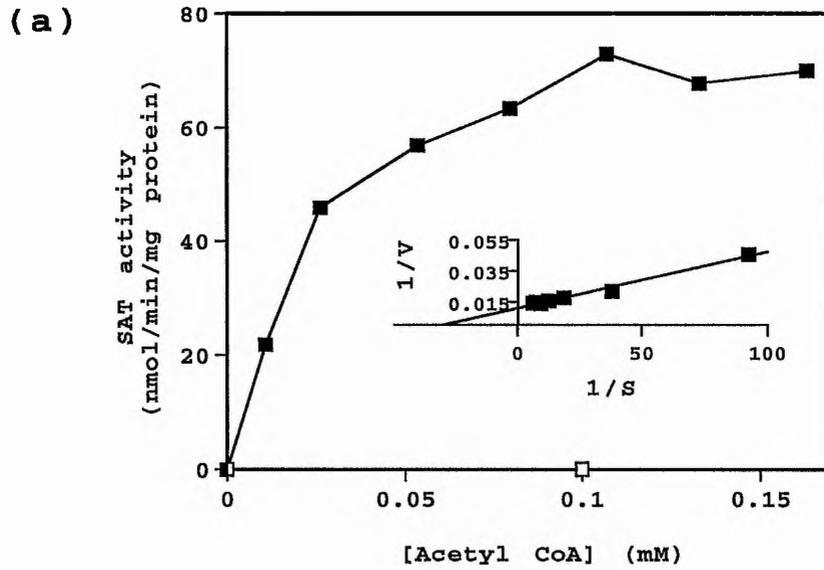
*Serine acetyltransferase activity in complemented, mutant and wild-type Escherichia coli strains*

Plasmid pSAT1 was able to transform the *E. coli cyse* strain JM15 to prototrophy (Table 3.3), which suggests that pSAT1 either confers serine acetyltransferase activity, or expresses a polypeptide which can bypass this step and allow growth of complemented JM15 strains on sulphate as sole sulphur source via another pathway. In order to examine these possibilities, the complemented strain SAT1 and strain JM15 were assayed for serine acetyltransferase activity by following the L-serine-dependent absorbance decrease at 232 nm which occurs upon cleavage of the thioester bond of acetyl CoA (Kredich & Tomkins 1966). Whereas no activity could be detected in crude cell extracts from strain JM15 (Figure 3.3), activity in strain SAT1 was found to be dependent on the two substrates acetyl CoA (Figure 3.3a) and L-serine (Figure 3.3b). Double reciprocal plots of substrate versus activity yielded Michaelis-Menten apparent  $K_m$  values of 0.043 mM for acetyl CoA and 3.47 mM for L-serine (Figure 3.3a and b, insets).



**Figure 3.2**  $\beta$ -galactosidase activity in crude cell-free extracts of *Escherichia coli* strain JM15

5  $\mu$ l (■) or 10  $\mu$ l (□) of crude protein extract from *E. coli* strain JM15 was assayed for  $\beta$ -galactosidase activity by following the accumulation at 420 nm of *o*-nitrophenol. Results are given as the change in OD<sub>420</sub> per mg protein per s. The data are means of triplicate assays, and representative of two separate experiments.



**Figure 3.3** Dependence of serine acetyltransferase activity in strain SAT1 extracts on acetyl CoA and L-serine concentration

Crude protein extracts from the *Escherichia coli* *cysE* mutant strain JM15 (□) and the complemented strain SAT1 carrying the clone *Sat-1* (■) were assayed for serine acetyltransferase (SAT) activity at different concentrations of the enzyme's substrates acetyl CoA (a) and L-serine (b). Insets show Lineweaver-Burk plots of the same data for each substrate. The data are means of triplicate assays, and representative of four separate experiments.

Serine acetyltransferase activity was also examined in crude extracts from another complemented strain, SAT2, and a wild-type sulphate assimilator, strain LE392 (Table 3.4). Levels of activity in strain SAT2 extracts were comparable to those found for strain SAT1 extracts. Serine acetyltransferase activity was slightly lower when strain SAT2 cultures were grown in the absence of the *lac* promoter inducer IPTG, but this difference was not substantial. Serine acetyltransferase activity in the complemented strains SAT1 and SAT2 (in the latter case with or without IPTG) was approximately twenty-fold higher than in the wild-type strain LE392. Ruffet et al. (1995) subsequently reported similar relative serine acetyltransferase specific activities from desalted crude extracts of the *cysE* mutant JM39 (no activity), the  $\lambda$ YES-complemented strains JM39pYES1-9 (activities ranging from 16-194 [average 74.4] nmol/min/mg protein) and the wild-type strain JM105 (activity: 6 nmol/min/mg protein). Correspondence of those results to the results in Table 3.4 is notable because a different enzyme assay which involved measuring production of *O*-acetylserine with high-performance liquid chromatography was employed by Ruffet et al. (1995). Together, the studies indicate that *E. coli cysE* mutant strains complemented by pYES-derived *A. thaliana* clones typically have higher levels of expression of serine acetyltransferase than wild-type *E. coli* strains.

Expression of genes cloned into the pYES vector is reported to be regulated tightly through the *lac* promoter in *E. coli* (Elledge et al. 1991). However, an *E. coli* purine biosynthetic pathway *purD* mutant (glycinamide ribonucleotide synthetase deficient) complemented with the  $\lambda$ YES cDNA library was found to exhibit substantial glycinamide ribonucleotide synthetase activity when grown without IPTG, although activity was increased 2.4-fold when induced by IPTG (Schnorr et al. 1994). In this study, serine acetyltransferase activity in SAT2 was found to be increased only marginally by induction of the *lac* promoter by IPTG (Table 3.4). Expression

---

**Table 3.4** Serine acetyltransferase activity of *Escherichia coli* strains SAT1, SAT2, JM15 and LE392 (wild-type)

<u>Strain</u>	<u>SAT activity (nmol/min/mg protein)</u>	
SAT1	92.43±29.03	(n=12)
SAT2	101.9±26.09	(n=11)
SAT2 (-IPTG)	77.88±19.41	(n=11)
JM15 (+ or -IPTG)	0.000±0.000	(n=11)
LE392 (-IPTG)	4.778±3.330	(n=11)

Serine acetyltransferase (SAT) activity was assayed at substrate concentrations of 100  $\mu$ M acetyl CoA and 1 mM L-serine. Except where stated, cultures were grown in the presence of 1 mM IPTG which induces expression from the *lac* promoter. Results are given as mean  $\pm$  standard error and were obtained from at least two independent experiments.

---

of plasmid cDNA inserts cloned in opposite orientation to the *lac* promoter is feasible in *E. coli* (Senecoff & Meagher 1993), and might provide an explanation for lack of induction by IPTG here. However, subsequent sequence analysis showed that the pSAT2 cDNA insert was expressed from the *lac* promoter (Section 3.2.3.1). Therefore, transcriptional and translational control of plasmid inserts can become independent of plasmidic regulatory genes, as noted by Senecoff and Meagher (1993).

The kinetic constant values obtained using crude bacterial extracts of SAT1 (Figure 3.3a & b, insets) are the same order of magnitude as the range of published apparent  $K_m$  values for plant serine acetyltransferase: 0.09 - 0.83 mM for acetyl CoA and 0.03 - 9.1 mM for L-serine (Smith & Thompson 1971; Ngo & Shargool 1974; Brunold & Suter 1982; Nakamura et al. 1987, 1988; Nakamura & Tamura 1990; Ruffet et al. 1994). Several of those values were derived from assays in which the purified or partially purified serine acetyltransferase was associated with O-acetylserine (thiol)-lyase (Smith & Thompson 1971; Nakamura et al. 1987, 1988; Nakamura & Tamura 1990; Droux et al. 1992). In the supramolecular complex of serine acetyltransferase and O-acetylserine (thiol)-lyase formed in enterobacteria and higher plants, kinetic properties of the individual component enzymes are modified (Kredich et al. 1969; Saito et al. 1995; Chapter 1.2.4). Kinetic data from separated serine acetyltransferase is therefore not likely to be physiologically relevant. It is not known whether the putative *A. thaliana* serine acetyltransferase polypeptide expressed in strain SAT1 and *E. coli* O-acetylserine (thiol)-lyase form a complex *in vivo*.

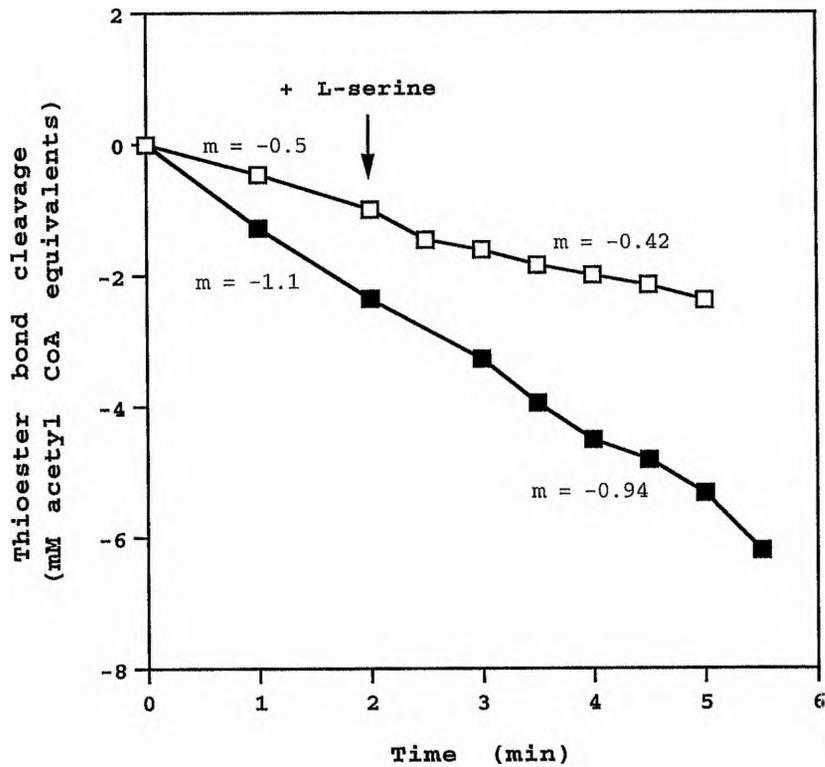
#### *Assay for serine acetyltransferase activity in Arabidopsis thaliana extracts*

Experiments were performed to characterise serine acetyltransferase activity from crude *A. thaliana* cell-free

extracts so that typical expression levels might be compared to those in *E. coli* (Figure 3.3). However, these assays were unsuccessful because the plant extracts contained very high background levels of thioester bond cleavage (Figure 3.4). Addition of L-serine to the crude extracts therefore had little observable effect on decrease in absorbance at 232 nm. Absorbance decrease at 232 nm without exogenous L-serine was equivalent to a rate of about 41 nmol/min/mg protein, but it was not ascertained whether the reaction was chemical or enzymatic. The spectrophotometric serine acetyltransferase assay, which defines activity as L-serine-dependent cleavage of the thioester bond of acetyl CoA observable as a decrease in absorbance at 232 nm, was thus not suitable for determining enzyme activity in *A. thaliana* extracts in this study.

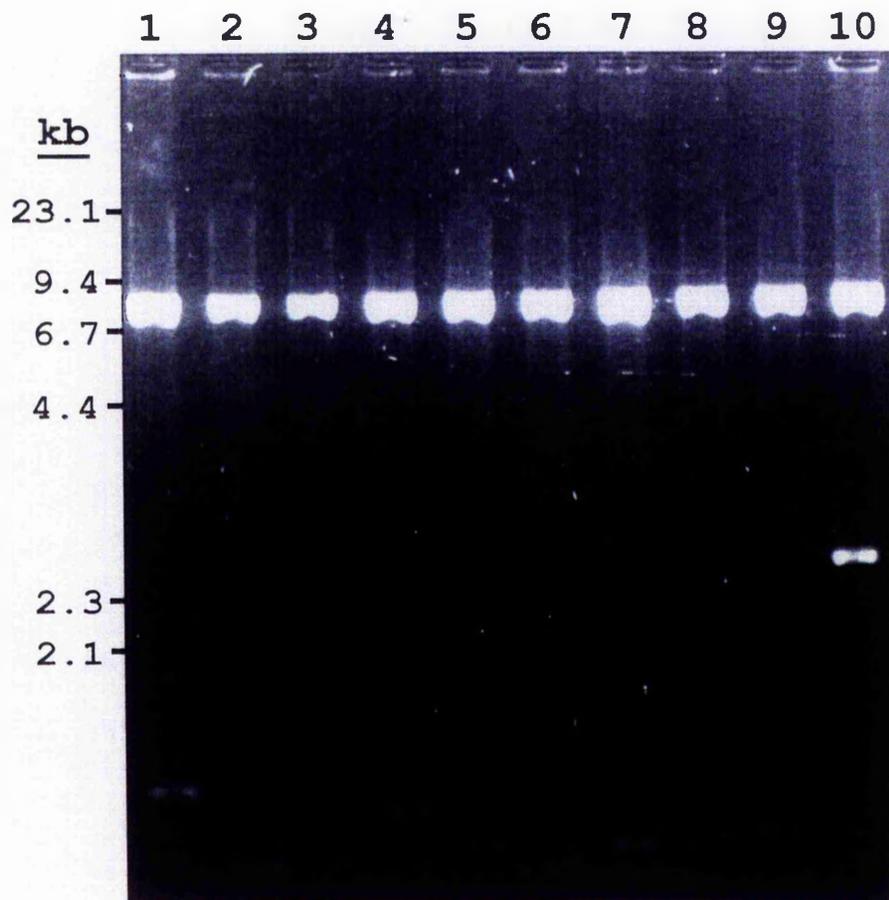
### 3.2.3 Analysis of Sat clones and cDNA sequence determination

Plasmids were extracted from ten complemented SAT strains and the cDNA clones contained within the plasmids were excised by digestion with the restriction enzyme *EcoRI*, for which there are two recognition sequences flanking the *XhoI* cDNA insertion site of  $\lambda$ YES (Figure 2.1). The size of these ten clones was estimated by fractionation through an agarose gel (Figure 3.5). Clones ranged from 1.0 kb (Sat-2 & Sat-40, lanes 2 & 6) to 2.6 kb (Sat-80, lane 10), but the majority of clones were 1.0 - 1.3 kb. One of the shortest clones, Sat-2, was selected initially for DNA sequencing because there was a greater possibility of obtaining sequence in a coding region that could provide further evidence for the identity of the complementing clone.



**Figure 3.4** Serine acetyltransferase assay using crude extracts of *Arabidopsis thaliana*

Thioester bond cleavage, quantified by the decrease in absorbance at 232 nm and converted to acetyl CoA equivalents, was assayed in 10 μl (□) or 20 μl (■) crude extracts from above-ground parts of *A. thaliana* plants (1 μl extract contained 1.134 μg protein). An arrow indicates the point at which 1 mM L-serine was added to the assay mixture, while assay curve slopes (m) before and after addition of L-serine are given for the two amounts of extract used. The data are means of triplicate assays from a single experiment.



**Figure 3.5** Electrophoretic analysis of pYES-derived plasmids extracted from complemented *Escherichia coli* JM15 strains

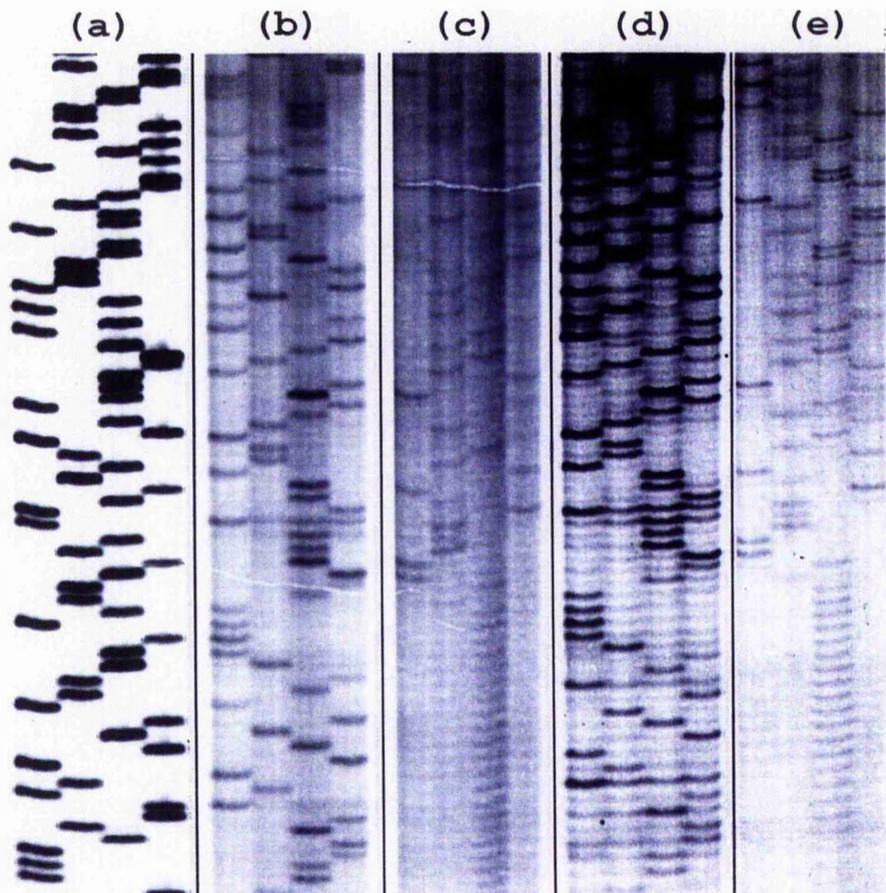
The pYES-derived plasmids pSAT1 (lane 1), pSAT2 (lane 2), pSAT10 (lane 3), pSAT20 (lane 4), pSAT30 (lane 5), pSAT40 (lane 6), pSAT50 (lane 7), pSAT60 (lane 8), pSAT70 (lane 9) and pSAT80 (lane 10), extracted from corresponding complemented *E. coli* JM15 strains, were digested with the restriction enzyme *EcoRI* and fractionated through a 1.2 % ( $v/v$ ) agarose gel. In each lane, linearised empty vector pYES (7.8 kb) and a cDNA insert can be visualised. Band positions of  $\lambda$ /*HindIII*F molecular weight markers are shown.

### 3.2.3.1 Analysis of clone Sat-2

#### *Method development for DNA sequencing of clone Sat-2*

Initial sequencing reactions with Sequenase Version 2.0 DNA polymerase (Chapter 2.8.3) using plasmid pSAT2 purified by the Promega Magic Minipreps system (Chapter 2.5.2) did not yield clear sequence information (Figure 3.6b & c). Use of single-stranded control DNA from the bacteriophage M13mp18 confirmed that the sequencing reactions *per se* were functioning well (Figure 3.6a), so it was concluded that the plasmid DNA template was of unsatisfactory quality. In general, plasmid DNA template does not yield as consistent or reliable results as single stranded DNA (Sambrook et al. 1989). Optimal results using plasmid DNA have been reported using caesium chloride gradient purified plasmids, or supercoiled plasmids uncontaminated by RNA or linear and open circular plasmid DNA (Chen & Seeburg 1985; Hattori & Sakaki 1986). Treatment of the purified pSAT2 with 100  $\mu\text{g/ml}$  RNase for 1 h at 37°C improved background levels of the sequencing reactions (Figure 3.6d & e), but the results suggested that template DNA was still impure. Background contamination was reduced, but not completely eliminated, using plasmid template prepared by Qiagen Tip-100, QIAprep-spin miniprep or Midi plasmid purification kits (results not shown; Chapter 2.5.2). Qiagen purification kits were used subsequently for preparation of plasmid DNA for sequencing.

A further problem observed during initial sequencing reactions with pSAT2 was the presence of bands across all four dideoxynucleotide lanes in specific regions. Troubleshooting using the Sequenase Version 2.0 Protocol guidebook (United States Biochemical, 1993) eliminated insufficient mixing of reagents, prolonged labelling, labelling above room temperature, and cold or prolonged termination as possible causes of the banding. Using different amounts of template and primer in the sequencing



**Figure 3.6** DNA sequence obtained from pSAT2 purified by Promega Magic Minipreps and single-stranded control DNA

Sequencing reactions using a Sequenase Version 2.0 kit were performed using the following templates and primers: (a) 1  $\mu\text{g}$  M13mp18 single-stranded control DNA, and 0.5 pmol 17-mer forward primer (both supplied in the Sequenase Version 2.0 kit); (b) 10  $\mu\text{g}$  pSAT2 purified by Promega Magic Minipreps, and 1 pmol primer P2 (forward primer derived from pYES vector sequence); (c) 10  $\mu\text{g}$  pSAT2 purified by Magic Minipreps, and 1 pmol primer P1 (reverse primer derived from pYES vector sequence); (d) 10  $\mu\text{g}$  Magic Miniprep-purified pSAT2 treated with 100  $\mu\text{g}/\text{ml}$  RNase for 1 h at 37°C, and 1 pmol primer P2; and (e) 10  $\mu\text{g}$  Magic Miniprep-purified pSAT2 treated with 100  $\mu\text{g}/\text{ml}$  RNase for 1 h at 37°C, and 1 pmol primer P1. Autoradiographs of polyacrylamide gels in which sequence lanes were loaded in the order G, A, T, C are shown.

reaction showed that less banding was evident at the lowest amounts used, ie. 10  $\mu$ g pSAT2 and 1 pmol primer (results not shown). Use of smaller amounts of plasmid DNA produced faint bands (not shown), and the large size of the pYES plasmid (8 kb), and hence small insert to vector ratio, thus necessitated the relatively large amount of template in the sequencing reaction. Banding was likely to be indicative of Sequenase 2.0 "pause" sites caused by dissociation of enzyme from the DNA template at strong secondary structures. As recommended by the manufacturer, duration of labelling reactions with Sequenase 2.0 was reduced to 3.5 - 4 min to minimise banding. The nucleotide analogue dITP can be used in place of dGTP to eliminate artifacts such as compressions that result from secondary structure formation during gel electrophoresis (Mills & Kramer 1979). However, Sequenase 2.0 pausing at sites of template secondary structure is exacerbated when using dITP (Sequenase Version 2.0 Protocol guidebook, United States Biochemical, 1993). Use of dITP in sequencing reactions with pSAT2 as template caused greater banding across the four reaction lanes, but no difference in readable sequence (results not shown), confirming that banding was caused by regions of strong secondary structure in the template. Strategies employed to obtain unambiguous sequence information in regions of strong secondary DNA structure within clone Sat-1 are discussed in Section 3.2.3.2.

Two further modifications of the basic sequencing protocol were found to be useful (data not shown): firstly, the use of Sequenase Extension Mix, which increases the concentration of deoxynucleotides in the termination reaction, maximised the sequence information that could be determined from a single gel run (up to 470 bp from the primer, with an average of about 300 bp). Resolution was often limited by background levels of radioactivity, probably caused by impurities in plasmid preparations; and secondly, pyrophosphatase was routinely included in sequencing reactions, as recommended by

the suppliers of Sequenase 2.0 enzyme, to prevent pyrophosphorolysis that causes some bands to appear fainter (Tabor & Richardson 1990).

#### *Characterisation of clone Sat-2 and sequence comparisons*

The complete DNA sequence of the 990 bp clone Sat-2 was obtained by overlapping single-stranded sequence information from the 5' and 3' ends of the clone. The longest open reading frame encoded by Sat-2 was a 290 residue polypeptide (SAT-2), including 9 amino acid residues at the N-terminus encoded by the pYES cloning site (Figure 3.7). The SAT-2 fusion protein was initiated at the pYES AUG translation codon, while the single methionine encoded by plant cDNA is at residue 276, only 15 residues from the C-terminus of SAT-2 (Figure 3.7). Clone Sat-2 is thus truncated at the 5' end. Fusion protein-encoding cDNA clones derived from the  $\lambda$ YES cDNA library have been reported previously (Pang et al. 1992). The 144 bp 3' untranslated region of Sat-2 contains a GT-rich tract (35 out of 38 nucleotides from 923 to 960 in Figure 3.7) and a 31-nucleotide poly(A) tail at the extreme 3' end. Plant AAUAAA-like polyadenylation signals (Hunt 1994; Wu et al. 1995) could not be detected, suggesting that the GT-rich sequence probably determined the polyadenylation site (at nucleotide 1031 in Figure 3.7).

A BLASTP search for sequence similarity (Altschul et al. 1990) was conducted with the Genetics Computer Group (GCG) software package (Devereux et al. 1984) using the 281 residue plant cDNA-encoded polypeptide portion of SAT-2 (called "ATHSAT-2"). When the search was done at the end of 1993, six serine acetyltransferase sequences, all from prokaryotic organisms, were selected from protein and translated nucleic acid databases due to significant homology to plant ATHSAT-2. Pairwise comparisons using the GCG Gap program showed that ATHSAT-2 had 50.2 % identity (69.1 % homology) with CysE (serine acetyltransferase) from *E. coli* (Denk & Böck 1987;

```

1  TTTCACACAGGAAACAGGACTCTAGAGGATCTGGAGGAAAAAATTATGAATTCCTCGAGCTACGTCAGGGGTCCCTTTGCT
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   aaagtgtctcctttgtcctgagatctcctagacctcctttttaaacttaaggagctcgatgcagctcccaggaaacga
                                     M N S S S Y V R G P L L
81  TGAAGATCTCGATCGCGACGCTGAAGTCGATGATGTTTGGGCCAAAATCCGAGAAGAGGCTAAATCTGATATCGCCAAG
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   acttctagagctagcgtcgacttcagctactacaacccgggttttaggctcttctccgatttagactatagcggtttc
   E D L D R D A E V D D V W A K I R E E A K S D I A K E
161 AACCTATTGTTTCCGCTTATTATCACGCTTCGATTGTTTCTCAGCGTTCGTTGGAAGCTGCGTTGGCGAATACTTTATCT
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   ttgataacaaaggcgaataatagtgcaagctaacaaagagtcgcaagcaaccttcgacgcaaccggttatgaaataga
   F I V S A Y Y H A S I V S Q R S L E A A L A N T L S
241 GTTAAACTCAGCAATTTGAATCTTCCAAGCAACACGCTTTTCGATTGTTTCTCTGGTGTCTTCAAGGAAACCCAGATAT
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   caatttgagctgtaaaccttagaagggtcgtttgtagcaaaaagctaaacaagagaccacaagaagttcctttgggtctata
   V K L S N L A N L P S N T L F D L F S G V A L Q G T N P D I
321 TGTGAATCTGTCAAGCTAGATCTTTTAGCTGTTAAGGAGAGAGATCCTGCTTGTATAAGCTACGTTTATTGTTTCTTCTC
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   acaacttagacagttcgatctagaaaatcgacaattcctctctctagacgaacatattcgatgcaagtaacaagggaag
   V E S V K L D L L A V K E R D P A C I S Y V H C F L H
401 ACTTTAAAGGCTTCCCTCGCTTGTCAAGCGCATCGTATTGCTCATGAGCTTTGGACTCAGGACAGAAAATCCTAGCTTTG
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   tgaatttcgaaggagcgaacagttcgcgtagcataacgagtagctcgaacctgagtcctgtcttttaggatcgaaac
   F K G F L A C Q A H R I A H E L W T Q D R K I L A L
481 TTGATCCAGAACAGAGTCTCTGAAGCCTTCGCTGTTGATTTCCACCCTGGAGCTAAAATCGGTACCGGGATTGTCGTAGA
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   aactaggtcttctcagagacttcggaagcgaactaaagggtgggacctcgatttagccatggcctaaacagatct
   L I Q N R V S E A F A V D F H P G A K I G T G I L L D
561 CCATGCTACGGCTATTGTGATCGGTGAGACGGCGGTTGTGGGGAACAATGTTTCGATTCTCCATAACGTTACGCTTGGAG
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   ggtacgatgccgataaactagaccactctgcccgaacaccccttgttacaagctaagaggtattgcaatgcaaacctc
   H A T A I V I G E T A V V G N N V S I L H N V T L G G
641 GAACGGGGAAACAGTGTGGAGATAGGCACCCGAGATTGGCGATGGGGTTTTGATTGGAGCTGGGACTgtattttgggg
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   cttgcccctttGTACACCTCTATCCGTGGGCTTCTAACCGCTACCCCAAACCTAACCTCGACCTGAACATAAAACCCC
   T G K Q C G D R H P K I G D G V L I G A G T C I L G
721 aatatcacgattggtgaaggagcgaagattggtgccccggtcggtggttgaagacgtgccgccccgctacgacggctgt
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   TTATAGTGTAAACCACTTCTCGATTCTAACACGCCCCAGCCACCACAACCTTCTGCACGGCGCGCATGCTGCCGACA
   N I T I G E G A K I G A G S V V L K D V P P R T T A V
801 tggaaatccggcgaggttgcttggtggaataatccgaaaaacgcatgacaagattcctggtttgactatggaccaga
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   ACCTTTAGCCGCTCCAACGAACCACATTCTATTAGGCTTTTGGCTACTGTTCTAAGGACCAACTGATACCTGGTCT
   G N P A R L L G G K D N P K T H D K I P G L T M D Q T
881 cgctgcatatataccgagtggtcggtatgtaattgaaaaagtccttgttttgggtgtgtgtgtttatggctttt
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   GCAGCGTATATAGGCTCACCGCTAATACATTAAACTTTTTCAGAAACAAAACCAACAAACAAACAAATACCGAAAA
   S H I S E W S D Y V I *
961 cattgtctctcggcttctctgttattgaaagctggtggaagtatatgatatgcgatgatgctgtttccataaaaaaaaa
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   GTAACAAGAGAGCCGAAGAAGACAATAACTTCGACCACATCATATACTATACGCTACTACGACAAAGGTATTTTTTTTTT
1041 aaaaaaaaaaaaaaaaaaaaaa
   -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
   TTTTTTTTTTTTTTTTTTTTTT

```

**Figure 3.7** Nucleotide sequence and deduced amino acid sequence of clone Sat-2

The complete nucleotide sequence of *Arabidopsis thaliana* cDNA clone Sat-2 (990 bp) and the 5' flanking section of the pYES vector sequence (71 bp, underlined) are given. DNA sequence obtained from sequencing the 5' and 3' ends of the clone, with an overlap of 50 nucleotides, is indicated in capitals, while unsequenced complementary nucleotides are given in small letters. The longest open reading frame of 290 amino acid residues, including nine residues encoded by pYES vector DNA (double underlined), is shown. The stop codon is indicated by an asterisk.

Tei et al. 1990a). Using default parameters of BLASTP, the probability that the sequence alignment high score of 648 between ATHSAT-2 and *E. coli* CysE arose by chance in a similar size search was  $4.9 \times 10^{-85}$ . Whilst functional homology, but not sequence homology, between the bacterial and heterologous enzyme is essential for complementation (Snustad et al. 1988; Pang et al. 1992), the high level of sequence conservation observed between ATHSAT-2 and *E. coli* CysE provides strong evidence for the identity of the plant protein. ATHSAT-2 also had 48.7 % identity (68.8 % homology) with CysE from *Salmonella typhimurium* (Sivaprasad, A.V., Kuczek, E.S., Bawden, C.S., Rogers, G.E. 1991, unpublished SWISS-PROT entry P29847), 43.6 % identity (62.5 % homology) with *Buchnera aphidicola* CysE (Lai & Baumann 1992), 40.5 % identity (61.0 % homology) with CysE from *Bacillus subtilis* (Gagnon et al. 1984), 39.4 % identity (57.6 % homology) with NIFP (serine acetyltransferase) from *Azotobacter chroococcum* (Evans et al. 1991), and 38.3 % identity (57.7 % homology) with the *nif* gene cluster ORF7 (subsequently confirmed by sequence similarity to be a NIFP-like serine acetyltransferase - data not shown) from *Azotobacter vinelandii* (Jacobson et al. 1989). These sequences were aligned using the GCG PileUp program (Figure 3.8). ATHSAT-2 was found to be longer at the N-terminus compared with the prokaryotic serine acetyltransferases, while a consensus sequence common to all these enzymes showed highly conserved regions in the C-terminus half of ATHSAT-2 (Figure 3.8). Of the 54 residues conserved between ATHSAT-2 and the six serine acetyltransferases, a high proportion of leucine/isoleucine (29.6 %) and glycine (20.4 %) residues was noted. The functional significance of these residues is discussed in Section 3.2.3.3.

Preliminary analysis of ATHSAT-2 thus provided strong evidence that Sat-2 encodes an *Arabidopsis thaliana* serine acetyltransferase with extensive homology to several prokaryotic serine acetyltransferases. However, clone Sat-2

	1					60
A. chroococcum NIFP	.....	.....	.....	.....	.....	.....
A. vinelandii ORF7	.....	.....	.....	.....	.....	.....
B. subtilis CysE	.....	.....	.....	.....	.....	.....
E. coli CysE	.....msce	eleivwnnik	aeartladce	pmlasfyhat	llkhenlgsa	lsymlankls
S. typhimurium CysE	.....mpce	eleivwnnik	aeasaladce	pmlasfyhat	llkhenlgsa	lsymlankla
B. aphidicola CysE	.....mcsl	eelwlnmik	hkaqkilkke	pilsnfygks	ilnhkklshs	lscilsdkls
ATHSAT-2	piledlrdra	evddvkwakir	eeaksdiake	pivsayyhas	ivsrgrleaa	lantlsvklis
Consensus	-----	-----	-----	-----	-----	-----
	61					120
A. chroococcum NIFP	.....	.....	mlllaqwreD	ircvferDPA	arttfevltt	YpGvhAimly
A. vinelandii ORF7	.....	.....	mllltqwrD	ircvferDPA	arttfevltt	YpGvhAigly
B. subtilis CysE	.....	.....	.mffrmlkeD	idtvfdqDPA	arsyfevilt	YsG1hAiwah
E. coli CysE	spimpalair	evveeayaad	pemiasaacD	iqavrtrDPA	vdkystpily	LkGfhAlqay
S. typhimurium CysE	spimpalair	evveeayaad	pemiasaacD	iqavrtrDPA	vdkystpily	LkGfhAlqay
B. aphidicola CysE	tsmisekdiy	nifnkiyann	isiinsvkdD	ikaasqrDPv	vkhytltpily	LkGfhAleay
ATHSAT-2	nlnlpstlrf	dlfsgvlgqn	pdivesvklD	llavkerDPA	cisyvhcflh	FKGflAcqah
Consensus	-----	-----	-----	-----	-----	-----
	121					180
A. chroococcum NIFP	RlahrLWrpn	alprpaavvr	.arlvsnVD	iHPgAvIGar	fFiDHgacVV	IGEtaeigrd
A. vinelandii ORF7	RlanrLWra	awryparlls	fvarmlsnVD	iHPgAtIGer	fFiDHgacVV	IGEtaeIgrd
B. subtilis CysE	RiahaLYkrk	.fyflarlis	qvsrfftgIE	iHPgAtIGrr	fFiDHgmVv	IGEtaeIgrn
E. coli CysE	RighwLWnqg	rrolaiflqn	qvsvtf.qVD	iHPaAkIGrg	iMLDHatgIV	VGEtavIend
S. typhimurium CysE	RighwLWnkg	rrolaiflqn	qvsvsf.qVD	iHPaAkIGrg	iMLDHatgIV	VGEtavIedd
B. aphidicola CysE	RlshyLWnik	ryelsaylqs	ristvf.sVD	iHPaAsIGsg	iMLDHatgIV	IGEGviIend
ATHSAT-2	RiaheLWtqd	rkilalliqn	rvseaf.avD	fHPgAkIGtg	iLlDHataIV	IGEtaVggnn
Consensus	R----LW---	-----	-----	-HP-A-IG--	---DH---IV	IGE---I---
	181					240
A. chroococcum NIFP	VtLYhgVTLG	GTtgakGk.R	HptlqdvvlV	gaGakILGpI	tIGAnarVGA	nSVVvqdvPe
A. vinelandii ORF7	VtLYhgVTLG	GTswnkGk.R	HptlrdgvlV	gaGakILGpI	tVGagarVGA	nSVVvqdvPd
B. subtilis CysE	VtvFqgVTLG	GTgkekGk.R	HptikddalI	atGakVLGSI	tVGegskIGA	gSVVlhdvPd
E. coli CysE	VsILqsVTLG	GTgksG.dR	HPkiregvmI	gaGakILGnI	eVGrgakIGA	gSVVlqpVpp
S. typhimurium CysE	VsILqsVTLG	GTgktsG.dR	HPkiregvmI	gaGakILGnI	eVGrgakIGA	gSVVlqpVpp
B. aphidicola CysE	VsIFhsVTLG	GTgsntGknR	HPiirknvtI	gaGakILGnI	eVGgqvkvGA	gSIVlknIPp
ATHSAT-2	VsILhnVTLG	GTgkqCG.dR	HPkigdgvlI	gaGacILGnI	tIGegakIGA	gSVVlkdVpp
Consensus	V--L--VTLG	GT----G--R	HP-----I	--G--ILG-I	-VG----IGA	-SVV---VP-
	241					300
A. chroococcum NIFP	gcTvvGiPpk	vvklreaggl	npygidldhh	lipdpvgkai	aclleridsl	ekrveagglv
A. vinelandii ORF7	gcTvvGiPpk	vvkvreagrp	npygidldhh	lipdpvgkai	aclleridsl	ekrveagglv
B. subtilis CysE	fstTvvGiPgr	vv.vqngkkv	rr...dlnhq	dldpvpadr	ksleqqilel	kaeledrker
E. coli CysE	htTaaGvPar	ivgk.pdsdk	psmdmdqghfn	ginhtfeygd	gi.....	.....
S. typhimurium CysE	htTaaGvPar	ivgk.pgsdk	psmdmdqghfn	ginhtfeygd	gi.....	.....
B. aphidicola CysE	fvTvvGvPak	iikkiknsnk	nlfqkek..	.....	.....	.....
ATHSAT-2	rtTavGnPar	llggkdnpkt	hdkipgltd	qtshisewsd	yvi.....	.....
Consensus	--T--G-P--	-----	-----	-----	-----	-----
	301					352
A. chroococcum NIFP	aaaasstfye	gnpndnsice	tnlrrsapws	sgrrrppaha	gdrvsgrakg	sd
A. vinelandii ORF7	tvgsqqqfyg	vcnsdnsice	....sdcag	gttaqaqqsa	grrravptpv	ae
B. subtilis CysE	inqk.....	.....	.....	.....	.....	.....
E. coli CysE	.....	.....	.....	.....	.....	.....
S. typhimurium CysE	.....	.....	.....	.....	.....	.....
B. aphidicola CysE	.....	.....	.....	.....	.....	.....
ATHSAT-2	.....	.....	.....	.....	.....	.....
Consensus	-----	-----	-----	-----	-----	-----

**Figure 3.8** Amino acid sequence alignment of ATHSAT-2 with prokaryotic serine acetyltransferases

ATHSAT-2, the 281 residue polypeptide encoded by *Arabidopsis thaliana* clone Sat-2, was aligned with prokaryotic serine acetyltransferases NIFP from *Azotobacter chroococcum* (Evans et al. 1991; SWISS-PROT entry P23145), ORF7 (orthologous to NIFP) from *Azotobacter vinelandii* (Jacobson et al. 1989; GenBank translation PID:g142363 from M20568), and CysE from *Bacillus subtilis* (Gagnon et al. 1994; SWISS-PROT entry Q06750), *Escherichia coli* (Denk & Böck 1987; Tei et al. 1990a; SWISS-PROT entry P05796), *Salmonella typhimurium* (Sivaprasad, A.V., Kuczek, E.S., Bawden, C.S., Rogers, G.E. 1991; unpublished SWISS-PROT entry P29847) and *Buchnera aphidicola* (Lai & Baumann 1992; SWISS-PROT entry P32003) using the GCG PileUp program. A consensus sequence shows amino acids common to the seven aligned polypeptides.

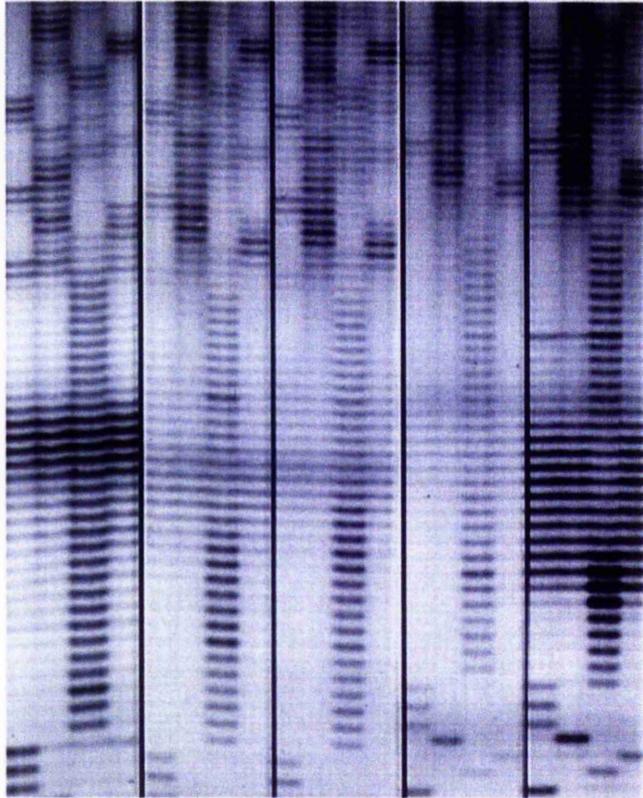
is truncated at the 5' end and ATHSAT-2 could not supply complete information on the plant enzyme. Therefore, attempts were made to obtain sequence from a full-length complementing clone, commencing with the 1.3 kb clone Sat-1 (Figure 3.5).

### 3.2.3.2 Analysis of clone Sat-1

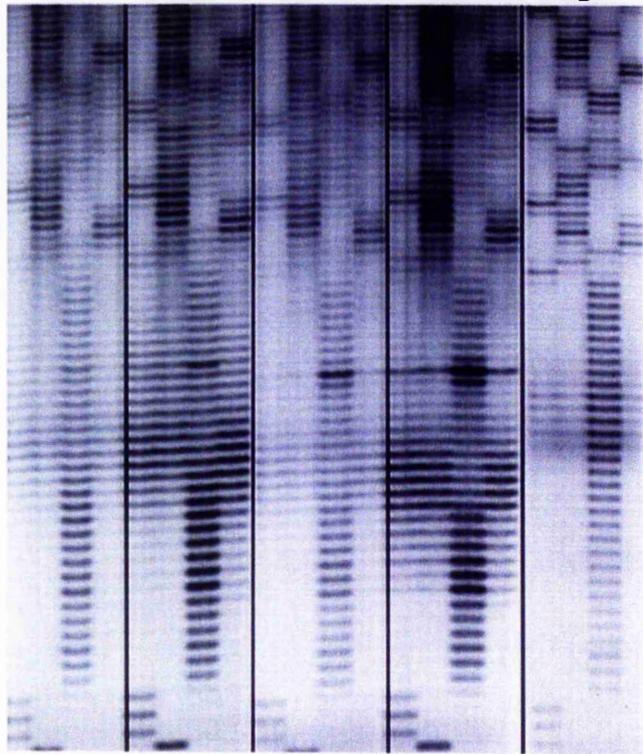
#### *Method development for DNA sequencing of clone Sat-1*

Both strands of the 1 278 bp clone Sat-1 contained within pSAT1 were sequenced completely. Using standard conditions optimised for clone Sat-2, sequence obtained from the vector-based primer P1 (5'-ACT TTA ACG TCA AGG AG), which reads the reverse strand from the 3' end of Sat-1, was difficult to interpret accurately due to the presence of bands across more than one tract at several positions but especially in the region of a poly(T) tract (Figure 3.9a). Fresh stocks of the primer and higher annealing temperatures were tested to eliminate primer contamination or denaturation and mispriming at a secondary site, respectively (not shown), leaving strong secondary structure of the template as a likely cause of the banding (Sequenase Version 2.0 Protocol guidebook, United States Biochemical, 1993). DNA sequencing with the TaqTrack Sequencing System (Promega Corporation, UK), which uses Sequencing Grade Taq DNA polymerase at an elevated reaction temperature of 70°C where secondary structure of DNA templates is decreased (Innis et al. 1988), was therefore attempted. Parameters varied with the TaqTrack Sequencing System included: (i) amount of template (Figure 3.9b and c); (ii) amount of primer (Figure 3.9d and e); (iii) addition of extra Taq DNA polymerase at two different primer concentrations (Figure 3.9f and g); and (iv) performing a chase step to eliminate false bands, as described in Chapter 2.8.4 (Figure 3.9h and i). None of the sequencing reactions performed with the TaqTrack Sequencing system were effective at producing clear unambiguous results. Finally, best

(a) (b) (c) (d) (e)



(f) (g) (h) (i) (j)



sequence results were obtained using a greater amount of Sequenase Version 2.0 (6.5 units per reaction, compared with the standard 3.25 units) (Figure 3.9j). Additional Sequenase probably allowed sufficient enzyme to continue through regions of strong secondary structure so as to reduce false banding through pausing (Sequenase Version 2.0 Protocol guidebook, United States Biochemical, 1993). Sequence data at the 3' end of Sat-1 obtained from the reverse strand was verified using the complementary strand sequence. Increasing Sequenase concentration to resolve strong secondary DNA structure was effective also in other regions of the Sat-1 clone sequence (not shown).

---

◀ **Figure 3.9** Optimisation of DNA sequencing protocols for sequencing the 5' end of the noncoding strand of clone Sat-1

The 5' end of the reverse strand of *Arabidopsis thaliana* clone Sat-1 was sequenced initially under standard conditions using a Sequenase Version 2.0 kit with 1 pmol of the pYES-derived primer P1, 5 µg pSAT1 and 3.25 units Sequenase Version 2.0 T7 DNA polymerase (a). Subsequently, a TaqTrack Sequencing System was employed with the following variations: (b) 7 µg pSAT1, 1 pmol P1 and 5 units Sequencing Grade Taq DNA polymerase; (c) 3.5 µg pSAT1, 1 pmol P1 and 5 units Sequencing Grade Taq DNA polymerase; (d) 7 µg pSAT1, 0.5 pmol P1 and 5 units Sequencing Grade Taq DNA polymerase; (e) 7 µg pSAT1, 5 pmol P1 and 5 units Sequencing Grade Taq DNA polymerase; (f) 7 µg pSAT1, 1 pmol P1 and 10 units Sequencing Grade Taq DNA polymerase; (g) 7 µg pSAT1, 5 pmol P1 and 10 units Sequencing Grade Taq DNA polymerase; (h) 7 µg pSAT1, 1 pmol P1 and 10 units Sequencing Grade Taq DNA polymerase, with termination reaction followed by a chase step in which 1 unit Sequencing Grade Taq DNA polymerase was added and the reaction continued for 15 min at 70°C to remove false bands; and (i) 7 µg pSAT1, 5 pmol P1 and 10 units Sequencing Grade Taq DNA polymerase, followed by a chase step as described for (h). Finally, the Sequence Version 2.0 kit was implemented using 3.5 µg pSAT1, 0.5 pmol P1 and 6.5 units Sequenase Version 2.0 T7 DNA polymerase. Autoradiographs of polyacrylamide gels in which sequence lanes were loaded in the order G, A, T, C are shown.

### *Characterisation of clone Sat-1*

The 1278 bp clone Sat-1 extended 111 bp at the 5' end beyond Sat-2 (Figure 3.10). The two clones were identical for 960 nucleotides over the entire Sat-2 sequence upstream from the Sat-2 polyadenylation site, but Sat-1 continued another 176 bp at the 3' end to a 32-residue polyadenylation tail (Figure 3.10). In addition to the GT-rich tract found in both Sat-1 (bp 965 to 1006) and Sat-2, Sat-1 contained another GT-rich tract (27 of 32 nucleotides from 1142 to 1178 bp). Except for GT-richness, the Sat-1 untranslated 3' end did not contain any known far-upstream (eg. TATTTGTA) or near-upstream (AATAAA motifs) elements for polyadenylation (Hunt 1994; Wu et al. 1995). Transcript heterogeneity at the 3' end due to multiple polyadenylation sites is found in most single plant genes, and it is speculated that different sites may be preferred under different environmental conditions (Dean et al. 1986; Hunt 1994). It is not known which transcript of the gene represented by Sat-1 and Sat-2 is more abundant *in vivo*.

The longest open reading frame of Sat-1 codes for a polypeptide of 304 amino acids with a calculated molecular weight of 32.9 kDa (Figure 3.10), which corresponds with previously reported values of 31 - 33 kDa for plant serine acetyltransferase (Nakamura & Tamura 1990; Ruffet et al. 1994). The Sat-1 encoded polypeptide, which is 23 amino acids longer than ATHSAT-2 at the N-terminus (Figure 3.10), was translated in a different reading frame from the pYES vector initiation site and was therefore not a fusion protein. Re-initiation of translation was required for expression of the polypeptide, a frequent phenomenon with the pYES plasmids (eg. Schnorr et al. 1994; Ruffet et al. 1995; Stallmeyer et al. 1995). However, an *A. thaliana* EST clone 89K20T7 (for which the partial sequence was available in GenBank, entry T21091, and which was then obtained from the Arabidopsis Biological Resource Center at Ohio State University [USA] and sequenced manually at the 5' end by Mr E. Campbell in this

```

1
EST 89K20T7      1      gtcacaagtc gcgcaccact cacaatgtcc ctatatatgc tcggttcac tctccacac atcaatcac actccttctc 80
81
EST 89K20T7      81      tcttcctctc ttggtttctc ccaaattcaa acaccatact ttatctctc ctcctctccc tctctctctc cctcctatgg 160
161
EST 89K20T7      161     ctgctgcat cgacacctgc cgcactggta aaccccagat tctctctcgc GATTCTTCTA AACACCACGA CGATGAATCT 240
Sat-1            .....C GATTCTTCTA AACACCACGA CGATGAATCT
241
EST 89K20T7      241     GGCTTTCGTT ACATGAACCTA CTTCCTG.... 320
Sat-1            GGCTTTCGTT ACATGAACCTA CTTCCTGttat cctgatcgat cttccttcaa tggaaaccag accaaaaacc tccatactcg
Sat-1 ORF        M N Y F R Y P D R S S F N G T Q T K T L H T R
321
Sat-2            TCCTTTGCTT GAAGATCTCG ATCGCGACGC TGAAGTCGAT GATGTTTGGG CCAAATCCG AGAAGAGGCT AAATCTGATA 400
Sat-1            TCCTTTGCTT GAAGATCTCG ATCGCGACGC TGAAGTCGAT GATGTTTGGG CCAAATCCG AGAAGAGGCT AAATCTGATA
Sat-1 ORF       P L L E D L D R D A E V D D V W A K I R E E A K S D I
401
Sat-2            TCGCCAAAGA ACCTATTGTT TCCGCTTATT ATCAGCCTTC GATTGTTTCT CAGCGTTCGT TGAAGAGTGC GTTGGCGAAT 480
Sat-1            TCGCCAAAGA ACCTATTGTT TCCGCTTATT ATCAGCCTTC GATTGTTTCT CAGCGTTCGT TGAAGAGTGC GTTGGCGAAT
Sat-1 ORF       A K E P I V S A Y H A S I V S Q R S L E A A L A N
481
Sat-2            ACTTTATCTG TTAACCTCAG CAATTTGAAT CTTCACGCA ACACGCTTTT CGATTGTGTC TCTGGTGTTC TTCAAGGAAA 560
Sat-1            ACTTTATCTG TTAACCTCAG CAATTTGAAT CTTCACGCA ACACGCTTTT CGATTGTGTC TCTGGTGTTC TTCAAGGAAA
Sat-1 ORF       T L S V K L S N L N L P S N T L F D L F S G V L Q G N
561
Sat-2            CCCAGATATT GTTGAATCTG TCAAGCTAGA TCTTTAGCT GTTAAGGAGA GAGATCCTGC TTGTATAAGC TACGTTCAIT 640
Sat-1            CCCAGATATT GTTGAATCTG TCAAGCTAGA TCTTTAGCT GTTAAGGAGA GAGATCCTGC TTGTATAAGC TACGTTCAIT
Sat-1 ORF       P D I V E S V K L D L L A V K E R D P A C I S Y V H C
641
Sat-2            GTTTCCTTCA CTTTAAAGGC TTCTTCGCTT GTCAAGCGCA TCGTATTGCT CATGAGCTTT GACTCAGGA CAGAAAATTC 720
Sat-1            GTTTCCTTCA CTTTAAAGGC TTCTTCGCTT GTCAAGCGCA TCGTATTGCT CATGAGCTTT GACTCAGGA CAGAAAATTC
Sat-1 ORF       F L H F K G F L A C Q A H R I A H E L W T Q D R K I
721
Sat-2            CTAGCTTTGT TGATCCAGAA CAGAGTCTCT GAAGCCTTCG CTGTGATTT CCACCTCGA GCTAAAATCG GTACCAGGAT 800
Sat-1            CTAGCTTTGT TGATCCAGAA CAGAGTCTCT GAAGCCTTCG CTGTGATTT CCACCTCGA GCTAAAATCG GTACCAGGAT
Sat-1 ORF       L A L L I Q N R V S E A F A V D F H P G A K I G T G I
801
Sat-2            TTGCTAGAC CATGCTACCG CTATTGTGAT CGGTGAGAC GCGGTGTGG GGAACATGT TTCGATTCTC CATAACGITA 880
Sat-1            TTGCTAGAC CATGCTACCG CTATTGTGAT CGGTGAGAC GCGGTGTGG GGAACATGT TTCGATTCTC CATAACGITA
Sat-1 ORF       L L D H A T A I V I G E T A V V G N N V S I L H N V T
881
Sat-2            CGCTTGGAGG AACGGGGAAA CAGTGTGGAG ATAGGCACCC GAAGATTGGC GATGGGGTTT TGATTGAGC TGGGACTTGT 960
Sat-1            CGCTTGGAGG AACGGGGAAA CAGTGTGGAG ATAGGCACCC GAAGATTGGC GATGGGGTTT TGATTGAGC TGGGACTTGT
Sat-1 ORF       L G G T G K Q C G D R H P K I G D G V L I G A G T C
961
Sat-2            ATTTTGGGGA ATATCACGAT TGTGAAGGA GCTAAGATG GTGCGGGTTC GGTGGTGTG AAAGACGTGC CGCCGGTAC 1040
Sat-1            ATTTTGGGGA ATATCACGAT TGTGAAGGA GCTAAGATG GTGCGGGTTC GGTGGTGTG AAAGACGTGC CGCCGGTAC
Sat-1 ORF       I L G N I T I G E G A K I G A G S V V L K D V P P R T
1041
Sat-2            GACGGCTGTT GGAATCCGG CGAGGTGCT TGGTGTAAA GATAATCCGA AAACGCATGA CAAGATTCTT GGTGTGACTA 1120
Sat-1            GACGGCTGTT GGAATCCGG CGAGGTGCT TGGTGTAAA GATAATCCGA AAACGCATGA CAAGATTCTT GGTGTGACTA
Sat-1 ORF       T A V G N P A R L L G G K D N P K T H D K I P G L T M
1121
Sat-2            TGGACCAGAC GTCGCATATA TCCGAGTGGT CGGATTATGT AATTGAAA AGCTTTGTT TTTGGTGTG GTTGTGTTT 1200
Sat-1            TGGACCAGAC GTCGCATATA TCCGAGTGGT CGGATTATGT AATTGAAA AGCTTTGTT TTTGGTGTG GTTGTGTTT
Sat-1 ORF       D Q T S H I S E W G D Y V I *
1281
Sat-2            ATGGCTTTTC ATGTTCTCT CCGCTTCTTC TGTATTGAA GCTGGTGAAG TATATGATAT GCGATGATGC TGTTCCTATA 1280
Sat-1            ATGGCTTTTC ATGTTCTCT CCGCTTCTTC TGTATTGAA GCTGGTGAAG TATATGATAT GCGATGATGC TGTTCCTATA
1281
Sat-2            AaaAaAaaAa aaaaaaaaaa AaAaaAaAa ..... 1360
Sat-1            AtcAgAtcAt ttgtctctcc AtAAGAtAaC actgattaag atgggattgg aaagtaaact cggaaactct gaactctgtt
1361
Sat-1            tagtggagtt tctctgtggt ggttctctgat tagcagagac ttaagatag tttagtagtt aagtgttat ttggttaaac 1440
1441
Sat-1            catgatttct gtatcaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 1487

```

**Figure 3.10** Nucleotide sequence comparison of Sat-1, Sat-2 and EST 89K20T7 clones, and deduced amino acid sequence of clone Sat-1

An alignment of the 1278 bp *Arabidopsis thaliana* clone Sat-1 with clone Sat-2 and the EST clone 89K20T7 is shown, with common nucleotides in capitals. The longest open reading frame (Sat-1 ORF) deduced from Sat-1 is given beneath the Sat-1 nucleotide sequence, and the N-terminal residues extending beyond those encoded by Sat-2 are underlined. The stop codon is indicated by an asterisk.

laboratory), was found to have an identical sequence to Sat-1 at the 5' end, except that Sat-1 was truncated by at least 200 bp (Figure 3.10). In addition, an in-frame stop codon could not be found upstream from the first methionine of the polypeptide encoded by Sat-1. These findings indicated that neither the Sat-1 transcript nor the encoded polypeptide were full length.

An *A. thaliana* serine acetyltransferase cDNA clone, Sat1-6, with 97 % identity to Sat-1, but shorter by 11 bp at the 5' end and with a poly(A) site corresponding to that of Sat-2, was published by Hell's group subsequent to completion of this work (Bogdanova et al. 1995). Several discrepancies between Sat-1 and Sat1-6 sequence within the coding region were resolved when the Sat1-6 GenBank entry (X82888) was updated and corrected subsequent to the submission of the Sat-1 sequence to GenBank, but several differences in the 3' untranslated sequence prior to the poly(A) site remained: a sequence TCT (nucleotides 1019 - 1021 of Sat-1 and 909 - 911 of Sat-2) is not present in Sat1-6, and a single nucleotide A (nucleotide 1039 of Sat-1 and 939 of Sat-2) is represented by a G in Sat1-6. Although Sat-1 and Sat1-6 may represent distinct genes encoding serine acetyltransferase, the possibility of errors in the Sat1-6 sequence cannot be excluded. Barroso et al. (1995) also have raised the latter conclusion after noting disparities in nucleotide and deduced protein sequences of an *O*-acetylserine (thiol)-lyase gene cloned by their group and that of Hell (Hell et al. 1994). Nevertheless, evidence presented above which suggests that Sat-1 is not full-length applies also to the Sat1-6 clone of Bogdanova et al. (1995).

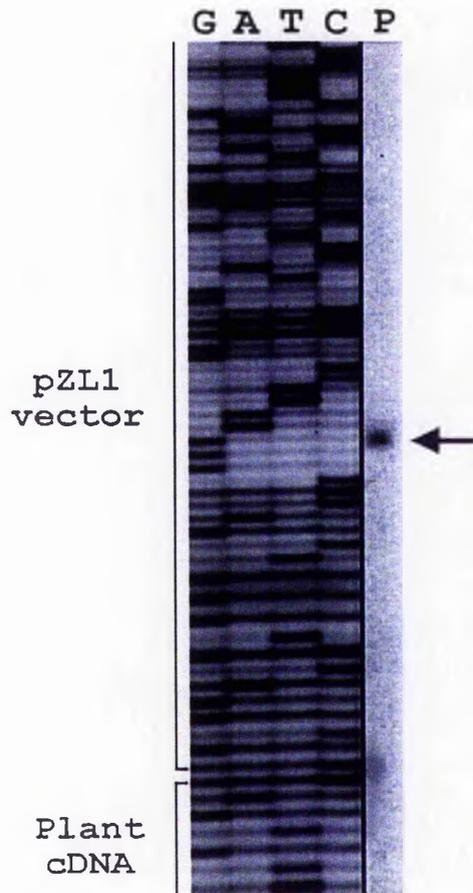
#### *Primer extension analysis of EST clone 89K20T7*

To determine experimentally whether the Sat-1 homologue EST clone 89K20T7 was full-length, the transcription initiation site was determined by primer extension analysis (Chapter

2.11). A [ $\gamma$ - $^{32}$ P]-ATP end-labelled 18 bp oligonucleotide P172 (5'-AGG AAA GAG TGA TGA TTG), complementary to bp 61 to 79 from the 5' end of EST clone 89K20T7, was annealed to *A. thaliana* total RNA template and extended using reverse transcriptase (Figure 3.11). A single transcription start site about 22 bp upstream from the 5' end of clone 89K20T7 was detected, indicating that the EST clone itself did not represent a full length transcript. Analysis of the 5' end of several *A. thaliana* EST clones has shown that the majority of EST clones contain a translation initiation codon for the corresponding complete polypeptides, but that the cDNA itself might not be full length (Newman et al. 1994). Unequivocal confirmation of a full-length open reading frame for a previously uncharacterised gene, however, can be attained from a cDNA sequence only if it is shown to represent a complete transcript or if an in-frame stop-codon is found upstream from an ATG initiation codon.

*5'RACE amplification of an Arabidopsis thaliana putative full-length Sat-1 cDNA clone*

The RACE (Rapid Amplification of cDNA Ends) technique is a PCR-based method for obtaining gene-specific full-length cDNAs. For cloning the 5' end of cDNAs, the original RACE method (Frohman et al. 1988, Belyavsky et al. 1989) and several adaptations (Loh et al. 1989; Ohara et al. 1989; Harvey & Darlison 1991) involved first-strand synthesis of cDNAs by reverse transcriptase, followed by addition of homopolymeric (dA) tail at the 3' end of synthesised first-strand cDNAs, second-strand cDNA synthesis in the presence of a poly(dT)-adaptor primer, and PCR amplification using complimentary gene-specific and adaptor primer oligonucleotides. Homopolymeric tailing was a limiting step in those protocols, so an improved strategy termed SLIC (single strand ligation to single-stranded cDNA), which enabled the ligation of a single-stranded oligonucleotide anchor to the 3' end of single-stranded cDNA, was developed



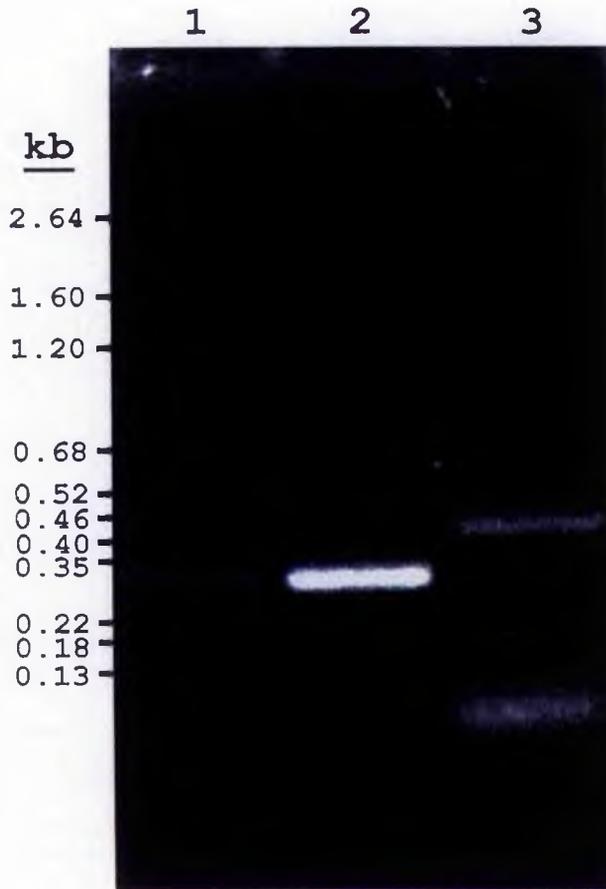
**Figure 3.11** Determination of the putative transcription initiation site of EST clone 89K20T7 by primer extension

Primer extended products (P) using 30  $\mu$ g *Arabidopsis thaliana* total leaf RNA were run in parallel with primer-labelled dideoxynucleotide sequencing reaction (G, A, T, C) of plasmid pZL1/89K20T7 on a polyacrylamide gel and exposed to X-ray film. Plant cDNA and vector sequences are indicated. A single band corresponding to the transcription start site is marked with an arrow.

(Edwards et al. 1991). The 5'-AmpliFINDER RACE Kit (Clontech Laboratories, USA), which is a modified version of the SLIC method, was employed here as described in Chapter 2.9 to obtain the 5' end of the truncated clone Sat-1.

Positive control reactions using reagents supplied with the 5'-AmpliFINDER RACE kit were performed prior to experimental work to test the efficacy of the kit. First-strand cDNA was synthesised from human placental poly(A) RNA using the primer TFR1 complementary to a sequence approximately 450 bp from the 5' end of the human transferrin receptor (TFR) gene. PCR amplification using the primers TFRQC, complementary to the extreme 5' end of TFR gene, and TFR2, complementary to the TFR gene slightly upstream from the TFR1 primer, was then performed using either unpurified or concentrated and purified first-strand cDNA as template. A single band of about 340 bp was observed upon agarose gel electrophoresis of both amplification reactions (Figure 3.12, lanes 1 and 2). As predicted, the 340 bp band was stronger using purified template (Figure 3.12, lane 2), indicating that first-strand cDNA synthesis and purification were successful. The AmpliFINDER anchor was then ligated to purified first-strand cDNA, and the product used as template in a PCR reaction using the anchor primer and TFR2 primer. An expected band of approximately 480 bp, representing the first 430 nucleotides of the TFR gene and 50 bp of the anchor primer and anchor sequence, was detected upon fractionation of the PCR products (Figure 3.12, lane 3). Low molecular weight primer dimers were also observed in the PCR amplification reactions. The positive control reactions thus demonstrated that the 5'AmpliFINDER RACE kit was operative.

First-strand cDNA can be synthesised from either total RNA or purified poly(A) RNA, but the latter template was recommended because mRNA species are more concentrated in poly(A) RNA and less PCR amplification is required in subsequent steps (5'-AmpliFINDER RACE Kit Protocol, Clontech Laboratories, 1993).

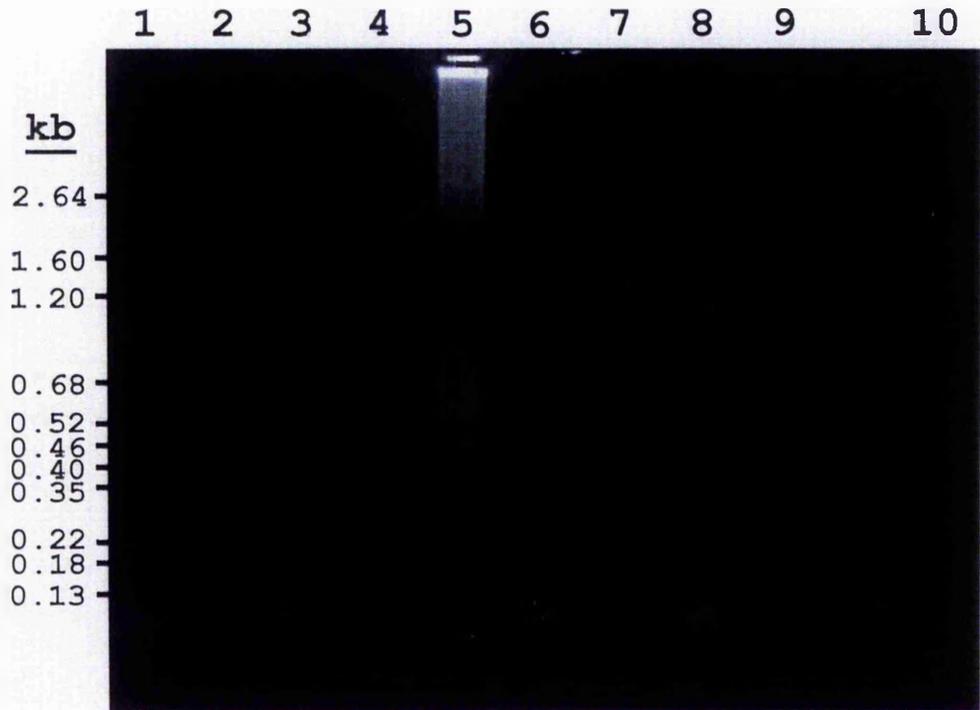


**Figure 3.12** PCR amplification of the 5' end of the human TFR gene using the 5'-AmpliFINDER RACE kit

A positive control experiment was performed using material provided with the 5'-AmpliFINDER RACE kit. Human placental poly(A) RNA was used as template for first-strand cDNA synthesis primed by the human transferrin receptor (TFR) gene-specific primer TFR1. Gene-specific primers TFRQC and TFR2 were used to PCR amplify unpurified cDNA (result shown in lane 1) and purified cDNA (result shown in lane 2). Purified first-strand cDNA was ligated to the AmpliFINDER anchor, and the anchor-ligated cDNA amplified by PCR using TFR2 and the anchor primer (lane 3). PCR products were fractionated through a 1.8 % (w/v) agarose gel. Band positions of pGEM molecular weight markers are shown.

Poly(A) RNA was therefore isolated from above-ground parts of *A. thaliana* using a Dynabeads mRNA DIRECT kit (DynaL, UK; Chapter 2.7.3). Dynabeads are magnetic oligo (dT) microspheres that bind to poly(A) RNA particles in crude plant extracts, thus allowing poly(A) RNA purification without the requirement for prior total RNA isolation (Jakobsen et al. 1990). While suitable for a wide range of tissue, reports indicate that Dynabeads are effective for purifying intact poly(A) RNA from *A. thaliana* (eg. Kaldenhoff et al. 1993; Villand et al. 1993). Small aliquots of poly(A) RNA isolated here were too dilute to be visualised by denaturing agarose gel electrophoresis. Autoradiography of fractionated <sup>32</sup>P-labelled first-strand cDNA synthesised from poly(A) RNA (Chapter 2.7.5) demonstrated that mRNA species of up to 3 kb were present in the purified poly(A) RNA sample (not shown). The isolated poly(A) RNA was therefore considered to be suitable for use as template in the 5'-AmpliFINDER RACE procedure.

A 30-mer oligonucleotide sequence, PR1, complementary to Sat-1 approximately 250 bp from the 5' end of the clone, was selected for first-strand cDNA synthesis from poly(A) RNA, the initial stage of the 5' RACE procedure. Primer PR2, a 28-mer about 80 bp upstream from PR1, was then used for amplification of the anchor-ligated cDNA. PR1 and PR2 were chosen from regions where there was little identity between Sat-1 and *A. thaliana* genes encoding different isoforms of serine acetyltransferase which were becoming available in the GenBank/EMBL databases (Section 3.2.3.3): both primers could therefore be regarded as gene-specific. Based on primer extension analysis (Figure 3.11), clone Sat-1 was truncated by an estimated 241 bp, so the amplified 5' cDNA segment from the RACE protocol was expected to be approximately 400 bp in size. However, only a low molecular weight primer-dimer band was visible after agarose gel fractionation of products from the initial PCR amplification using PR2 and the AmpliFINDER anchor primer (Figure 3.13, lane 1). Variation of the PCR



**Figure 3.13** RACE amplification of the 5' end of clone Sat-1. I. Use of primer PR2 and cDNA synthesised from poly(A) RNA

Following first-strand cDNA synthesis from *Arabidopsis thaliana* poly(A) RNA and anchor ligation using the 5'-AmpliFINDER kit, attempts were made to amplify the 5' end of clone Sat-1 using the anchor primer and a gene-specific primer PR2. Two basic PCR programs were used, with cycle parameters varied as stated: (i) denaturation at 94°C for 45 s, annealing at 60°C (lanes 1 & 2), 55°C (lane 3) or 65°C (lane 4) for 45 seconds, extension at 72°C for 2 min, cycled 35 times (lane 1) or 45 times (lanes 2 - 4) with a final extension time of 7 min; and (ii) initial steps of denaturation at 94°C for 5 min, annealing at 50°C (lane 5 & 10), 51°C (lane 6), 52°C (lane 7), 53°C (lane 8) or 55°C (lane 9) for 2 min, extension at 72°C for 3 min, followed by 35 cycles of denaturation at 94°C for 1 min, annealing for 2 min at temperatures as in initial step, extension at 72°C for 3 min, and a final extension time of 10 min. A PCR reaction lacking anchor-ligated cDNA was included as a negative control (lane 10). PCR products were fractionated through a 1.8 % (w/v) agarose gel. Band positions of pGEM molecular weight markers are shown.

parameters denaturation time, annealing temperature and time, extension time and number of cycles did not result in amplification of a band of the expected size (Figure 3.13, lanes 2 -9). Annealing temperatures of 50 and 52°C (but not 51°C) generated non-specific amplification of DNA, evident as smeared products (Figure 3.13, lane 5). In the absence of template DNA in the amplification reaction at an annealing temperature of 50°C, no smear was present (Figure 3.13, lane 10).

Another series of PCR amplifications were performed using a different gene-specific primer, P9, which is a 17-mer complementary to a sequence 203 bp from the 5' end of Sat-1. A product of at least 450 bp representing a full-length 5' end fragment of the Sat-1 gene would be expected. From a range of cycling parameters attempted (Figure 3.14, lanes 1 - 8), a distinct band of about 460 bp was amplified at an annealing temperature of 54°C (Figure 3.14, lane 6). The 460 bp band was also amplified at annealing temperatures of 52 and 53°C, but at these temperatures other PCR products were generated (Figure 3.14, lanes 4 & 5; visible in original gel). The 460 bp band was not present after PCR reactions without template DNA at annealing temperatures of 50 and 54°C (Figure 3.14, lanes 9 & 10). The 460 bp PCR product amplified at an annealing temperature of 54°C, termed Sat-1.5.1, was isolated by excision from the agarose gel and purified.

The Sat-1 clone contains a single *EcoRV* restriction enzyme recognition site (GAT↓ATC) 26 bp upstream from the 5' end of the P9 primer sequence. Several unsuccessful attempts were made to subclone the 460 bp Sat-1.5.1 fragment digested with restriction enzymes *EcoRI* (for which there is a recognition site in the AmpliFINDER anchor primer sequence) and *EcoRV* into pBluescript II KS (Stratagene, UK) which was digested with the same restriction enzymes in the order specified by the supplier. Subsequently it was noted that an appendix in the 1994 and 1995 Stratagene catalogues on restriction



**Figure 3.14** RACE amplification of the 5' end of clone Sat-1. II. Use of primer P9 and cDNA synthesised from poly(A) RNA

Attempts were made to amplify the 5' end of clone Sat-1 using anchor-ligated cDNA synthesised from *Arabidopsis thaliana* poly(A) RNA, the anchor primer and a gene-specific primer P9. As in Figure 3.13, two basic PCR programs were used: (i) denaturation at 94°C for 45 s, annealing at 65°C (lane 1) for 45 seconds, extension at 72°C for 2 min, cycled 45 times (lane 1) with a final extension time of 7 min; and (ii) initial steps of denaturation at 94°C for 5 min, annealing at 50°C (lane 2 & 9), 51°C (lane 3), 52°C (lane 4), 53°C (lane 5), 54°C (lane 6 & 10), 55°C (lane 7) or 56°C (lane 8) for 2 min, extension at 72°C for 3 min, followed by 35 cycles of denaturation at 94°C for 1 min, 2 min annealing at temperatures as in initial step, extension at 72°C for 3 min, and a final extension time of 10 min. PCR reactions lacking anchor-ligated cDNA were included as negative controls (lanes 9 & 10). PCR products were fractionated through a 1.8 % ( $\frac{w}{v}$ ) agarose gel. Band positions of pGEM molecular weight markers are shown.

endonuclease cleavage activity near DNA termini for the cloning site of pBluescript II contained incorrect information. The order of digestions in the appendix was reversed compared with experimental findings of Kaufman and Evans (1990). Thus, cleavage of the pBluescript II cloning site first by *EcoRV* and then by *EcoRI*, as recommended by Stratagene to achieve 70 % or higher efficiency, would actually result in less than 15 % cleavage (Kaufman & Evans 1990). (Stratagene have been informed of this misinformation.)

Due to the difficulty experienced in subcloning Sat-1.5.1 into pBluescript II, the PCR product was purified and sequenced directly (reviewed in Rao 1994). Very faint sequence bands, which required about six weeks exposure to X-ray film for visualisation, were obtained when either double- or single-stranded DNA was used as template (not shown). During this prolonged exposure time, the DNA sequence of Sat-1.5.1 was obtained by standard plasmid sequencing after a *ClaI/EcoRI*-digested fragment was successfully subcloned into the vector pGEM-7Zf(+) (Promega Corporation, UK). No homology between the 305 bp Sat-1.5.1 fragment (Figure 3.15) and Sat-1 or EST clone 89K20T7 could be found. However, a BLAST search with the Sat-1.5.1 sequence revealed highly significant identity with several plant chloroplast genome sequences for NADH dehydrogenase subunit 2 (not shown). Fragment Sat-1.5.1 was thus an artifact introduced by contamination of isolated mRNA, contamination of the PCR step or during subcloning.

The 5'-RACE protocol was repeated using total RNA extracted from *A. thaliana* leaf tissue. Intactness of isolated RNA was determined by analysis of ribosomal RNA bands following denaturing agarose gel electrophoresis (Section 3.2.4.2). As previously, first-strand cDNA was synthesised from total RNA using the Sat-1 gene-specific primer PR1, ligated to the AmpliFINDER anchor sequence, and amplification by PCR was conducted using the AmpliFINDER anchor primer and either PR2

```

TAGATCAACTAAGCCCTCTCGGGGACTTGCTTAAGAATAAGAAAGAGCAATCTCATGTAAATACCATGGAATAAGGTTTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1 atctagttgattcgggagagccctgaacgaattcttattcttctctcgtagagtacatttatgggtaccttattccaaaa

a * I N * A L S G T C L R I R K S N L M * I P W N K V L -
b R S T K P S R G L A * E * E R A I S C K Y H G I R F Y -
c D Q L S P L G D L L K N K K E Q S H V N T M E * G F -
d S * S L G R P S K S L F L F S C D * T F V M S Y P K -
e I L * A R E P V Q K L I L F L L R M Y I G H F L T K -
f L D V L G E R P S A * S Y S L A I E H L Y W P I L N -

ATCCTATTTCATGGGATTCGGTAAATATTCATTCCAAAAATCAAAAATGGTTTTTTTTTTGGAGATTGGATGCAGTTACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
81 taggataagtaccctaaggcatttataaggtaagggttttagtttttaaccAAAAAAACCTTAACCTACGTCAATGA

a S Y S W D S V N I P F Q K S K I G F F L E I G C S Y * -
b P I H G I P * I F H S K N Q K L V F F W R L D A V T -
c I L F M G F R K Y S I P K I K N W F F F G D W M Q L L -
d I R N M P N R L Y E M G F I L F Q N K K P S Q I C N S -
e D * E H S E T F I G N W F D F I P K K K S I P H L * -
f * G I * P I G Y I N W E L F * F N T K K Q L N S A T V -

AATTCATGATCTGGCATGTACAGAATGAAAATTCATTCTCgattctacgagaattttttgaaagcctttcatttgcttc
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
161 TTAAGTACTAGACCGTACATGTCTTACTTTTAAAGTAAGAGCTAAGATGCTCTTAAAAAACTTTCGGAAAGTAAACGAAG

a F M I W H V Q N E N F I L D S T R I F * K P F I C F -
b N S * S G M Y R M K I S F S I L R E F F E S L S F A S -
c I H D L A C T E * K F H S R F Y E N F L K A F H L L L -
d I * S R A H V S H F N * E R N * S F K K F A K * K S -
e * N M I Q C T C F S F K M R S E V L I K Q F G K M Q K -
f L E H D P M Y L I F I E N E I R R S N K S L R E N A E -

tcttcgatggaagttttattttccagaatgtatcctaatttttggcctaattcttctctgatg
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
241 AGAAGCTACCTTCAAAATAAAAGGGTCTTACATAGGATTAAAAACCGGATTAAGAAGAAGACTAC 305

a S S M E V L F S Q N V S * F L A * F F F * -
b L R W K F Y F P R M Y P N F W P N S S S D -
c F D G S F I F P E C I L I F G L I L L M -
d R K S P L K I K G S H I R I K P R I R R R I -
e E E I S T K N E W F T D * N K A * N K R Q H -
f R R H F N * K G L I Y G L K Q G L E E E S -

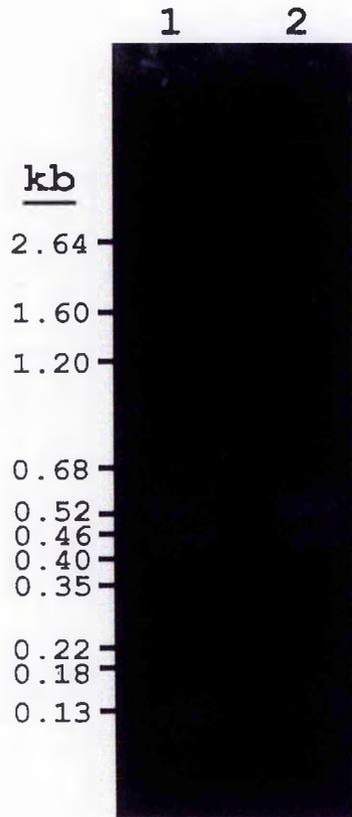
```

**Figure 3.15** DNA sequence and deduced reading frames of clone Sat-1.5.1

Clone Sat-1.5.1, amplified using the 5'-AmpliFINDER RACE kit (Figure 3.14, lane 7) and subcloned into a pGEM-7Zf(+) plasmid, was sequenced from 5' and 3' ends using vector-based primers. Sequenced nucleotides of the Sat-1.5.1 insert are shown in capitals, while unsequenced complementary nucleotides are in small letters. Nucleotide translations in three forward (a, b and c) and three reverse (d, e and f) reading frames are given. Stop codons are indicated by asterisks. The longest open reading frame of 43 amino acids (reading frame b, commencing at nucleotide 176) is not terminated by a stop codon within the Sat-1.5.1 sequence.

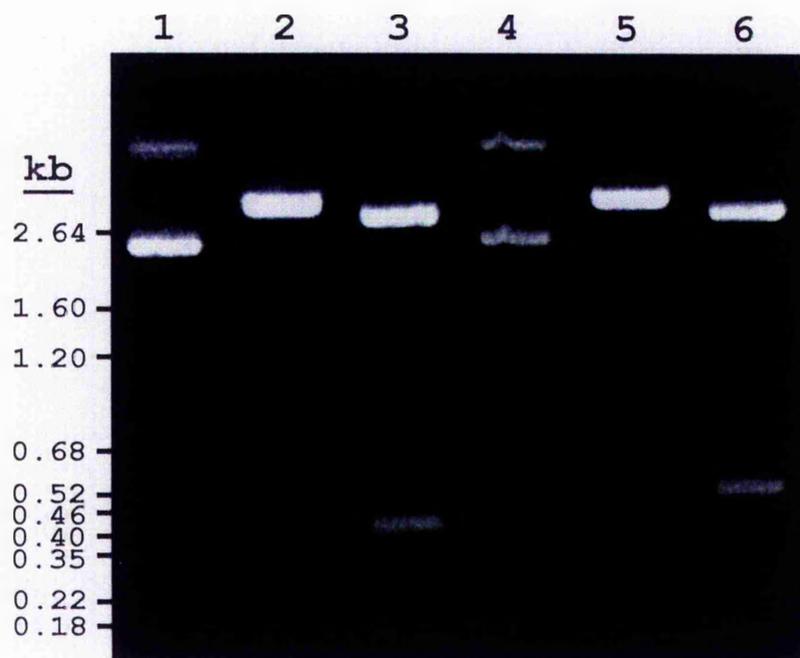
or P9. At an annealing temperature of 55°C, two bands approximately 400 and 500 bp in size were amplified in both PCR reactions with the gene-specific primers PR2 and P9 (Figure 3.16). Discrimination of the 21 bp difference in size from amplification of homologous fragments using PR2 (Figure 3.16, lane 1) and P9 (Figure 3.16, lane 2) after fractionation through a 1.8 % (w/v) agarose gel is not possible (Sambrook et al. 1989). The fragments amplified by PR2 and the AmpliFINDER anchor primer were subcloned into pGEM-7Zf(+) (Figure 3.17). The 530 bp fragment, Sat-1.5.2 (Figure 3.17, lane 6), and the 400 bp fragment, Sat-1.5.3 (Figure 3.17, lane 3), were then sequenced using the standard double-stranded plasmid sequencing method.

The Sat-1.5.2 and Sat-1.5.3 sequences were 100 % identical with the 5' end of the original Sat-1 clone for about 180 bp (up to the PR2 primer sequence used for gene-specific PCR amplification), but were 237 and 119 bp longer, respectively, than Sat-1. Sat-1.5.2 was 28 bp longer at the 5' end than the EST clone 89K20T7, revealing a slight discrepancy between the primer extension experiment which had demonstrated that clone 89K20T7 was truncated by approximately 22 bp (Figure 3.11) and the length of Sat-1.5.2. Appending the Sat-1.5.2 sequence obtained from the 5'-RACE procedure to Sat-1 resulted in a 1515 bp Sat-1a clone that was designated as a full-length transcript (Figure 3.18; Roberts & Wray 1996; GenBank entry U22964). Subsequently, another *A. thaliana* EST clone, 133C22T7 (GenBank entry T45495), was identified that was significantly homologous to the 5' end of Sat-1a but was longer than Sat-1a by 50 bp (BLASTN Score = 1083;  $P = 2.3 \times 10^{-82}$ ). Although clone 133C22T7 is only 90 % identical with Sat-1a, differences can be attributed to unknown nucleotides ("n") in the EST clone sequence and an increase in sequencing errors after 300 bp in the EST clone sequence to up to 4 % (Newman et al. 1994). Confirmation that the two clones are isogenic is required, yet EST clone 133C22T7 does not contain an initiation codon in any three reading frames prior to the



**Figure 3.16** RACE amplification of the 5' end of clone Sat-1. III. Use of cDNA synthesised from total RNA and primers PR2 and P9

Attempts were made to amplify the 5' end of clone Sat-1 using anchor-ligated cDNA synthesised from *Arabidopsis thaliana* total RNA, the anchor primer and gene-specific primers PR2 (lane 1) or P9 (lane 2). Both samples were subjected to the following PCR cycles: initial steps of denaturation at 94°C for 5 min, annealing at 55°C for 2 min, extension at 72°C for 3 min, then 35 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 2 min, extension at 72°C for 3 min, with a final extension time of 10 min. A further 0.4  $\mu$ l (2 units) Taq DNA polymerase was added, and an extra 5 cycles (as for the latter 35 cycles) were performed. PCR products were fractionated through a 1.8 % ( $v/v$ ) agarose gel. Band positions of pGEM molecular weight markers are shown.



**Figure 3.17** Electrophoretic analysis of Sat-1.5.2 and Sat-1.5.3 fragments ligated into pGEM-7Zf(+)

DNA bands amplified by the 5'-AmpliFINDER kit using PR2, derived from clone Sat-1, and the anchor primer (Figure 3.16, lane 1) were subcloned into the vector pGEM-7Zf(+). Digestion of ligated vectors using the restriction endonucleases XbaI and SacI that flank the insertion site followed by fractionation through a 1.6 % ( $v/v$ ) agarose gel revealed the presence of linearised vector (2.9 kb) and two classes of insert size: a smaller fragment of 400 bp (designated Sat-1.5.3; lane 3) and a larger fragment of 530 bp (designated Sat-1.5.2; lane 6). Also shown are undigested plasmids containing Sat-1.5.3 (lane 1) and Sat-1.5.2 (lane 4) which could be linearised by SacI restriction (lanes 2 and 5, respectively). Band positions of pGEM molecular weight markers are shown.



overlap with Sat-1a, and the SAT-1 protein encoded by Sat-1a (Figure 3.18) can thus be regarded as full-length. Very recently, a 1236 bp serine acetyltransferase sequence (designated pYES SAT1 [Ruffet, M.-L., Lebrun, M., Droux, M., Douce, R. 1996, unpublished GenBank accession L78443]) with 99.9 % identity to Sat-1a was submitted to GenBank. However, the pYES SAT1 sequence is shorter by 32 nucleotides than Sat-1a at the 5' end, and is truncated at the 3' end at nucleotide 1268 relative to Sat-1a. The Sat-1a and pYES SAT1 sequences differ at a single nucleotide (nucleotide 576 of Sat-1a).

It is possible that less abundant longer Sat-1 5' ends, not visible on ethidium bromide stained gels, were amplified in the 5'-RACE protocol from total RNA. Sequencing of more clones or Southern blotting (Frohman et al. 1988) would be required to confirm this. It is unclear whether the length variation at the 5' end of the Sat-1-like transcripts, viz. Sat-1, Sat-1.5.2 and Sat-1.5.3 from this study, EST clones 89K20T7 and 133C20T7, pYES SAT1 (Ruffet et al. 1996, unpublished GenBank entry L78443) and Sat1-6 (Bogdanova et al. 1995), are caused by experimental artifacts or whether heterogeneity of transcript length, possibly regulated by tissue-specific factors (Jitrapakdee et al. 1996), occurs *in vivo*. From the experimental perspective, analysis of other publications in which the *A. thaliana* pYES library is used reveals some reports where cloned cDNAs were definitely not full-length (Pang et al. 1992; Schnorr et al. 1994; Ruffet et al. 1995), and others where clones were putatively full-length (Hsu et al. 1993; Jain & Leustek 1994; Leustek et al. 1994; Stallmeyer et al. 1995). Unequivocal demonstration that clones were full-length by primer extension or S1 mapping (Sambrook et al. 1989) was not provided in the latter studies, but the presence of an in-frame stop codon upstream from the initiation codon of the longest open reading frame, which confirms a full-length open reading frame, is often given as evidence for a full-length transcript. Therefore,

the relative abundance of full-length cDNAs in the  $\lambda$ YES cDNA library is difficult to quantify. Since plant N-terminal transit peptides can direct polypeptides expressed in *E. coli* to the periplasm (Overbeeke & Verrips 1993), it is feasible that functional complementation will favour selection of shorter cDNAs that encode functional polypeptides required for phenotypic rescue but lacking complete transit peptides.

### 3.2.3.3 Analysis of clone Sat-1a

#### *Characterisation of clone Sat-1a and sequence comparisons*

The longest open reading frame within the 1 515 bp Sat-1a cDNA sequence codes for a polypeptide, SAT-1, of 391 amino acids, with a calculated molecular weight of 42.7 kD (Figure 3.18). SAT-1 is 87 residues longer than the polypeptide encoded by Sat-1 (Figure 3.10) and 110 residues longer than ATHSAT-2 (Figure 3.7). None of the five ATGs within the first 300 nucleotides of Sat-1a have flanking sequences which show complete identity with the plant translation start site consensus sequence AACAATGGC (Joshi 1987; Lütcke et al. 1987). Three ATGs (coding for residues 1, 12 and 56) lie within six or seven of the nine consensus nucleotides, and two (coding for residues 1 and 12) could be regarded as being in optimal context for translation initiation, with an A or G nucleotide at position -3 relative to the ATG (Joshi 1987). In view of the "first AUG rule" (Kozak 1984), which holds true for 96 % of plant genes without downstream terminator codons before functional ATGs (Joshi 1987), it is likely that it is the first ATG of Sat-1a that initiates translation.

A BLASTP search with SAT-1 revealed that the protein has significant homology with all available serine acetyltransferase sequences (high scores range from 1156 to 181 and corresponding P values are  $6.6 \times 10^{-163}$  to  $6.2 \times 10^{-24}$ ). A GCG PileUp alignment of these enzymes, which originate from

higher plants and bacteria, is shown in Figure 3.19. As expected, SAT-1 has greatest similarity with higher plant serine acetyltransferases. SAT-1 is 76 % identical (87 % homologous) to the *A. thaliana* serine acetyltransferase isozyme SAT5 (Ruffet et al. 1995). Another isoform, SAT-52 (Howarth et al. 1997), has 55 % identity (70 % homology) with SAT-1. (Clone Sat-52 was obtained from the complementation experiment described in Section 3.2.1, but was sequenced and characterised by Mr J.R. Howarth in this laboratory.) SAT-1 also has 59 % identity (73 % homology) with SAT2 of *Citrullus vulgaris* (Saito et al. 1995). The next group of serine acetyltransferase enzymes showing significant homology to SAT-1 (as observed for ATHSAT-2 where the sequences were available) are the CysE proteins from a group of related proteobacteria: (i) *E. coli* (Denk & Böck 1987; Tei et al. 1990a) and *Salmonella typhimurium* (Sivaprasad, A.V., Kuczek, E.S., Bawden, C.S., Rogers, G.E. 1991, unpublished SWISS-PROT entry P29847); (ii) *Haemophilus influenzae* (Fleischmann et al. 1995); and (iii) *Buchnera aphidicola* (Lai & Baumann 1992), an endosymbiont of aphids (*Schizaphis graminum*). Two NIFP (serine acetyltransferase) proteins from proteobacteria belonging to the family Azotobacteriaceae, *Azotobacter chroococcum* (Evans et al. 1991) and *Azotobacter vinelandii* (Jacobson et al. 1989), also show significant homology with SAT-1, as do three serine acetyltransferases (CysE or SrpH proteins) from the cyanobacteria genera *Synechococcus* and *Synchocystis* (Adandan, S., Nalty, M.S., Cogdell, D.E., Golden, S.S. 1995, unpublished GenBank entry L41665; Nicholson et al. 1995; Sakamoto & Murata 1995). The CysE protein from a firmicutes bacterium of the family Bacillaceae, *Bacillus subtilis* (Gagnon et al. 1994), exhibited greater homology to the NIFP and cyanobacterial CysE enzymes than SAT-1.

The relationship of the serine acetyltransferase sequences in Figure 3.19 can be portrayed graphically from the GCG PileUp program as a phylogenetic tree constructed from sequence

C. vulgaris SAT  
 A. thaliana SAT-52  
 A. thaliana SAT-1  
 A. thaliana SAT5  
 E. coli CysE  
 S. typhimurium CysE  
 H. influenzae CysE  
 B. aphidicola CysE  
 A. chroococcum NIFP  
 A. vinelandii ORP7  
 Synechococcus CysE  
 Synechocystis CysE  
 B. subtilis CysE  
 Synechococcus SprH  
 Consensus

C. vulgaris SAT  
 A. thaliana SAT-52  
 A. thaliana SAT-1  
 A. thaliana SAT5  
 E. coli CysE  
 S. typhimurium CysE  
 H. influenzae CysE  
 B. aphidicola CysE  
 A. chroococcum NIFP  
 A. vinelandii ORP7  
 Synechococcus CysE  
 Synechocystis CysE  
 B. subtilis CysE  
 Synechococcus SprH  
 Consensus

C. vulgaris SAT  
 A. thaliana SAT-52  
 A. thaliana SAT-1  
 A. thaliana SAT5  
 E. coli CysE  
 S. typhimurium CysE  
 H. influenzae CysE  
 B. aphidicola CysE  
 A. chroococcum NIFP  
 A. vinelandii ORP7  
 Synechococcus CysE  
 Synechocystis CysE  
 B. subtilis CysE  
 Synechococcus SprH  
 Consensus

C. vulgaris SAT  
 A. thaliana SAT-52  
 A. thaliana SAT-1  
 A. thaliana SAT5  
 E. coli CysE  
 S. typhimurium CysE  
 H. influenzae CysE  
 B. aphidicola CysE  
 A. chroococcum NIFP  
 A. vinelandii ORP7  
 Synechococcus CysE  
 Synechocystis CysE  
 B. subtilis CysE  
 Synechococcus SprH  
 Consensus

C. vulgaris SAT  
 A. thaliana SAT-52  
 A. thaliana SAT-1  
 A. thaliana SAT5  
 E. coli CysE  
 S. typhimurium CysE  
 H. influenzae CysE  
 B. aphidicola CysE  
 A. chroococcum NIFP  
 A. vinelandii ORP7  
 Synechococcus CysE  
 Synechocystis CysE  
 B. subtilis CysE  
 Synechococcus SprH  
 Consensus

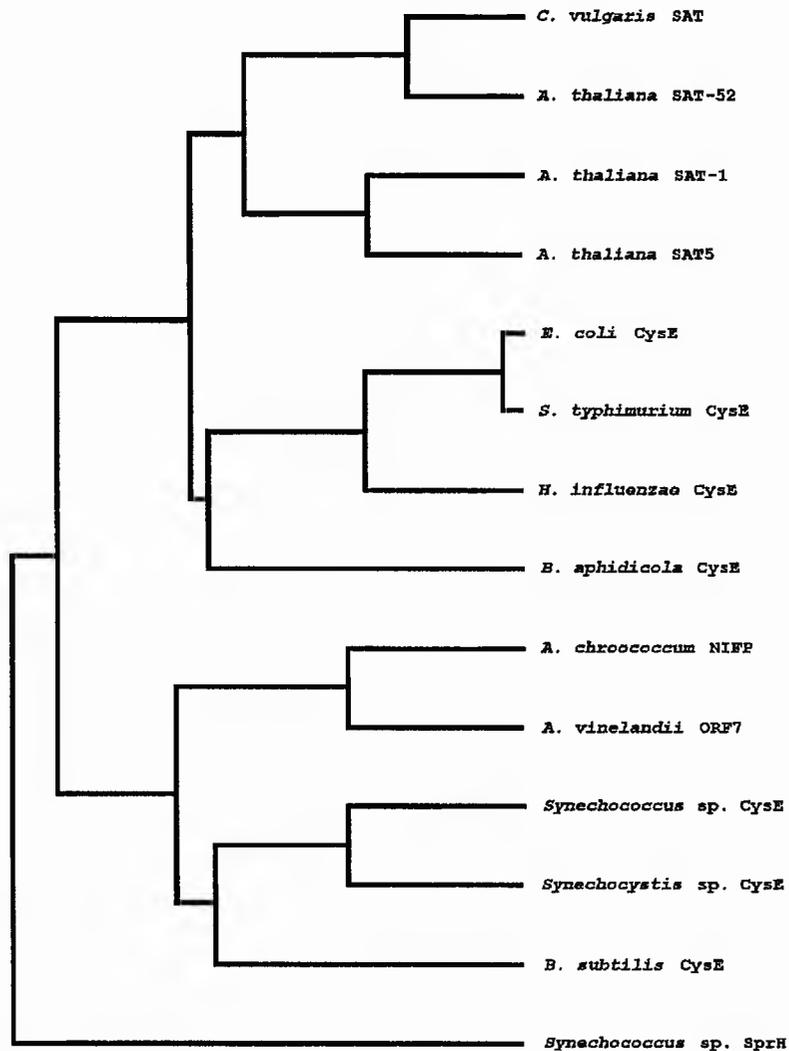
C. vulgaris SAT  
 A. thaliana SAT-52  
 A. thaliana SAT-1  
 A. thaliana SAT5  
 E. coli CysE  
 S. typhimurium CysE  
 H. influenzae CysE  
 B. aphidicola CysE  
 A. chroococcum NIFP  
 A. vinelandii ORP7  
 Synechococcus CysE  
 Synechocystis CysE  
 B. subtilis CysE  
 Synechococcus SprH  
 Consensus

similarity (Figure 3.20). The *E. coli* and *Salmonella typhimurium* CysE proteins, which are 97 % identical, are the most similar sequence pair (with the shortest branch length) to which other proteins were aligned progressively (Chapter 2.12; Figure 3.20). CysE proteins from the proteobacteria *Haemophilus influenzae* and *Buchnera aphidicola* are closely related to *E. coli* CysE, while the plant enzymes have sequences less divergent from the *E. coli* CysE cluster than the group encompassing prokaryotic NIFP, cyanobacterial CysE and *Bacillus subtilis* CysE serine acetyltransferases. The *Synechococcus* SprH protein has the most divergent enzyme sequence but is more closely related to the cluster encompassing NIFP proteins (Figure 3.20).

---

◀ **Figure 3.19** Amino acid sequence alignment of plant and bacterial serine acetyltransferases

Plant serine acetyltransferase isoforms from *Citrullus vulgaris* (SAT [Saito et al. 1995; GenBank translation PID:g1350550 from D85624]) and *Arabidopsis thaliana* (SAT-52 [Howarth et al. 1997; GenBank translation PID:g905391 from U30298]; SAT-1 [this study]; and SAT5 [Ruffet et al. 1995; GenBank translation PID:g608677 from Z34888]) were aligned with each other and the bacterial serine acetyltransferases CysE from *Escherichia coli* (Denk & Böck 1987; Tei et al. 1990a; SWISS-PROT entry P05796), *Salmonella typhimurium* (Sivaprasad, A.V., Kuczek, E.S., Bawden, C.S., Rogers, G.E. 1991; unpublished SWISS-PROT entry P29847), *Haemophilus influenzae* (Fleischmann et al. 1995; SWISSPROT entry P43886), *Buchnera aphidicola* (Lai & Baumann 1992; SWISS-PROT entry P32003), *Synechococcus* sp. (Adandan, S., Nalty, M.S., Cogdell, D.E., Golden, S.S. 1995; unpublished GenBank translation PID:g777763 from L41665), *Synechocystis* sp. (Sakamoto & Murata 1995; GenBank translation PID:g1100770 from D13777) and *Bacillus subtilis* (Gagnon et al. 1994; SWISS-PROT entry Q06750), NIFP from *Azotobacter chroococcum* (Evans et al. 1991; SWISS-PROT entry P23145), ORF7 (orthologous to NIFP) from *Azotobacter vinelandii* (Jacobson et al. 1989; GenBank translation PID:g142363 from M20568) and SprH from *Synechococcus* sp. (Nicholson et al. 1995; GenBank translation PID:g790255 from U23436) using the GCG PileUp program. A consensus sequence common to all 14 proteins is given. Three highly conserved regions are underlined: blocks [1] and [2] represent putative L-serine binding sites, while [3] contains a "hexapeptide-repeat containing-transferase signature" common to a family of acyl- and acetyltransferases (PROSITE entry PS00101; Bairoch & Bucher 1994).



**Figure 3.20** Phylogenetic tree of plant and bacterial serine acetyltransferases

A phylogenetic dendrogram comparing plant and bacterial serine acetyltransferases was constructed using the GCG PileUp program from the alignment shown in Figure 3.19. All sequences are as described in the legend to Figure 3.19. Horizontal branch lengths of the dendrogram are assumed to be proportional to evolutionary distance.

The two main groups of serine acetyltransferases separated in the phylogenetic tree (Figure 3.20) are not consistent with the biological relatedness of the organisms from which the enzymes are originated. Thus the serine acetyltransferases from *E. coli* (Superkingdom: Eubacteria, Phylum: Proteobacteria, Class: gamma division; Family: Enterobacteriaceae) and *Azotobacter vinelandii* (Superkingdom: Eubacteria, Phylum: Proteobacteria, Class: gamma division, Family: Azotobacteriaceae) have less sequence similarity than the *E. coli* enzyme has to that from dicotyledonous plants (Superkingdom: Eukaryotae, Kingdom: Viridiplantae, Phylum: Magnoliophyta, Class: Magnoliopsida). Such inconsistencies in relatedness between whole organisms and molecular sequences can in general be attributed to either disparate rates of change of an assumed ancestral enzyme along different lineages, horizontal gene transfers between organisms, or convergent sequence evolution (Doolittle 1994). Which of these reasons is likely to be true for the serine acetyltransferases?

All of the proteins depicted in Figure 3.20 are reported to be involved in cysteine biosynthesis, but closer inspection reveals heterogeneity in the precise role of the enzymes. The Cyse protein in *E. coli* is expressed constitutively, regulated by a feedback mechanism involving cysteine and is mandatory for sulphate autotrophy of the organism (Jones-Mortimer 1968; Denk & Böck 1987; Kredich 1987). Serine acetyltransferase in plants shares several physiological characteristics with the *E. coli* enzyme, including high levels of expression in the presence of excess sulphate (Bogdanova et al. 1995) and inhibition by cysteine (Brunold & Suter 1982; Saito et al. 1995). In contrast to the role of serine acetyltransferase in these organisms, disruption of the *Azotobacter chroococcum nifP* gene has shown that NIFP is not an obligatory enzyme for cysteine biosynthesis due to the presence of multiple serine acetyltransferases (Evans et al. 1991). It is postulated that NIFP enzymes have a role in

boosting rates of cysteine or methionine formation, especially during nitrogen fixing conditions where additional amino acids are required for synthesis of active nitrogenase components (Evans et al. 1991). Similarly, the *Synechococcus sprH* gene found on the endogenous 48.5 kb plasmid pANL encodes an auxiliary serine acetyltransferase that is expressed during times of sulphur deprivation and is probably involved with other plasmid-encoded genes in scavenging sulphur compounds for cysteine biosynthesis (Nicholson et al. 1995). The *in vivo* functions of the cyanobacterial and *Bacillus subtilis* CysE proteins have not been elucidated, although the latter enzyme is known to be regulated differently from its *E. coli* homologue (Gagnon et al. 1994). It can thus be postulated that the phylogenetic groups into which serine acetyltransferases segregate are dictated by structural features conserved between enzymes which have a similar physiological role. The enzymes in the plant/enterobacterial group are required in the main pathway of cysteine biosynthesis, while the enzymes in the *Azotobacter*/cyanobacterial group are auxiliary. Further information on several serine acetyltransferases is required to examine the validity of this explanation for the phylogenetic groupings in Figure 3.20.

Forty-four amino acid residues are conserved between the aligned serine acetyltransferases, with most of the common residues found in the middle and carboxyl domain of the enzymes (Figure 3.19). Two adjacent blocks contain the highest number of consecutive conserved residues: (1) [VI]-[DE]-[IF]-H-P-[GA] and (2) [IV]-V-[IV]-GE (alignment residues 263 to 269 and 285 to 287, respectively). In addition, ten of the last eleven conserved amino acids form part of a 29-residue signature (residues 332 to 360 of alignment) documented in the PROSITE database (Accession PS00101; Bairoch & Bucher 1994). The consensus pattern of this "bacterial hexapeptide-repeat containing-transferase signature", [LIV] - [GAED] -X(2) - [STAV] -X- [LIV] -X(3) - [LIVAC] -X-

[LIV] - [GAED] -X(2) - [STAVR] -X- [LIV] - [GAED] -X(2) - [STAV] -X- [LIV] - X(3) - [LIV], was matched perfectly in all serine acetyltransferases (Figure 3.19), except for a single disparate residue in the *Synechococcus* SprH protein. Raetz and Roderick (1995) have proposed that the transferase signature domain is involved in binding acetyl CoA and similar substrates. It is thus possible that the conserved blocks (1) and (2) found in the serine acetyltransferases are required for binding L-serine, but further studies are required to establish empirically the substrate binding sites.

The transferase signature defines a family of yeast and bacterial acyl- and acetyltransferases (Downie 1989; Parent & Roy 1992; Vaara 1992; Vuorio et al. 1994) - members of this family are detected by BLAST homology searches using SAT-1. Transferase enzymes homologous to SAT-1 include:

- (i) *Rhizobium leguminosarum* and *R. meliloti* NodL protein (Surin & Downie 1988; Canter Cremers et al. 1989; Baev & Kondorosi 1992; Bloemberg et al. 1994; Journet et al. 1994), an acetyltransferase implicated in the O-acetylation of Nod factors that play a key role in symbiotic association between the microbe and its legume host;
- (ii) *E. coli* galactosidase acetyltransferase (LacA) (Hediger et al. 1985), involved in the synthesis of lactose;
- (iii) *Staphylococcus aureus* CapG (Lin et al. 1994), required for the biosynthesis of type 1 capsular polysaccharide;
- (iv) *Staphylococcus aureus* Vat and VatB (Allignet et al. 1993; Allignet, J., El Solh, N. 1996, unpublished GenBank entry U19459), which inactivate virginiamycin-like antibiotics;
- (v) *Bordetella pertussis* Bp1B (Allen & Maskell 1996), involved in lipopolysaccharide biosynthesis;
- (vi) *Agrobacterium tumefaciens*, *E. coli*, *Morganella morganii* and *Enterobacter aerogenes* chloramphenicol acetyltransferase (CAT) (Tennigkeit & Matzura 1991; Parent & Roy 1992; Kupzig, S., Bennett, P.M. 1994, unpublished GenBank entry X82455;

Bunny, K.L., Hall, R.M., Stokes, H.W. 1995, GenBank entry U13880), evolutionary unrelated to the main family of CAT enzymes and an effector of chloramphenicol resistance;

(vii) *Escherichia coli* NeuD (Annunziato et al. 1995), thought to be responsible for acetylation of polysialic acid in the formation of K1 capsular polysaccharide;

(viii) *Enterococcus faecium* streptogramin A acetyltransferase (Rende-Fournier, R., Leclercq, R., Galimand, M., Duval, J., Courvalin, P., GenBank entry L12033), an effector of streptogramin A antibiotic resistance;

(ix) *Escherichia coli*, *Listeria monocytogenes*, *Bacillus subtilis* and *B. calcohyticus* N-acetylglucosamine-1-phosphate uridyltransferase (GlmU, GcaD or Tms) (Walker et al. 1984; Nilsson et al. 1989; Gouin, E., Mengaud, J., Cossart, P. 1994, GenBank entry M92842; Krath, B.N., Hove-Jensen, B. 1995, unpublished SWISS-PROT entry P42817), required for peptidoglycan and lipopolysaccharide biosynthesis;

(x) *Bacillus sphaericus* macrolide acetyltransferase (Monod et al. 1987), involved in antibiotic resistance;

(xi) *Yersinia enterocolitica* and *Haemophilus influenzae* UDP-N-acetylglucosamine acyltransferase (LpxA) (Coleman & Raetz 1988; Vuorio et al. 1994; Fleischmann et al. 1995), catalysing the first step in the synthesis of lipid A that is an integral part of the outer membrane;

(xii) *E. coli* UDP-N-acetylglucosamine acyltransferase (LpxD, FirA or Ssc) (Dicker & Seetharam 1991; Hirvas et al. 1991), catalysing the third step of lipid A biosynthesis;

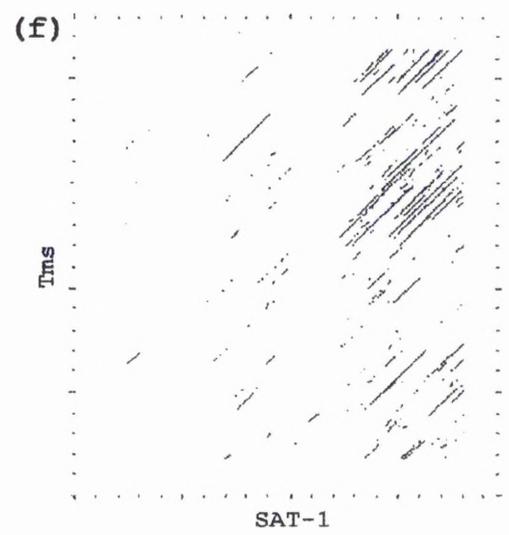
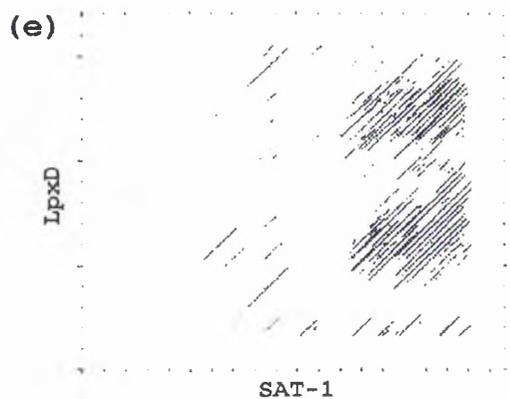
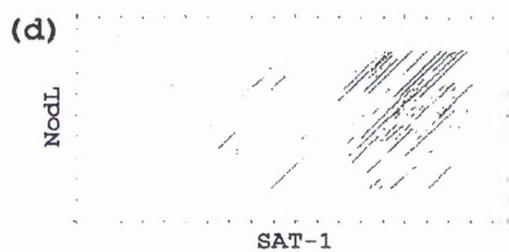
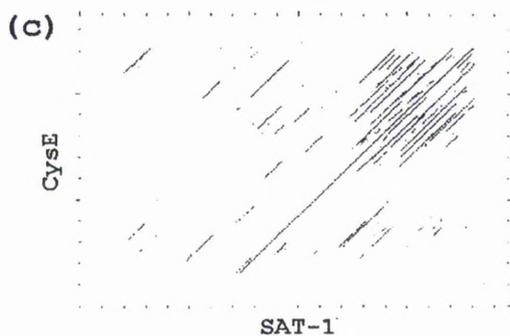
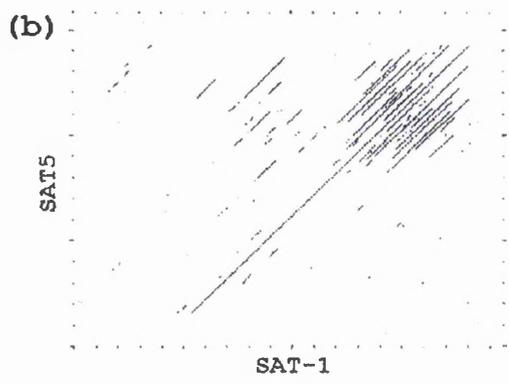
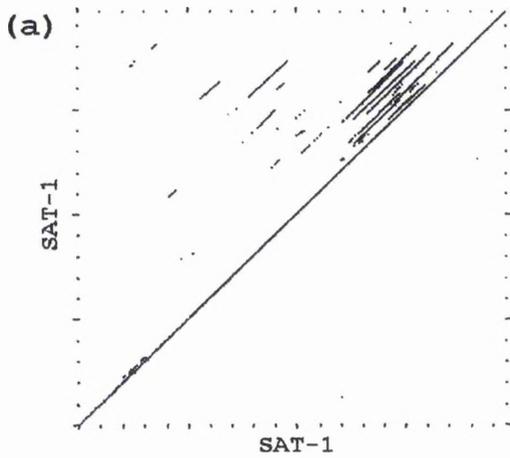
(xiii) *Klebsiella oxytoca* CymB (Fiedler et al. 1996), probably involved in starch metabolism;

(xiv) *Vibrio cholerae* RfbO (Stroehner et al. 1992), involved in serotype determination; and

(xv) hypothetical transferase proteins from *Bacillus subtilis* (Quirk et al. 1994), *Saccharomyces cerevisiae* (Vandenbol et al. 1994), *Schizosaccharomyces pombe* (Devlin, K., Churcher, C.M., Barrell, B.G., Rajandream, M.A., Walsh, S.V. 1995, unpublished SWISS-PROT entry Q09707) and *Synechocystis* sp. (Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T.,

Miyajima, N., Sugiura, M., Tabata, S. 1995, GenBank entry D64006).

Further analysis of the homology of these transferase proteins reveals that the identifying signature domain is based on a fourfold repeat of a [LIV]-G-X(4) motif (Dicker & Seetharam 1992; Vaara 1992; Vuorio et al. 1994). The "hexapeptide repeat" (Vuorio et al. 1994) or "isoleucine patch" (Dicker & Seetharam 1992) is composed of an aliphatic residue (L, I or V) at the first position, followed often but not always by a glycine residue, and a small residue (A, S, C, V, T or N) at the fifth position. The frequency of occurrence of tandem arrangements of the repeat is different in the transferase enzymes in which it is found: in SAT-1, the largest block of consecutive hexapeptide repeats is six (SAT-1 residues 312 to 346, which includes the transferase signature; Figure 3.18), while in LpxA the hexapeptide repeat occurs 28 times within the N-terminal 186 residues (Coleman & Raetz 1988). The presence of internal repeats can be demonstrated graphically using dot-plots, where identity between two sequences compared along their successive residues is scored by a dot and matching blocks appear as lines of dots (Collins & Coulson 1987). A dot-plot comparison of SAT-1 with itself shows several diagonal lines in the C-terminal region off-set from the main self-identity line, indicating internal repeats (Figure 3.21a). Dot-plots comparing SAT-1 with *Arabidopsis thaliana* SAT5 (Figure 3.21b) and *E. coli* CysE (Figure 3.21c) demonstrates the extensive identity which occurs between these serine acetyltransferases and also shows the presence of shared internal repeats in their C-termini. When SAT-1 is compared with *Rhizobium leguminosarum* NodL (Figure 3.21d), *E. coli* LpxD (Figure 3.21e), *Bacillus subtilis* Tms (Figure 3.21f) and *Agrobacterium tumefaciens* CAT (Figure 3.21g), less overall identity is evident but common internal repeats can be observed. Using a window of 60 and a stringency of 25, it was found that background noise was eliminated almost



completely, so the internal repeats visible either in the C-terminus (serine acetyltransferases, NodL and CAT) or throughout the protein (TMS and LpxD) correlate with the frequency of hexapeptide repeats. Furthermore, strong internal homologies were not detected when SAT-1 was compared with several sequences in which the hexapeptide repeat theme is not found. One example, *E. coli* spermidine acetyltransferase, from a different family of acetyltransferases, is shown in Figure 3.21h. When SAT-1 was compared with an *Arabidopsis thaliana* O-acetylserine (thiol)-lyase, no internal homologies were evident (not shown).

---

◀ **Figure 3.21** Dot-plot comparisons of SAT-1 against itself, related enzymes and an unrelated sequence

Dot-plot comparisons were constructed with the deduced protein sequence SAT-1 against (a) itself, (b) *Arabidopsis thaliana* SAT5 (Ruffet et al. 1995; GenBank translation PID:g608677 from Z34888), (c) *Escherichia coli* CysE (Denk & Böck 1987; Tei et al. 1990a; SWISS-PROT entry P05796), (d) *Rhizobium leguminosarum* NodL (Surin & Downie 1988; SWISS-PROT entry P08632), (e) *Escherichia coli* LpxD (Dicker & Seetharam 1991; SWISS-PROT entry P21645), (f) *Bacillus subtilis* Tms (Nilsson et al. 1989; SWISS-PROT entry P14192), (g) *Agrobacterium tumefaciens* CAT (Tennigkeit & Matzura 1991; SWISS-PROT entry P23364) and (h) *Escherichia coli* spermidine acetyltransferase (SA), from a different family of acetyltransferases (Fukuchi et al. 1994; GenBank translation PID:g517105 from D25276). Comparisons were made using the GCG Compare and DotPlot programs with a window of 60 and a stringency of 25. The scale of the axes is 20 residues per tick.

Recently, characterisation of the three dimensional x-ray structure of LpxA revealed that domains containing hexapeptide repeats were folded into novel left-handed parallel  $\beta$ -strand helices, and that  $\beta$ -helices of individual LpxA monomers align with each other to form a homotrimer functional protein (Raetz & Roderick 1995). The possibility that the hexapeptide repeat could be involved in dimerisation, heterodimerisation or oligomerisation of transferases was raised previously after analysis of space-filling molecular models of LpxD (Dicker & Seetharam 1992). While extrapolation of the structure of LpxA to other hexapeptide-containing transferases remains speculative, it is conceivable that the hexapeptide domains of serine acetyltransferases are involved in the *in vivo* formation of a bifunctional enzyme complex with *O*-acetylserine (thiol)-lyase (Kredich et al. 1969; Baecker & Wedding 1980; Nakamura et al. 1988; Saito et al. 1995). The predicted secondary structure (Chou and Fasman 1978; Garnier et al. 1978) of SAT-1 in the C-terminus region with the largest block of hexapeptide repeats was not consistent with the  $\beta$ -strand helix model as there was not a clear predominance of  $\beta$ -strands (data not shown), but it should be noted that the accuracy of these prediction methods has been reported at less than 55 % (Nishikawa 1983).

Transferases containing hexapeptide repeats catalyse very different reactions (Vaara 1992). However, most of the enzymes use an acyl-coenzyme A (eg. serine acetyltransferase) or an acyl-acyl carrier protein (eg. LpxA) as co-substrate, which both have a phosphopantotheryl moiety (Raetz & Roderick 1995). Although the catalytic site of LpxA is not known, it has been suggested that the parallel  $\beta$ -helices of the enzyme contribute towards binding of the acyl-acyl carrier protein substrate (Raetz & Roderick 1995). That suggestion presumes both a structural and catalytic role of the hexapeptide repeats in LpxA, but caution should be exercised in assuming that the hexapeptide repeats from other transferases have a

catalytic, rather than structural, significance (Doolittle 1994).

#### *Subcellular location of SAT-1*

SAT-1 is longer at the N-terminus compared with aligned serine acetyltransferases of prokaryotic organisms, eg. by 116 amino acids for *E. coli*, *Salmonella typhimurium* and *Buchnera aphidicola* CysE proteins, and by 50 amino acids for the *Synechococcus* SprH protein (Figure 3.20). Of the three *A. thaliana* isoforms, SAT-1 is extended at the N-terminus by 76 amino acids beyond SAT5, which has been assigned a cytoplasmic location (Ruffet et al. 1995), and by 79 amino acids beyond SAT-52. SAT-1 is also 97 residues longer at the N-terminus than *Citrullus vulgaris* SAT2, but that sequence has not been shown to be full-length (Saito et al. 1995). In addition, the calculated molecular weight of 42.7 kD for SAT-1 is higher than the 31 - 33 kD size of purified serine acetyltransferases (Nakamura & Tamura 1990; Ruffet et al. 1994). These data strongly suggest that the N-terminal extension of SAT-1 may contain an organellar targeting peptide. Identity between the first 11 residues of the putative cytoplasmic isoform SAT5 and residues 56 - 66 of SAT-1 implies that a targeting peptide would be located within the first 50 - 60 residues of SAT1.

In the first 60 residues of SAT-1, there are no acidic amino acids and a high concentration of hydroxylated amino acids (23 %), a characteristic feature of both mitochondrial and chloroplast targeting peptides (Gavel & von Heijne 1990; von Heijne 1992). A comparison of the fraction of Ser ( $f_{\text{Ser}} = 0.18$ ) residues with the formula  $0.07 + 1.4 \times f_{\text{Arg}}$  ( $= 0.14$ ) predicts with 90 % accuracy that the putative transit peptide is chloroplastic (von Heijne et al. 1989). However, no consensus chloroplastic cleavage sites (Gavel & von Heijne 1990) are evident. Furthermore, conflicting results were obtained when the first 55 residues of SAT-1 were analysed

using PSORT, an automated program for protein localisation (Nakai & Kanehisa 1992), as the residues were predicted to represent a mitochondrial targeting peptide. The accuracy of PSORT is between 20 and 80 % for predicting mitochondrial localisation signals depending on the subcompartment (51 test sequences), and averages 60 % for predicting all subcellular sites (401 test sequences) (Nakai & Kanehisa 1992).

After analysis of a small number of known chloroplast transit peptides, early studies were able to identify sequence similarities (Karlin-Neumann & Tobin 1986; Schmidt & Mishkind 1986). As new sequences accumulated, it was found that those common sequences were not ubiquitous, and new hypotheses for peculiar peptide characteristics had to be considered (Keegstra et al. 1989; von Heijne 1992). Using a sample size of 18 chloroplastic and 37 mitochondrial transit peptides, von Heijne et al. (1989) found a statistically significant predominance of the residues D, E, S, T and Y in the former and A, D, E, I, K, L, R and S in the latter. It has been suggested since that chloroplastic transit peptides are designed to be devoid of secondary or tertiary structure (von Heijne & Nishikawa 1991), in which case attempting to find similarities at amino acid level becomes more elusive. For proteins with putative targeting sequences that do not have previously described targeting motifs or cleavage sites, unequivocal demonstration of subcellular localisation requires experimental methods, such as immunocytochemical localisation (reviewed by Herman 1988), rather than sequence analysis.

*Is a second polypeptide expressed from the Sat-1a sequence?*

An open reading frame of 135 amino acid residues present in the complementary strand of Sat-1a and overlapping with the opposite open reading frame encoding SAT-1 (nucleotides 1 241 - 833, Figure 3.18) encodes a conceptual polypeptide, termed SAT-X, with a calculated molecular mass of 16.1 kD. A BlastP

search with SAT-X revealed only one significant match (high score of 126 and associated  $P = 1.8 \times 10^{-10}$ ), with the hypothetical *E. coli* protein CysX whose function is unknown (Tei et al. 1990b). The 16 kD CysX polypeptide is encoded by the complementary strand in the *cysE* gene opposite to CysE, and both CysX and CysE have been shown to be expressed *in vivo* in *E. coli* (Tei et al. 1990b). An alignment of SAT-X and CysX, which have an overall identity of 33.8 % (61 % homology), shows that the two polypeptides are not similar over their entire length but that the C-terminus of SAT-X (residues 60 to 135) overlaps with the N-terminus of CysX (residues 2 to 76) (Figure 3.22).

Residues 60 to 135 of SAT-X correspond on the opposite coding strand of Sat-1a with residues 274 to 348 of SAT-1, a region containing 11 hexapeptide repeats including the consensus transferase signature (Figure 3.18). Similarly, residues 2 to 76 of CysX are opposite to residues 157 to 232 of CysE that includes several hexapeptide blocks and the transferase signature (Tei et al. 1990b). Therefore, it is feasible that homology between SAT-X and CysX might be a coincidence that has arisen because of structural residues present in the opposite strands that encode SAT-1 and CysE, respectively. This possibility was examined further by comparing SAT-X and CysX with reverse reading frames in the signature-containing region of other serine acetyltransferases (Figure 3.23). None of the other serine acetyltransferases had reverse open reading frames in this region longer than 82 residues (reverse of SAT-52). Several residues were conserved between the aligned reverse reading frames, even though the residues did not constitute expressed polypeptides. The data in Figure 3.23 support the argument that similarities in the reverse strand of serine acetyltransferases may be due to the conserved hexapeptide repeats in the coding strands.

Whether or not homology between SAT-X and CysX is coincidental, it remains to be established if SAT-X is

	1				50
<i>E. coli</i> CysX	.....	.....	.....	.....	.....
<i>A. thaliana</i> SAT-X	mkshkqqqqp	kktkfsnyii	rplgymrrlv	hsqtrnlvmr	friifttkqp
Consensus	-----	-----	-----	-----	-----
	51				100
<i>E. coli</i> CysX	.....mR	RHrLQHHgtc	aNLraapNfn	IaedfraraN	hhtftNFRVt
<i>A. thaliana</i> SAT-X	rrisnsrtrR	RHvFQHHrpr	tNLssftNrd	IpqntspssN	qnpiaNLRVp
Consensus	-----R	RH--QHH---	-NL----N--	I-----N	-----N-RV-
	101				150
<i>E. coli</i> CysX	VtTrFtRtaK	RhrLqNRyVV	fnhRRFTndd	acrVvkhhta	tnfccrvnid
<i>A. thaliana</i> SAT-X	IstlFpRssK	RnvMeNRnIV	phnRRLTdhn	srsMV.....	.....
Consensus	--T-F-R--K	R---NR--V	---RR-T---	---V-----	-----
	151			188	
<i>E. coli</i> CysX	lerhrnlvlq	kdcqcatpli	pqpvtdaigl	qgmktlqv	
<i>A. thaliana</i> SAT-X	.....	.....	.....	.....	
Consensus	-----	-----	-----	-----	

**Figure 3.22** Amino acid alignment of the hypothetical proteins SAT-X and CysX

The hypothetical *Arabidopsis thaliana* protein SAT-X deduced from a reverse reading frame of Sat-1a (Figure 3.18) was compared with the hypothetical *Escherichia coli* protein CysX (Tei et al. 1990b; SWISS-PROT entry P20343) using the GCG PileUp program. A consensus sequence shows residues conserved between SAT-X and CysX.

	1				50
A. thaliana SAT-X	.....	*lwkqhhris	ytspasitee	aertMKshkq	qqqpktkTfs
A. thaliana sat5 RF -3	.....	.....	.....	....Lfhhhq	lqerkkyTfv
A. thaliana Sat-52 RF -2	.....	*fkaiscieh	rkmfcklqiq	qssltsaply	krmvrKfftf
C. vulgaris Sat RF -2	.....	.....	...*qghngn	*nplLKw*rd	enlehsykil
Synechocystis cysE RF -1	.....	.....	.....	.....	.....
E. coli CysX	.....	.....	.....	.....	.....
Consensus	.....	.....	.....	----LK----	-----T--
	51				100
A. thaliana SAT-X	.NyIIRPLGy	mRrLVHsqtr	NLvMrfR.ii	Ftt..kQPRR	IsNsRrtRRH
A. thaliana sat5 RF -3	lNhIIRPLG*	iRcLVHsqtr	NLiMfSR.iL	Fpt..nQPRR	IsNrRrtRRH
A. thaliana Sat-52 RF -2	lydVI*PFrd	e*sMIHrfsr	tFliMnRwlL	Fps..dksRR	IpNrststRH
C. vulgaris Sat RF -2	sNyVI*PFgy	eRcMVHglSR	NileL*W.lL	Lsp..hQPCg	IahgssssWWH
Synechocystis cysE RF -1	...*fnsF**	*anyfalriW	*LTMFqRinp	FpgrmhhapW	hphhRkicWH
E. coli CysX	.....	.....	.....	.....	.....mRRH
Consensus	-N--I-PFG-	-R--VH---R	NL-MF-R--L	FP----QPRR	I-N-R--RRH
	101				150
A. thaliana SAT-X	VFQHHRPrTN	LSSFTNrdIp	QNTspSSNQn	pIaNLrvPIS	TLFprSSKrn
A. thaliana sat5 RF -3	ILnHHRP*sN	LSSLTdcyIp	QyttpSSNQH	TITNLRmPit	TLFpCSSqgH
A. thaliana Sat-52 RF -2	VnQHnRtssy	FStcTNLhIs	kNsRsSSNQt	TVTdLRmsIS	TsFtCSt*cH
C. vulgaris Sat RF -2	IneHHRscpd	LgSfSNLNiA	kNgRsStd*n	TIpNLWmPVS	ThLsCpSKed
Synechocystis cysE RF -1	Vp*HH*PrpN	adiiTDgNVt	QylstSShhH	ifpggRvtla	TFfsCapqsy
E. coli CysX	rLQHhgTcaN	LraapNFNiA	edfRaraNhh	TfTnFRvtVt	TrFtrtaKrH
Consensus	VLQHHRP--N	LSSFTNLNI-	QN-R-SSNQH	TITNLR-PIS	TLF-CSSK-H
	151				200
A. thaliana SAT-X	vMeNRNIVPh	nRRLTDHnsr	sMV*qNPgtd	FssRveINsE	gfrdSVLdqg
A. thaliana sat5 RF -3	sv*NRNIVsN	HRRlsDHhar	rMV*kNPFsd	LrsgmnIdgE	rf*ySVLdq*
A. thaliana Sat-52 RF -2	mvkd*NIVPN	HRCfSDynsg	cvV*kyPFsd	LrCWmdINsE	yigdS*v*s*
C. vulgaris Sat RF -2	mMkn*dIVay	HcRLThHytg	sMIkqNaFpn	scCRvNvNgE	disnarLkcg
Synechocystis cysE RF -1	pLvNqrIVPh	dRRFTDHyph	aMVtehaLan	LgpWmdfNag	ekagkVgh*a
E. coli CysX	rLqNRyVVfN	HRRFTnddac	rvVkhdtatn	FcCRvNIdLE	rhrnlVLqkd
Consensus	---NRNIVPN	HRR-TDH---	-MV--NPF--	L-CR--IN-E	----SVL---
	201				250
A. thaliana SAT-X	s*DFsvLsPk	LmSntMrLts	eEAFkVKEtM	nvayTSRISL	lns*.....
A. thaliana sat5 RF -3	syDFsvLfPe	gmSysMsLts	eEAFeaqEaM	nvayTSWVSF	fdcykillrg
A. thaliana Sat-52 RF -2	C*wLp*LcP*	Lv*yTMrLnS	*EAFVIEEtM	rerdTSRItN	tssa*.....
C.vulgaris Sat RF -2	C*rpt*LiPq	LvShTMsLag	*EsFVIEEaM	.....	.....
Synechocystis cysE RF -1	rkEr*pesik	tmrksMy**S	lqARItkqhl	*piagSgISL	knnpkvsn*.
E. coli CysX	CqcatpLiPq	pvtDaigLqg	mktLqV*....	.....	.....
Consensus	C-DF--L-P-	L-S-TM-L-S	-EAFV--E-M	----TSRISL	-----

**Figure 3.23** Alignment of the hypothetical proteins SAT-X and CysX with corresponding reverse translations of four serine acetyltransferases

The hypothetical proteins SAT-X from *Arabidopsis thaliana* (Figure 3.18) and CysX from *Escherichia coli* (Tei et al. 1990b; SWISS-PROT entry P20343) were aligned with amino acids derived from translating reverse reading frames of *A. thaliana* sat5 (Ruffet et al. 1995; GenBank entry Z34888) and Sat-52 (Howarth et al. 1997; GenBank entry U30298) clones, *Citruillus vulgaris* Sat clone (Saito et al. 1995; GenBank entry D85624) and the *Synechocystis* sp. cysE gene (Sakamoto & Murata 1995; GenBank entry D13777) using the GCG PileUp program. Where appropriate, reverse reading frames are identified. Stop codons are indicated by asterisks. The consensus line shows amino acid residues which are common to at least three of the aligned sequences.

expressed *in vivo*. The first methionine residue of SAT-X has 78 % identity with the plant translation start site consensus sequence AACCAATGGC, and is in an optimal context for translation initiation (Joshi 1987; Lütcke et al. 1987). Expression studies, as have been performed to confirm expression of CysX (Tei et al. 1990b), are required to investigate the existence of the hypothetical polypeptide SAT-X.

#### 3.2.3.4 Sequence analysis of Sat-80

The 5' end (Figure 3.24a) and 3' end (Figure 3.24b) of the 2.6 kb clone Sat-80 were sequenced manually in order to determine whether the clone represented another isoform of serine acetyltransferase. A BLASTN search using the 5' end sequence revealed identity (94.6 %) to the *A. thaliana* clone Sat-52 (Howarth et al. 1997). Sat-80 is shorter by 21 bp than Sat-52. It is likely that the difference between the two clones was caused by errors in the single forward reaction used to sequence Sat-80, and that Sat-80 and Sat-52 encode the same isoform of serine acetyltransferase. However, the 3' end sequence of Sat-80, which does not contain a poly(A) tail (Figure 3.24b), was not similar to the 3' end of Sat-52 (not shown). Clone Sat-80 is 1.4 kb longer than Sat-52, which indicates that Sat-80 may be a chimeric clone with two cDNA transcripts artifactually fused together during library construction.

#### 3.2.4 Nucleic acid blotting with the Sat-1 probe

##### 3.2.4.1 Southern analysis

Genomic DNA, extracted from *A. thaliana* leaves using a CTAB method (Dean et al. 1992; Chapter 2.6.1), was observed as a single intact band above the 23.1 kb  $\lambda$ /*Hind*III molecular

(a)

```
AGGAGATAGAAATCGAATCTTATCGCCGCGTTAATATGCCACCGGCGGAGAACTCCGACA
1 -----+-----+-----+-----+-----+-----+-----+ 60
tcctctatcttttagcttagaatagcggcgcaattatacgggtggcgcctcttgaggctgt

TCAATCTCCATCAAAGGAGAACTATCTTCCGTTACCCAATCCGATGAAGCAGAAGCAGC
61 -----+-----+-----+-----+-----+-----+ 120
agtttagaggtagtttcctctttgatagaaggcaatggggttaggctacttcgtcttcgtcg

GTCAGAGCGATATCTGCGGCAGTGCAGATGCGGAAGTGCCGGATTATGGACACAGATCAA
121 -----+-----+-----+-----+-----+-----+ 180
cagtctcgctatagacgcgctcacgtctacgccttcacggcctaatacctgtgtctagtt

GGCGGAAGTCGCCGTATGTGAGG
181 -----+-----+----- 203
ccgccttcagcggcatacactcc
```

(b)

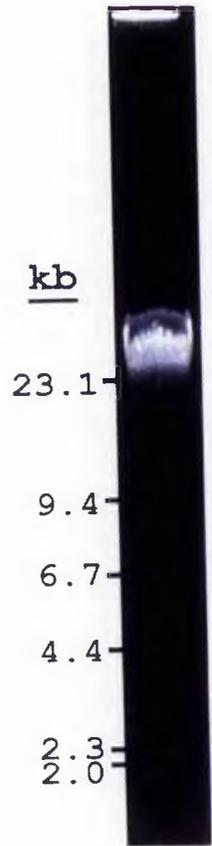
```
cttcctcatttttaggtggagagtaagcggcggaggaagaaaactaagaaacgaatgaaaa
1 -----+-----+-----+-----+-----+-----+ 60
GAAGGAGTAAAATCCACCTCTCATTGCGCCGCTCTTCTTTTGATTCTTTGCTTACTTTT

tcgtttgaagccattgctcgagagaagagctacggaggctcgactgaagctagtggactg
61 -----+-----+-----+-----+-----+-----+ 120
AGCAAACCTTCGGTAACGAGCTCTTCTTCGATGCCTCCGAGCTGACTTCGATCACCTGAC

aactgagaaaaacgggtcatagcgtctcgagttaggtggtatacttaagcaaattctga
121 -----+-----+-----+-----+-----+-----+ 177
TTGACTCTTTTGCCAGTATCGCAGAGCTCAATCCACCATATGAATTCGTTTAAGACT
```

**Figure 3.24** Nucleotide sequence of the 5' and 3' ends of clone Sat-80

*Arabidopsis thaliana* clone Sat-80 was sequenced manually using plasmid pSAT80 as template and the pYES-derived primers P1 or P2. Single strand DNA sequence obtained from the 5' end (a) and 3' end (b) of clone Sat-80 is shown in capitals, while the complementary sequence is given in small letters.



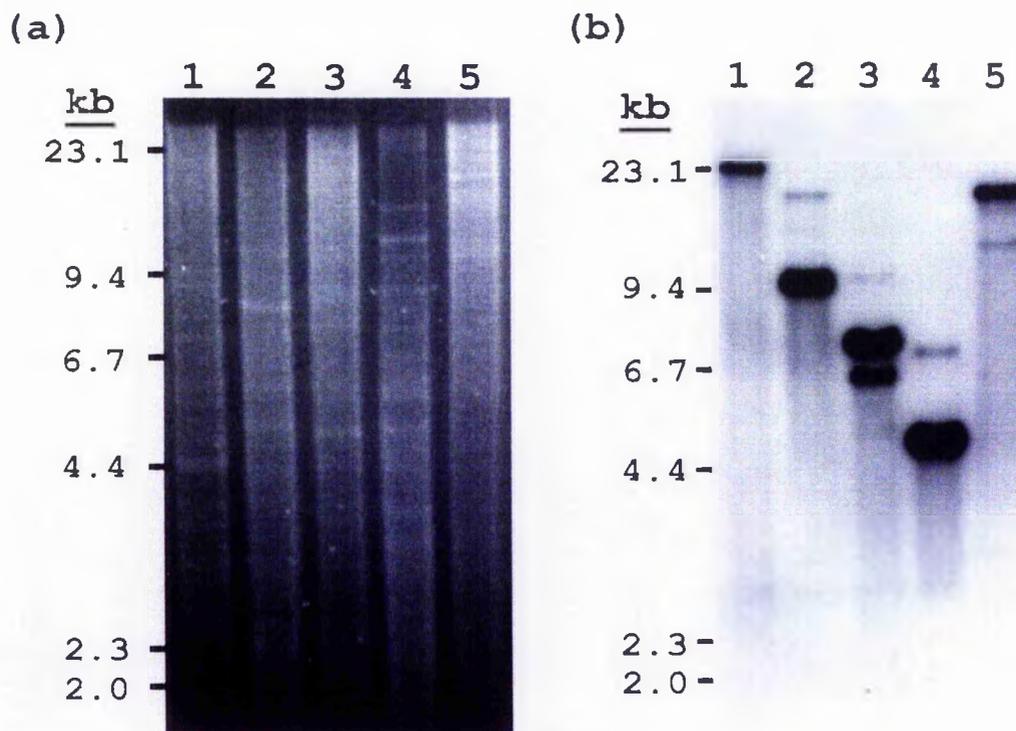
**Figure 3.25** Agarose gel analysis of *Arabidopsis thaliana* genomic DNA

1  $\mu$ g of genomic DNA extracted from *A. thaliana* leaves was fractionated through a 0.4 % ( $w/v$ ) agarose gel. Band positions of  $\lambda$ /*Hind*III molecular weight markers are shown.

weight marker (Figure 3.25). The DNA was sufficiently pure to allow restriction enzyme digestion for Southern blotting.

*A. thaliana* genomic DNA digested with the restriction enzymes *Bam*HI, *Ban*II, *Bst*EII, *Hind*III and *Pst*I, which do not cut within the Sat-1a cDNA sequence, was subjected to agarose gel electrophoresis for Southern blotting. DNA digest quantities were confirmed to be approximately equivalent (Figure 3.26a). Hybridisation using <sup>32</sup>P-labelled clone Sat-1 revealed single strongly hybridising bands and weaker secondary bands when washed under either low (not shown) or high stringency conditions (Figure 3.26b). Identical results were obtained when an internal 800 bp *Ava*II/*Hae*III fragment of Sat-1 from within the coding region was used as probe (not shown). The weaker hybridising bands (Figure 3.26b) may indicate cross-hybridisation of Sat-1 with the *A. thaliana* gene encoding SAT5 (Ruffet et al. 1995), as the cDNAs have 71.1 % identity at the nucleotide level. Ruffet et al. (1995) found that *A. thaliana* cDNAs encoding three serine acetyltransferase isoforms did not cross-hybridise with the non-homologous genes on Southern filters washed at very high stringency, but that at lower stringency pYES SAT5 did cross-hybridise to the gene corresponding to pYES SAT3. The pYES SAT3 cDNA encodes the same isoform as a recently released nucleotide sequence of pYES SAT1 (GenBank entry L78443; Ruffet et al. 1995). The pYES SAT1 cDNA probably is derived from the same gene as Sat-1a (Section 3.2.3.3), providing support for the view that the Sat-1 cDNA is cross-hybridising with a gene corresponding to the pYES SAT5 isoform at the high stringency washing conditions used here (Figure 3.26b).

A gene copy number reconstruction experiment was undertaken to determine the copies per haploid genome of the Sat-1 gene. Amount of clone Sat-1 corresponding with 1, 2, 5 and 10 gene copies per haploid genome in 5 µg of *A. thaliana* genomic DNA were calculated according to Sabelli and Shewry (1993) (Chapter 2.6.3). *A. thaliana* genomic DNA digested with *Pst*I,



**Figure 3.26** Southern blot analysis of *Arabidopsis thaliana* genomic DNA with the clone Sat-1

(a) *A. thaliana* genomic DNA was digested with the restriction endonucleases *Bam*HI (lane 1), *Ban*II (lane 2), *Bst*EII (lane 3), *Hind*III (lane 4) and *Pst*I (lane 5) and 5  $\mu$ g of each digestion was fractionated through a 0.8 % ( $v/v$ ) agarose gel.

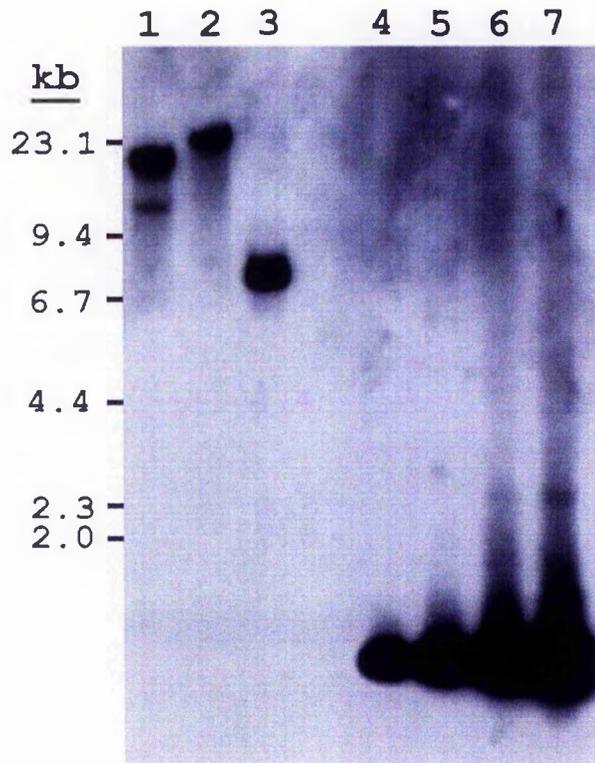
(b) Following Southern blotting of the gel shown in (a), the resultant filter was hybridised with radioactively labelled clone Sat-1 as probe and washed at high stringency. Lanes are as for (a).

Band positions of  $\lambda$ /*Hind*III molecular weight markers are shown in (a) and (b).

*Bam*HI and *Bst*EII, and gene copy equivalents were fractionated by agarose gel electrophoresis, transferred to a nylon filter by Southern blotting and hybridised with <sup>32</sup>P-labelled Sat-1. Hybridising band intensity of the digested genomic DNA was equivalent to, or slightly less than, the single gene copy reconstruction (Figure 3.27), confirming that the gene encoding SAT-1 is present as a single copy in the *A. thaliana* genome.

Southern analysis was utilised to determine whether the *A. thaliana* clone Sat-1 could detect serine acetyltransferase-like sequences present in other higher plant species. Several hybridising bands were present when <sup>32</sup>P-labelled Sat-1 was used as a probe against Southern blots of genomic DNA digests from three cultivars of *Brassica napus* (oil-seed rape), a member of the Brassicaceae family to which *A. thaliana* also belongs (Figure 3.28b). Sulphur secondary compounds have a fundamental ecological role in *B. napus*, as the species produces sulphur-rich glucosinolates as a chemical defence mechanism (Bennett & Wallsgrove 1994). *B. napus* is an important agronomic species, and biochemistry of its secondary metabolites such as glucosinolates has attracted considerable research (Schnug 1993). The presence of several bands hybridising to clone Sat-1 suggests that *B. napus* may have several serine acetyltransferases analogous to SAT-1, or at least a small gene family encoding different serine acetyltransferase isoforms. Differences in the number or position of hybridising bands could not be detected between the three cultivars, which differ in their glucosinolate content. However, incomplete digestion of genomic DNA in some lanes (Figure 3.28a) prevented a thorough analysis of cultivar variation.

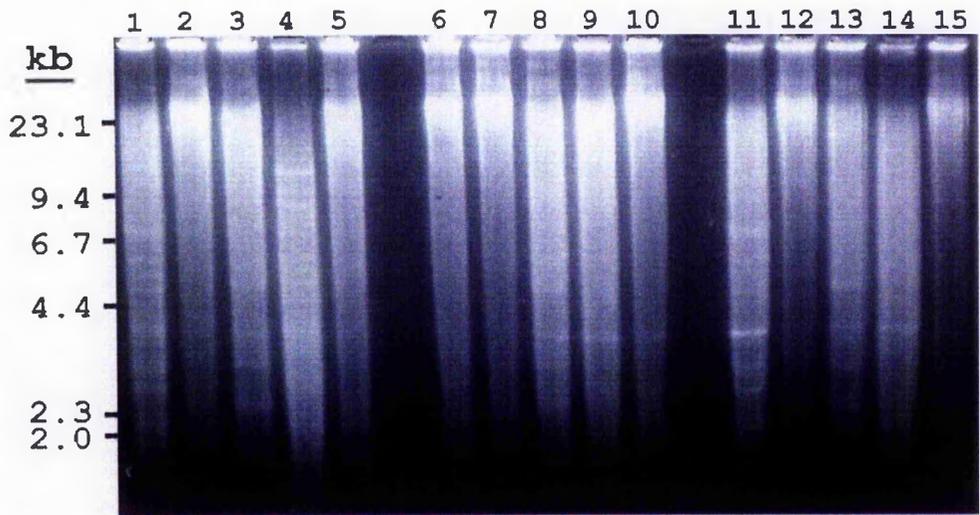
After several weeks of autoradiography exposure, the Sat-1 probe detected faint single hybridising bands in Southern blots of restriction enzyme-digested genomic DNA from *Spinacia oleracea* (Chenopodiaceae family) and *Pachyrizus*



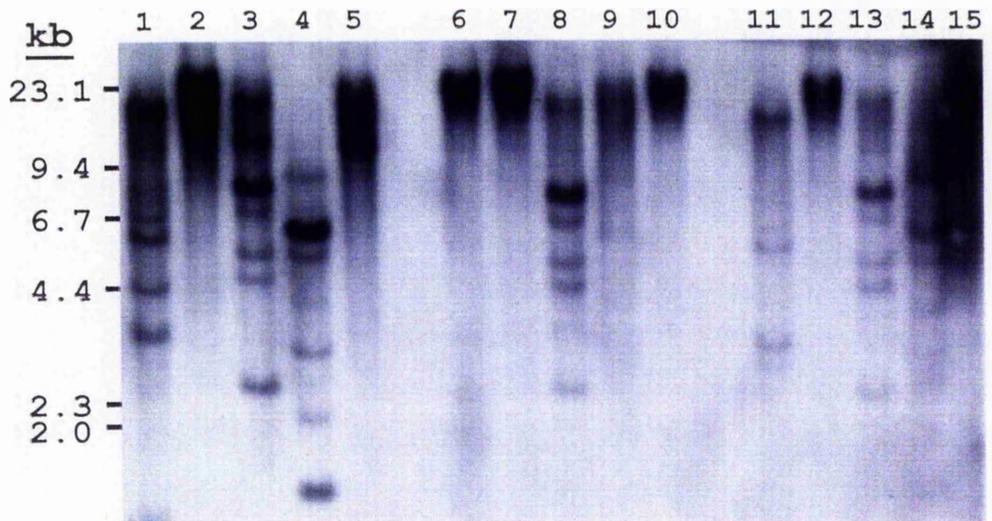
**Figure 3.27** Gene copy number reconstruction experiment with the clone Sat-1

*Arabidopsis thaliana* DNA (5  $\mu$ g/lane) digested with the restriction endonucleases *Pst*I (lane 1), *Bam*HI (lane 2) and *Bst*EII (lane 3) and clone Sat-1 DNA corresponding to 1 (70 pg, lane 4), 2 (140 pg, lane 5), 5 (350 pg, lane 6) and 10 (700 pg, lane 7) gene copy equivalents per haploid genome were fractionated through a 0.8 % ( $v/v$ ) agarose gel. Following Southern blotting of the gel, the resultant filter was hybridised with radioactively labelled clone Sat-1. Band positions of  $\lambda$ /*Hind*III molecular weight markers are shown.

(a)



(b)



**Figure 3.28** Southern blot analysis of *Brassica napus* genomic DNA with the clone Sat-1

(a) *Brassica napus* genomic DNA from the cultivars Comet (lanes 1 - 5), Falcon (lanes 6 - 10) and Pasha (lanes 11 - 15) was digested with the restriction endonucleases BamHI (lanes 1, 6 & 11), BanII (lanes 2, 7 & 12), BstEII (lanes 3, 8 & 13), HindIII (lanes 4, 9 & 14) and PstI (lanes 5, 10 & 15) and 7.5  $\mu$ g of each digestion was fractionated through a 0.8 % ( $v/v$ ) agarose gel.

(b) Following Southern blotting of the gel shown in (a), the resultant filter was hybridised with radioactively labelled clone Sat-1 as probe and washed at high stringency. Lanes are as for (a).

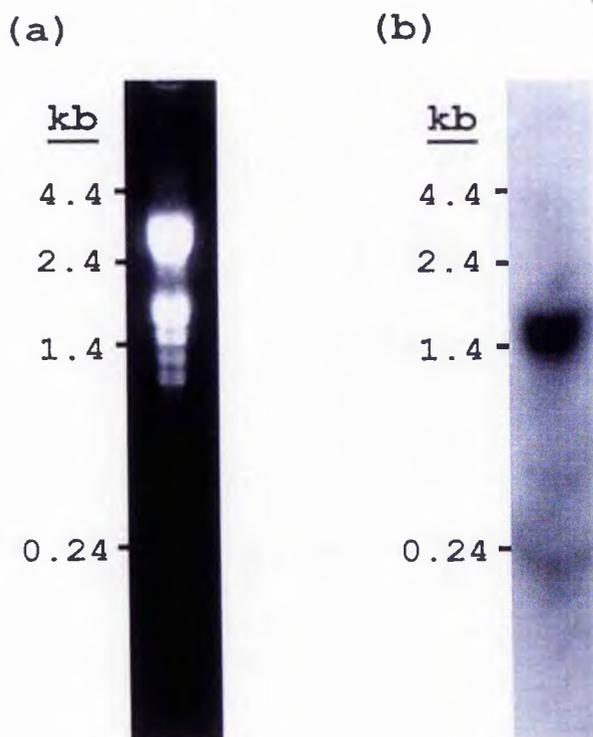
Band positions of  $\lambda$ /HindIII molecular weight markers are shown in (a) and (b).

*tuberosus* (Fabaceae family) (results not shown). Higher amounts of DNA from these species would be required to determine if less homologous sequences hybridise to Sat-1. Nevertheless, these preliminary experiments indicate the presence of SAT-1-like serine acetyltransferase genes in a range of higher plant species.

#### 3.2.4.2 Northern analysis

Examination of total RNA extracted from *A. thaliana* leaf or above-ground tissue and fractionated by denaturing gel electrophoresis (Chapter 2.7) showed discrete bands for the main 25S and 18S rRNAs and smaller organellar rRNAs (Figure 3.29a). The ratio of 25S to 18S rRNA was assessed qualitatively to be greater than two, which confirms the intactness of the RNA sample (Correa-Rotter et al. 1992).

A single message of approximately 1.5 kb was detected when Sat-1 was used as a probe against total RNA transferred onto a nylon filter (Figure 3.29b). It remains to be determined whether the message identified by Sat-1 corresponds with the Sat-1a cDNA or other Sat-1-like cDNAs that are of different length. The results in Figure 3.29b suggest that a single mRNA species encoding SAT-1 is expressed in abundance in above-ground tissue of *A. thaliana*. Bogdanova et al. (1995) report that their SAT1-6 cDNA, which is identical to Sat-1 over the coding region, reveals a message of 1.1 kb that is expressed in *A. thaliana* aerial tissue but not root tissue. Although SAT1-6 is polyadenylated 176 bp upstream from Sat-1, it is unlikely that this difference could explain the size discrepancy of 400 bp in message detected by SAT1-6 and Sat-1. An independent northern experiment has verified the size of message detected in *A. thaliana* total RNA by Sat-1 (not shown). Interestingly, expression of the transcript detected by SAT1-6 is regulated by light and sulphur availability (Bogdanova et al. 1995).



**Figure 3.29** Northern blot analysis of *Arabidopsis thaliana* total RNA with the clone Sat-1

(a) *A. thaliana* total RNA (20  $\mu$ g) was fractionated through a 1.5 % ( $v/v$ ) agarose denaturing gel.

(b) Following northern blotting of the gel shown in (a), the resultant filter was hybridised with radioactively labelled clone Sat-1 as probe. Band positions of RNA molecular weight markers are shown in (a) and (b). (Note: 9.5 and 7.5 kb markers were difficult to resolve and are not shown.)

### 3.3 Conclusions

Functional complementation of an *Escherichia coli* cysteine auxotroph has resulted in the cloning of 110 putative *Arabidopsis thaliana* serine acetyltransferase cDNAs. One of these clones, Sat-1 (and a full-length isogenic transcript, Sat-1a), has been characterised extensively. Three lines of evidence reported here provide confirmation that Sat-1a encodes an *A. thaliana* serine acetyltransferase:

(i) the auxotrophic *E. coli* mutant strain JM15, defective in the *cysE* gene encoding serine acetyltransferase, regains the ability to grow with sulphate as sole sulphur source when transfected with pSAT1, which contains the Sat-1 insert (Figure 3.1; Table 3.3);

(ii) the ability of strain SAT1, a JM15-derived strain containing the plasmid pSAT1, to grow on sulphate as sole sulphur source was associated with restoration of serine acetyltransferase activity. Crude extracts from strain SAT1 exhibited serine acetyltransferase activity that followed Michaelis-Menten kinetics, with apparent  $K_m$  values of 0.043 mM for acetyl CoA and 3.47 mM for L-serine, whereas no activity could be detected in the mutant strain JM15 (Figure 3.3; Table 3.4); and

(iii) clone Sat-1a encodes a protein, SAT-1, with significant homology to other plant and bacterial serine acetyltransferases (Figure 3.19). Putative binding sites for acetyl CoA and L-serine could be identified from conserved residues common to all identified serine acetyltransferases.

Within the 110 cDNA clones obtained here, at least three members of an *A. thaliana* multigene family encoding different serine acetyltransferase isoenzymes are present: Sat-1 (Roberts & Wray 1996), Sat-52 (Howarth et al. 1997) and Sat-53 (identified by Mr J. Howarth, pers. comm.), a pYES SAT5-like clone (Ruffet et al. 1995). It appears that these isoenzymes are associated with different subcellular locations, in agreement with the suggestion that cysteine

biosynthesis occurs in each subcellular compartment (Lunn et al. 1990; Leustek 1996). The  $\lambda$ YES system (Elledge et al. 1991) is thus a powerful tool for isolation of heterologous genes by functional complementation where a suitable screen for phenotypic rescue is available. However, the  $\lambda$ YES cDNA library appears to contain several clones which are not full-length, and selection by functional complementation will not eliminate these incomplete transcripts especially when transit peptides or N-termini of encoded polypeptides are not essential for enzyme activity.

CHAPTER 4: CHARACTERISATION OF ARABIDOPSIS THALIANA cDNAs ENCODING PAPS REDUCTASE HOMOLOGUE ISOFORMS OBTAINED BY FUNCTIONAL COMPLEMENTATION OF ESCHERICHIA COLI STRAINS JM81A (DEFICIENT IN APS KINASE) AND JM96 (DEFICIENT IN PAPS REDUCTASE)

4.1 Introduction

Whether sulphate assimilation in higher plants proceeds via a bound intermediate pathway (Chapter 1.2.3.1) or free intermediate pathway (Chapter 1.2.3.2) is an unresolved question. A major discrepancy between the two proposed pathways concerns the primary sulpho donor for sulphate reduction: APS is hypothesised to be the primary sulpho donor for the bound intermediate pathway while PAPS is advanced as the primary sulpho donor for the free intermediate pathway (Chapter 1.2.3.3). After more than two decades of intensive study, this discrepancy regarding the primary sulpho donor for plant sulphate reduction has not been resolved using biochemical methods because of the lability of enzyme preparations and ambiguity caused by spontaneous reaction of sulphur-containing compounds (Schiff & Hodson 1973; Suter et al. 1992; Schiff et al. 1993). Furthermore, the possibility that both APS and PAPS could serve as primary sulpho donors for sulphate reduction in plants cannot be excluded following critical analysis of reported data (Chapter 1.2.3.3). As has been noted by Schmidt and Jäger (1992), the most effective strategy for elucidation of the pathway(s) of sulphate assimilation in plants probably will be isolation of the genes involved and specific downregulation of their products using antisense techniques.

If sulphate assimilation in plants proceeds via the free intermediate pathway with PAPS as primary sulpho donor, then the pathway is analogous to that described in bacteria, yeasts and fungi (Kredich 1971; Peck & Lissolo 1988; Schwenn

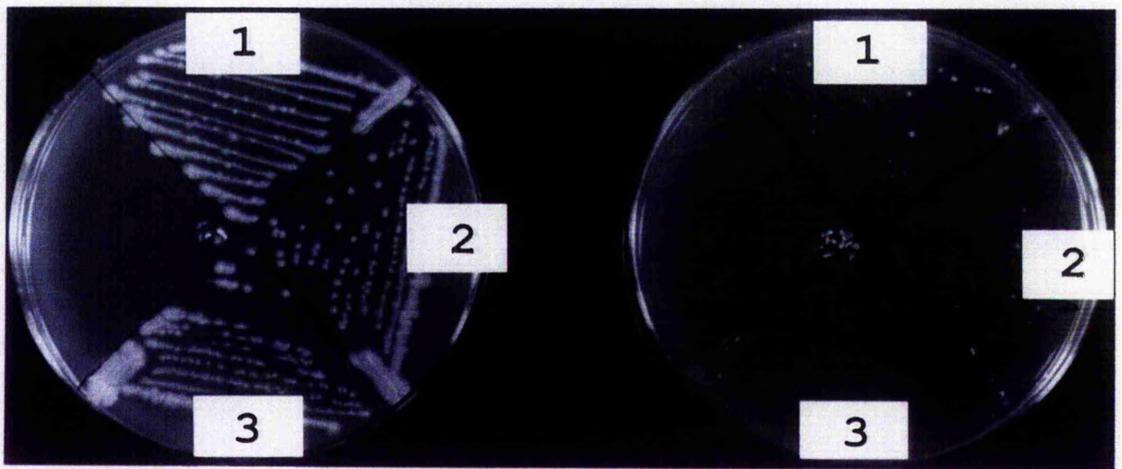
1994). Genetic evidence reveals that the free intermediate pathway is the sole pathway for sulphate assimilation in *Escherichia coli* (Jones-Mortimer 1968; Kredich 1987), and genes encoding enzymes required for all steps of sulphate assimilation in *E. coli* have been isolated (Li et al. 1987; Ostrowski et al. 1989a, b; Krone et al. 1990b, 1991). Homology between several plant and *E. coli* sulphate assimilation enzymes has been reported at the nucleotide level following hybridisation analysis (Krone et al. 1990a; Gisselmann et al. 1992) or at a functional level (Brunold 1990). It was therefore apparent that isolation of plant sulphate assimilation genes encoding enzymes required for phosphorylation of APS and reduction of PAPS would theoretically be achievable by functional complementation of appropriate *E. coli* mutant strains. This chapter describes the isolation and characterisation of *Arabidopsis thaliana* cDNA clones that are able to restore cysteine auxotrophy of the *E. coli* mutant strains JM81A (defective in APS kinase) and JM96 (defective in PAPS reductase) to prototrophy.

## 4.2 Results and Discussion

### **4.2.1 Complementation of the *Escherichia coli cysC* mutant JM81A to prototrophy with the $\lambda$ YES *Arabidopsis thaliana* cDNA library**

*Growth characteristics of Escherichia coli strain JM81A and preparation of strain JM81A/ $\lambda$ KC lysogens*

*E. coli* strain JM81A (*cysC92*, *tfr-8*), defective in APS kinase and unable to grow with sulphate as sole sulphur source (Wheldrake & Pasternak 1965), was obtained from Dr B. Bachmann, *E. coli* Genetic Stock Center, Yale University, USA. When plated onto M9 minimal medium supplemented with 0.2 % ( $w/v$ ) mannitol, growth of small numbers of colonies with sulphate as sole sulphur source indicated that strain JM81A had a low level of spontaneous reversion of the *cysC* lesion (Figure 4.1). Additional strain JM81A cultures obtained from the *E. coli* Genetic Stock Center also displayed this "leaky" phenotype. Previous work in which functional complementation was employed to isolate heterologous genes using leaky *E. coli* mutants suggested that ampicillin selection prevented the growth of prototrophs arising from spontaneous reversion (Delauney & Verma 1990). Elledge et al. (1991) argue that even if prototrophy arises from simultaneous spontaneous reversion and transformation by ampicillin resistance-conferring pYES plasmids, tight regulation of the *lac* promoter would allow identification of complemented cells and transformed spontaneous revertants. However, evidence presented in this study shows that expression of cDNA inserts in pYES need not be under control of the *lac* promoter (Chapter 3.2.2), and therefore it might not be feasible to eliminate false positives caused by spontaneous reversion if they were ampicillin resistant. Nevertheless, false positives arise infrequently with the dual selection criteria of phenotypic reversion and ampicillin resistance if spontaneous reversion rate is low (Snustad et al. 1988; Delauney & Verma



**Figure 4.1** Growth characteristics of *Escherichia coli* cysteine auxotrophic strain JM81A

Each of three colonies (1, 2 & 3) of *E. coli* strain JM81A, deficient in APS kinase, was plated onto M9 minimal medium supplemented with 0.5 mM cysteine and 0.2 % (v/v) mannitol (left) and M9 minimal medium containing sulphate as sole sulphur source and 0.2 % (v/v) mannitol (right). Plates were incubated at 37°C for three days.

1990).

$\lambda$ KC lysogens of strain JM81A were prepared prior to transfection with  $\lambda$ YES (Chapter 2.4.4). Strain JM81A cells proved more resistant to lysogenisation than strain JM15 cells (Chapter 3.2.1), even though the two strains were reported to have identical genetic characteristics except for their respective *cys* mutations (Wheldrake & Pasternak 1965; Jones-Mortimer 1968). Two separate experiments in which the input multiplicity ranged from  $2.6 \times 10^4$  -  $2.6 \times 10^6$  pfu/0.4 ml cells (approximately 0.04 - 0.0004  $\lambda$ /cell) did not yield any kanamycin-resistant lysogens. A third experiment, in which the method was changed by growing strain JM81A cells to mid-logarithmic phase before starvation in 10 mM  $MgSO_4$  for 35 min, yielded kanamycin-resistant JM81A/ $\lambda$ KC colonies at a rate of approximately 1 lysogen/ $10^6$  pfu (n=3). Three strain JM81A/ $\lambda$ KC kanamycin-resistant colonies, from which glycerol stocks were prepared, were both kanamycin-resistant and cysteine-dependent when plated onto selective minimal medium containing kanamycin and with or without cysteine (not shown). The latter observation indicated that none of the lysogens were spontaneous revertants.

*Complementation of Escherichia coli strain JM81A/ $\lambda$ KC with the  $\lambda$ YES cDNA library*

Infection of an *E. coli* strain JM81A/ $\lambda$ KC lysogen with the  $\lambda$ YES *A. thaliana* cDNA library at a rate of  $10^8$  pfu/ $10^9$  cells resulted in the growth after three days on M9 minimal medium, with sulphate as the sole sulphur source and supplemented with 50  $\mu$ g/ml ampicillin and 1 mM IPTG, of over 1000 colonies for each of two separate experiments. Assuming that all colonies contained complementing clones, efficiency of complementation can be estimated as  $3.0 \times 10^{-5}$ , which is ten-fold higher than the efficiency obtained in this study for serine acetyltransferase (Table 3.2). Two sizes of strain JM81A/ $\lambda$ KC/ $\lambda$ YES colonies were observed, and 100 each of the

larger and smaller colonies were replated onto fresh selection medium. All of the replated colonies retained both the ability to grow with sulphate as sole sulphur source and were ampicillin resistant, although several of the 100 smaller colonies did not grow well. Glycerol stocks were prepared of each of these 200 JM81A/ $\lambda$ KC/ $\lambda$ YES independent strains, with the larger strains designated APSK1 - 100 and the smaller strains APSK101 - 200.

*Retransformation of pAPSK14 and pAPSK15 into Escherichia coli strain JM81A*

*E. coli* strains APSK14 and APSK15 were two of the complemented strains which showed more vigorous growth on selection medium. Plasmids pAPSK14 and pAPSK15 were isolated from strains APSK14 and APSK15, respectively, and tested for their ability to confer cysteine-independence on the mutant strain JM81A. Initial experiments revealed that strain JM81A was difficult to transform with up to 500 ng of the plasmids pAPSK14, pAPSK15 and the empty plasmid pYES using established methods (Chapter 2.5.3). Three days after exposure to these plasmids, no colonies could be detected on M9 minimal medium supplemented with 50 or 100  $\mu$ g/ml ampicillin, 0.5 mM IPTG, 0.2 % ( $w/v$ ) mannitol and either with or without 0.5 mM cysteine. Therefore, the transformation and complementation components of the retransformation experiment were separated by first selecting transformed cells on LB medium supplemented with 50  $\mu$ g/ml ampicillin, then replating transformants onto M9 minimal medium supplemented with 0.2 % ( $w/v$ ) mannitol and either with or without 0.5 mM cysteine to test cysteine dependence. Using 500 ng of pAPSK14, pAPSK15 and pYES, transformation rate of strain JM81A was  $2.8 \times 10^3$  -  $3.6 \times 10^3$  transformants per  $\mu$ g DNA following an overnight incubation on LB medium supplemented with 50  $\mu$ g/ml ampicillin. This transformation efficiency is comparatively low (compare Chapter 3.2.1). A sample of eight colonies transformed by each of pAPSK14, pAPSK15 and pYES were

replated onto minimal selection medium. Strain JM81A/pAPSK14 and JM81A/pAPSK15 colonies were capable of comparable growth on minimal medium with or without cysteine. However, strain JM81A/pYES colonies could grow on both media types, albeit with considerably reduced ability in the absence of cysteine (results not shown).

Results from the retransformation experiment of strain JM81A using pAPSK14, pAPSK15 and pYES are not as unequivocal as those obtained for pSAT1 (Chapter 3.2.1). Low transformation ability of strain JM81A and inability to select transformants and complemented cells on minimal medium prevented direct assessment of whether plasmids pAPSK14 and pAPSK15 confer cysteine independence. Selection of transformants on complete LB medium containing ampicillin followed by selection of cysteine independence on minimal medium lacking ampicillin may have allowed sufficient spontaneous reversion of the *cysC* lesion in strain JM81A to enable colonies of strain JM81A/pYES to grow in the absence of cysteine. Differences in the ability of strains JM81A/pAPSK14, JM81A/pAPSK15 and JM81A/pYES to grow in the absence of cysteine imply still that the pAPSK plasmids conferred restoration of prototrophy in complemented APSK strains.

Following the completion of this retransformation work, an *A. thaliana* cDNA encoding an APS kinase isoform was isolated by Arz et al. (1994) using a PCR approach. To provide genetic proof for the identity of their clone, Arz and co-workers attempted to complement the APS kinase lesion of *E. coli* strain JM81A, the same strain used in this study, with the plasmid pAKSS1 containing their *A. thaliana* APS kinase cDNA. A transformation efficiency of  $9 \times 10^{-2}$  transformants per  $\mu\text{g}$  DNA was reported (Arz et al. 1994), which is about 10000-fold lower than that obtained in this study. Arz et al. (1994) attributed the inefficient rate of transformation of strain JM81A to its wild-type restriction system which could cleave exogenous DNA. However, strain JM15 which was transformed

with pSAT1 at efficiencies of up to  $10^5$  transformants per  $\mu\text{g}$  DNA also has wild-type restriction system (Chapter 3.2.1), so it remains unclear why strain JM81A is resistant to transformation.

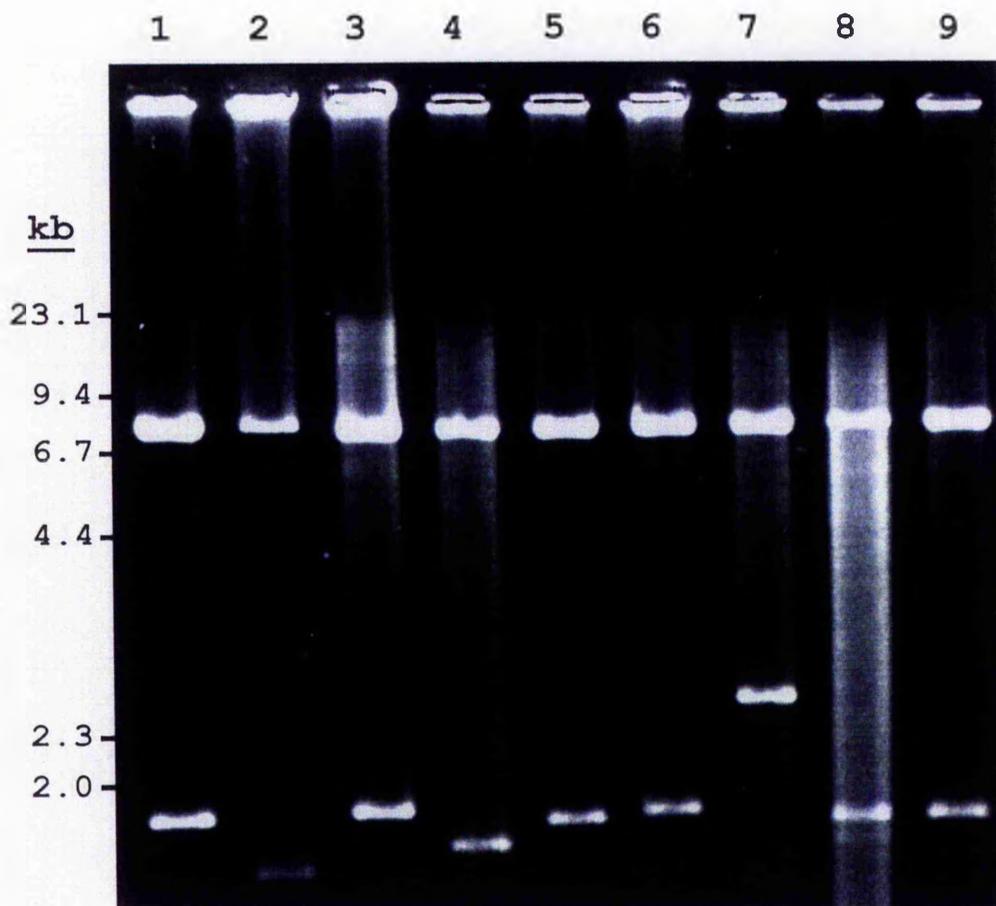
#### 4.2.2 Analysis of Apsk clones and cDNA sequence determination

Plasmids were extracted from nine APSK strains that grew well on selection medium containing ampicillin and with sulphate as sole sulphur source. The cDNA inserts within these plasmids were excised with the restriction endonuclease *EcoRI*, and insert sizes estimated following fractionation through an agarose gel (Figure 4.2). Approximate sizes ( $\pm 0.15$  kb) of the cDNA clones were: Apsk-12, 1.9 kb; Apsk-14, 1.6 kb; Apsk-15, 1.9 kb; Apsk-17, 1.7 kb; Apsk-27, 1.9 kb; Apsk-30, 1.9 kb; Apsk-57, 2.6 kb; Apsk-77, 1.9 kb; Apsk-83, 1.9 kb (Figure 4.2, lanes 1 - 9, respectively). Two clones, Apsk-14 and Apsk-15, were selected for sequence analysis as identification of the shorter clone Apsk-14 was expected to expedite characterisation of the Apsk complementing clones, while the larger clone Apsk-15 was of the same size as clones Apsk-12, Apsk-27, Apsk-30, Apsk-77 and Apsk-83.

##### 4.2.2.1 Analysis of clones Apsk-14 and Apsk-15

###### *Characterisation of clones Apsk-14 and Apsk-15*

Clones Apsk-14 and Apsk-15 were sequenced partially by primer extension using the dideoxy chain-terminating method (Sanger et al. 1977) with a Sequenase Version 2.0 sequencing kit (Chapter 2.8.3). Using sequencing conditions optimised for clones Sat-1 and Sat-2 (Chapter 3.2.3.1 and 3.2.3.2), 709 nucleotides of the 5' end coding strand and 224 nucleotides of the 3' non-coding strand were obtained from clone Apsk-14 (Figure 4.3). The longest open reading frame encoded by clone



**Figure 4.2** Electrophoretic analysis of pYES-derived plasmids extracted from complemented *Escherichia coli* JM81A strains

The pYES-derived plasmids pAPSK12 (lane 1), pAPSK14 (lane 2), pAPSK15 (lane 3), pAPSK17 (lane 4), pAPSK27 (lane 5), pAPSK30 (lane 6), pAPSK57 (lane 7), pAPSK77 (lane 8) and pAPSK83 (lane 9), extracted from corresponding complemented *E. coli* JM81A strains, were digested with the restriction enzyme *Eco*RI and fractionated through a 1.2 % (v/v) agarose gel. Linearised empty vector pYES (7.8 kb) and a cDNA insert can be seen in each lane. Band positions of  $\lambda$ /*Hind*III molecular weight markers are shown.

(a)

```
CAAGACAAAGGATTCAATGATTCTCTTTGCGGCAACAATGGTAGCAGAAATTGCAGAGGAAGTTGAAGTGGTTGAGATT
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
gtttctgtttcctaagttactaaggagaacgcggtgttaccatcgtctttaacgtctccttcaactcaccactctaa
      M I P L A A T M V A E I A E E V E V V E I
GAGGATTTGAAGAGCTTGCTAAGAAGTTAGAGAATGCTTCACCTCTTGAGATTATGGACAAAGCTCTTGAGAAATACGG
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 160
ctcctaaaacttctcgaacgattcttcaatctcttacgaagtggagaactctaatacctgtttcgagaactctttatgcc
E D F E E L A K K L E N A S P L E I M D K A L E K Y G
GAACGATATCGCCATTGCATTTAGTGGTGCAGAAGATGTTGCTCTTATTGAGTACGCTCATTTGACTGGGAGGCCATTTA
161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
cttgctatagcggtaacgtaaatcaccacgtcttctacaacgagaataactcatgagtaactgacctcgggtaaat
N D I A I A F S G A E D V A L I E Y A H L T G R P F R
GAGTATTTAGTTTGGATACAGGGAGGTTGAATCCTGAGACGTATCGGTTTTTTCGATGCGGTGGAGAAGCACTATGGGATT
241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 320
ctcataaatcaaacctatgtccctccaacttaggactctgcatagccaaaaagctacgccacctcttcgtgataccttaa
V F S L D T G R L N P E T Y R F F D A V E K H Y G I
AGGATTGAGTATATGTTTCTGATTCTGTTGAGGTTCAAGGTTTGGTTAGGAGCAAGGGATTGTTCTCTTTTTATGAGGA
321 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 400
tcctaactcatatacaaaaggactaagacaactccaagttccaaccaatcctcgttccctaacaagagaaaaatactcct
R I E Y M F P D S V E V Q G L V R S K G L F S F Y E D
TGATCATCAGGAGTGTTCGCCGTTCGAAAGGTGAGACCTTTGAGGCGTGCTCTCAAGGGTTTAAAGGCTTGGATTACTG
401 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
accagtagtctcacaacggcacaagctttccactctggaactccgcacgagagttcccaaatctccgaacctaagac
G H Q E C C R V R K V R P L R R A L K G L K A W I T G
GTCAGAGGAAAGATCAATCTCCGGGGACAAGGCTGAGATTCCGGTTGTTGAGGTTGATCCGGTGTTTGAAGGTTTGGAT
481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 560
cagtctcctttctagtttagaggccctgttccagactctaaggccaacaagtccaactaggccacaaacttccaaaccta
Q R K D Q S P G T R S E I P V V Q V D P V F E G L D
GGTGGAGTTGGTAGTTTGGTGAAGTGAATCCGGTTGCGAATGTTGAAGGGAATGATGTTTGAAGTCTTTGAGGACTAT
561 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 640
ccacctcaaccatcaaacacttcaccttagccaacgcttacaacttcccttactacaaaccttgaagaactcctgata
G G V G S L V K W N P V A N V E G N D V W N F L R T M
GGATGTTCCGGTTAACACATTGCATCCGGCAGGTTATATATCGATTGGATGTGAGCCTTGCACGAAAGC
641 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 709
cctacaaggccaattgtgtaacgtaggccgtcccatatagctaactacactcggaaacgtgctttcg
D V P V N T L H P A G Y I S I G C E P C T K
```

(b)

```
tattatatcttcgtctcgttccagattgctagaacagatgaagttggcgtttggagaatcaattcaaagctttggatgac
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
ATAATATAGAAGCAGAGCAAGGTCTAACGATCTTGTCTACTTCAACCGCAAACCTCTTAGTTAAGTTTCGAAACCTACTG
aaagtcagagataagctctgctacaaagcttgagtaattgcagattgtaagattataaagttgcatagtggtgattc
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 160
TTTCAGTCTTATTCGAGACGATGTTTCGAACCTATTAACGTCTATAACATTCTAAATATTTCAACGCTATCACAACCTAAG
ttcatgtaaatgtgtggtgtctattaaaaatcaagcctctgtttcttcttgcattaaaaaaaaa
161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 224
AAGTACATTTACACACCAGATAATTTTGTTCGGAGACAAAGAAGAAACGTAATTTTTTTTT
```

**Figure 4.3** Nucleotide sequence of the 5' and 3' end of clone Apsk-14 and deduced amino acid sequence

Approximately 60 % of the *Arabidopsis thaliana* clone Apsk-14 was sequenced manually using the plasmid pAPSK14 as template and pYES- and clone-derived primers. Single strand DNA sequence obtained from the 5' end (a) and 3' end (b) of clone Apsk-14 is shown in capitals, while the complementary sequence is given in small letters. The longest forward open reading frame, a polypeptide of 230 amino acids designated APSK-14, is shown below the encoding 5' end nucleotide sequence of clone Apsk-14 (a).

Apsk-14, designated APSK-14, is a polypeptide of 230 amino acids that is truncated at the C-terminus at the end of the partial 5' nucleotide sequence (Figure 4.3a). The first methionine residue of APSK-14 is not in a favourable context for translation initiation and the flanking sequences have only 56 % identity with the plant translation start site consensus sequence (Joshi 1987; Lütcke et al. 1987). The second methionine, residue 8 of APSK-14, has flanking sequences that are 89 % identical with the start site consensus sequence. However, no in-frame stop codons were found upstream from the longest open reading frame and the possibility that APSK-14 was truncated at the N-terminus was not excluded. The 3' end of clone Apsk-14 was terminated by a short poly(A) tract of nine residues which may have represented a polyadenylation site (Figure 4.3b). Plant polyadenylation signals and far-upstream polyadenylation elements (Hunt 1994; Wu et al. 1995) could not be detected in the 3' end of Apsk-14.

DNA sequencing of clone Apsk-15 yielded 673 nucleotides of the 5' end coding strand and 244 nucleotides of the 3' non-coding strand (Figure 4.4). The longest open reading frame encoded by clone Apsk-15, termed APSK-15, is a polypeptide of 209 amino acid residues terminated at the C-terminus by the partial DNA sequence (Figure 4.4a). Although there were no stop codons upstream from the first methionine of APSK-15, flanking sequences of this methionine had 78 % identity to the plant translation start site consensus sequence (Joshi 1987; Lütcke et al. 1987). The second methionine of APSK-15 was 100 residues downstream from the first (Figure 4.4a). Two putative AATAAT poly(A) signals (Hunt 1994; Wu et al. 1995) were present in the 3' end sequence of clone Apsk-15 (Figure 4.4b). However, neither of these putative signals were functional in clone Apsk-15, as no downstream poly(A) tracts were found.

(a)

```
AAAAAACTTAAGAACATTGTTTCGAGAGATTTGGTTTTGGAGTGAAAATGGCACTAGCAATCAACGTTTCTTCATCTTCTT
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
tttttgaattcttgtaacaaagctctctaaaccaaactcacttttaccgtgatcggttagtgcaagaagttagaagaa
M A L A I N V S S S S S
CTTCTGCGATCTCAAGCTCTAGCTTCCTTCTTCAGATCTCAAAGTAACAAAAATCGGATCATTGAGGTTATGAAATCGT
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 160
gaagacgctagagttcgagatcgaaggaagaagtctagagtttcattgtttttagcctagtaactccaataacttagca
S A I S S S S F P S S D L K V T K I G S L R L L N R
ACCAATGTCTCTGCGCTTCTCTGAGTTTGTCCGGGAAGAGATCTCCGTGAAAGCTCTTAATGTGCAATCAATTACAAA
161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
tggttacagagacgccgaagagactcaaacagccctctctaggaggcactttcagaattacacgttagttaaagttt
T N V S A A S L S L S G K R S S V K A L N V Q S I T K
GGAATCCATGTGTGCTTCTGAGTTACAGAGAAGCTAGATGTGGTGGAAAGTTGAAGACTTTGAGGAACTAGCAAAGAGAT
241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 320
ccttaggtaacaacgaagactccaatgtctctctgatctacaccaccttaacttctgaaactccttgatcgtttctcta
E S I V A S L E V T E K L D V V E V E D F E E L A K R L
TAGAGAATGCTTCTCTCTCTGAGATCATGGATAAAGCTCTTGAGAAGTTGGGAATGACATGCTATTGCCTTTAGTGA
321 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 400
atctcttaagaagaggagaactctagtagctatttcgagaactcttcaaaccttactgtaacgataacggaaatcacct
E N A S P L E I M D K A L E K F G N D I A I A F S G
GCTGAAGATGTTGCTCTGATTGAGTATGCTCATTAACTGGAAGACCTTATAGGGTATTCAGTTTGGATACAGGGAGATT
401 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
cgacttctacaacgagactaactcatacagagtaaattgacctctggaatatccataagtcacaacctatgctcctctaa
A E D V A L I E Y A H L T G R P Y R V F S L D T G R L
GAATCCAGAAACATATAGACTCTTCGATACCGTGGAGAAACATTACGGTATTCCGATGAGTATATGTTTCTCTGATGCTG
481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 560
cttaggtctttgtatctgagaagctatggcactctttgtaatgccataagcctcaatacacaaggaactacagac
N P E T Y R L F D T V E K H Y G I R I E Y M F P D A V
TTGAGGTTCAAGCTTTAGTTAGGAACAAGGGTTTGTCTCTTTCTATGAAGATGGTCACCAGGAGTGTGCCGTATTAGA
561 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 640
aactccaagttcgaatatcaatccttgttcccaacaagagatacttctaccagtggctctcacaacggcataatct
E V Q A L V R N K G L F S F Y E D G H Q E C C R I R
AAGGTGAGACCTTTAGAGGCGTGCCTTGAAGGG
641 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 673
ttccactctggaatctccgcacgcaacttccc
K V R P L E A C V E G
```

(b)

```
cgaagagttcataaatagtcacatccagctagttocagattgctagaacagatgaagaaggcgaagctgagaatctctat
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
GCTTCTCAAGTTATTATCAGGTAGGTGATCAAGGTCAAACGATCTTGCTACTTCTTCCGCTTCGACTCTTAGAGATA
ttggcgcttttagggctttgtatagctgtggttaaaaataatacgccaaaggttttgctttttgtgcttatttttacca
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 160
AACCGGAAATCCCGAAACATATCGACACCAATTTTATTAATGCGGTTCCAAACCGGAAACACGCGAATAAAAATGGT
atcttcagtgattactcctttttgctttgtgatggatgatgaacaagtttgatgatgaaagtgactcattaccctgacg
161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
TAAAAGTCCACATAATGAGGAAAAACCGGAAACACTACCTATACCTTGTTCAAACCTATACCTTTCACTGAGTAATGGGACTGC
tagc
241 ---- 244
ATCG
```

**Figure 4.4** Nucleotide sequence of the 5' and 3' end of clone Apsk-15 and deduced amino acid sequence

Approximately 50 % of the *Arabidopsis thaliana* clone Apsk-15 was sequenced manually using the plasmid pAPSK15 as template and pYES- and clone-derived primers. Single strand DNA sequence obtained from the 5' end (a) and 3' end (b) of clone Apsk-15 is shown in capitals, while the complementary sequence is given in small letters. The longest forward open reading frame, a polypeptide of 209 residues designated APSK-15, is shown below the encoding 5' end nucleotide sequence of clone Apsk-15 (a). Two putative poly (A) signals at the 3' end of clone Apsk-15 are double underlined (b).

### *Subcellular location of APSK-15*

An alignment of APSK-14 with APSK-15 revealed that the two sequences had 80.4 % identity (88.5 % homology), strongly suggesting that the two polypeptides represented isoforms of the same enzyme (Figure 4.5). APSK-15 extended 61 residues beyond APSK-14 at the N-terminus, which implies either that APSK-15 has an N-terminal transit peptide not present in the APSK-14 isoform, or that clone Apsk-14 is truncated at the 5' end. A perfect match to the semi-conserved chloroplast cleavage site motif [VI]-X-[CA]↓A for stromal processing protease (SPP) (Gavel & von Heijne 1990) was found in the N-terminal domain of APSK-15 (residues 41 - 44 in Figure 4.5). Complete identity with the SPP cleavage site was detected in only 33 % of known chloroplast transit peptides, but if present reliably predicted the chloroplast transit peptide cleavage (Gavel & von Heijne 1990; von Heijne 1992). The presence of one or more arginine residues within ten residues upstream of the cleavage site is a common feature of chloroplast transit peptides (Gavel & von Heijne 1990): APSK-15 has two arginine residues in this predicted region. Analysis of the putative 43 residue N-terminal transit peptide of APSK-15 shows a high frequency of alanine, serine and leucine residues and a low frequency of aspartate, glutamate, histidine, tryptophan and tyrosine residues, as expected for chloroplast transit peptides (von Heijne et al. 1989). Finally, comparison of the fraction of Ser ( $f_{Ser} = 0.32$ ) residues in the first 43 residues of APSK-15 with the formula  $0.07 + 1.4 \times f_{Arg}$  ( $= 0.14$ ) predicts that the N-terminal extension of APSK-15 targets to the chloroplast rather than the mitochondrion (von Heijne et al. 1989). (The convincing evidence that the N-terminal domain of APSK-15 contains a chloroplast transit peptide is contradicted by the PSORT program for protein localisation [Nakai & Kanehisa 1992] which predicts that APSK-15 is targeted to the vacuole.)

	1					50
APSK-14	.....	.....	.....	.....	.....	.....
APSK-15	malainvsss	sssaissssf	pssdlkvtki	gslrllnrtn	vsaaslslsq	
Consensus	-----	-----	-----	-----	-----	-----
	51					100
APSK-14	.....	.mIplaatmv	aEiaEevEVV	EIEDFEELAK	kLENASPLEI	
APSK-15	krssvkalnv	qsItkesiva	sEVtEk1DVV	EVEDFEELAK	rLENASPLEI	
Consensus	-----	--I-----	-E--E---VV	E-EDFEELAK	-LENASPLEI	
	101					150
APSK-14	MDKALEKYGN	DIAIAFSGAE	DVALIEY AHL	TGRPFRV FSL	DTGRLNPETY	
APSK-15	MDKALEKFGN	DIAIAFSGAE	DVALIEY AHL	TGRPYRV FSL	DTGRLNPETY	
Consensus	MDKALEK-GN	DIAIAFSGAE	DVALIEY AHL	TGRP-RV FSL	DTGRLNPETY	
	151					200
APSK-14	RFFDaVEKHY	GIRIEYMFPD	sVEVQgLVRS	KGLFSFYEDG	HQECCRVRKV	
APSK-15	RLFdtVEKHY	GIRIEYMFPD	aVEVQaLVRn	KGLFSFYEDG	HQECCRIRKV	
Consensus	R-FD-VEKHY	GIRIEYMFPD	-VEVQ-LVR-	KGLFSFYEDG	HQECCR-RKV	
	201					250
APSK-14	RPLrralkGl	kawitgqrkd	qspgtrseip	vvqvdpvfeg	ldggvgslvk	
APSK-15	RPLeacveG.	.....	.....	.....	.....	
Consensus	RPL-----G-	.....	.....	.....	.....	
	251					291
APSK-14	wnpvanvegn	dvwnflrtmd	vpvntlhpag	yisigcept	k	
APSK-15	.....	.....	.....	.....	.....	
Consensus	.....	.....	.....	.....	.....	

**Figure 4.5** Amino acid comparison of the APSK-14 and APSK-15 polypeptides

A GCG PileUp alignment of the amino acid sequences APSK-14 and APSK-15, partial polypeptides deduced from clones Apsk-14 and Apsk-15, respectively, is shown. Common residues are given in a consensus sequence beneath the polypeptides.

### Sequence comparisons with APSK-14 and APSK-15

Due to the presence of a putative transit peptide at the N-terminus of APSK-15 and the availability of additional sequence towards the C-terminus of APSK-14, the latter polypeptide was selected to run BLAST sequence similarity searches (Altschul et al. 1990) using the NCBI electronic mail server (Chapter 2.12). When the searches were performed during 1994, APSK-14 was found to have most significant identity (high score of 96; the probability that this score arose by chance in a similar size search was  $5 \times 10^{-5}$ ) with PAPS reductase from the purple bacterium *Thiocapsa roseopersicina* (Haverkamp, T., Gisselmann, G., Schwenn, J.D. 1993, unpublished SWISS-PROT entry P52672). Meaningful alignment probabilities were obtained also, in decreasing statistical significance, with the PAPS reductases from *Synechococcus* sp. (Niehaus et al. 1992), *Salmonella typhimurium* (Ostrowski et al. 1989b), *E. coli* (Krone et al. 1991) and *Saccharomyces cerevisiae* (Thomas et al. 1990). When compared using the GCG Gap program, APSK-14 had 29 % identity with PAPS reductase from *T. roseopersicina*, 23 % identity with PAPS reductase from *Synechococcus* sp., 26 % identity with *Salmonella typhimurium* PAPS reductase, 27 % identity with *Saccharomyces cerevisiae* PAPS reductase, and 24 % identity with *E. coli* PAPS reductase. A PileUp alignment of APSK-14, APSK-15 and the PAPS reductases mentioned above is shown in Figure 4.6. A consensus sequence reveals that there are not extensive regions of identity between APSK-14, APSK-15 and the other reported PAPS reductase sequences, a reason for the relatively low observed BLAST high scores. However, it has been noted previously that PAPS reductases have few and dispersed areas of amino acid similarity (Berendt et al. 1995), and it is significant that most of these conserved patches are found in APSK-14 and APSK-15. Furthermore, a highly conserved tyrosine residue which has been identified experimentally as being important for catalytic activity of PAPS reductase (Berendt et al. 1995) was found to be present

	1					60
A. thaliana APSK-14	.....	.....	.....	.....	.....	.....
A. thaliana APSK-15	malainvsss	sssaissssf	pssdlkvtki	gsrlrllnrtn	vsaaslsislsg	krsssvkalnv
E. coli CysH	.....	.....	.....	.....	.....	.....
S. typhimurium CysH	.....	.....	.....	.....	.....	.....
T. roseopersicina CysH	.....	.....	.....	.....	.....	.....
Synechococcus Par	.....	.....	.....	.....	.....	.....
S. cerevisiae MET16	.....	.....	.....	.....	.....	.....
Consensus	-----	-----	-----	-----	-----	-----
	61					120
A. thaliana APSK-14	.miplaatmv	aeiaeevev	eiedFeElak	kLen.asple	imdkaLeKY.	gndIaiafsG
A. thaliana APSK-15	qsitkesiva	sevtekladv	evedFeElak	rLen.asple	imdkaLeKF.	gndIaiafsG
E. coli CysH	...msklldl	nalnelpkvd	rilaLaEtna	eLeKld.aeg	rvawAldnLp	geyVlsssfG
S. typhimurium CysH	....sqldl	nalnelpkvd	rvmaLaEtna	qLekls.aee	rvawAlenLp	geyVlsssfG
T. roseopersicina CysH	...mskpdld	daflh....g	ddaalREtnr	rLesmp.aed	rvrwAlehlp	pqhVlsssfG
Synechococcus Par	.....m	pallp....	...sLTeina	qLadqa.atq	iiqwAateFg	sglVlstsfg
S. cerevisiae MET16	...mktyhl	nn...diiv	tqeqlDhwn	qLlkletpge	iiawsivtFp	.hlfqttafG
Consensus	-----	-----	---L-E---	---L---	---A---	---V---G
	121					180
A. thaliana APSK-14	aed.Valiey	ahltgrp..f	rVfslDTGrL	nPETYrFfda	vekhY...g	irieymfpds
A. thaliana APSK-15	aed.Valiey	ahltgrp..y	rVfslDTGrL	nPETYrLfdd	vekhY...g	irieymfpda
E. coli CysH	iqaaVsLhlv	nqirpdi...	pViltDTGyL	fPETYrFide	ltdkL...k	lnlkvyrat.
S. typhimurium CysH	iqaaVsLhlv	nqirpdi...	pViltDTGyL	fPETYqFide	ltdkL...k	lnlkvyrag.
T. roseopersicina CysH	tqsaVmLhlv	srqmpei...	pVilvDTGyL	fPETYrLvda	ltdrF...g	lnlkvyrpa.
Synechococcus Par	iqsaVmLhla	tqvqpd...	pViwiDTGyL	ptETYrFaae	lterL...k	lnlkvyqse.
S. cerevisiae MET16	ltglVtidml	sklsekyymp	ellfiDTlhh	fPqTlLkne	iekkYyqgkn	qtihvkykpdg
Consensus	---V-L---	-----	---DTG-L	---PETY-F---	-----	-----
	181					240
A. thaliana APSK-14	vevqglvrsk	glfsFYE...	.DghqeCcrv	rKVrPLrRAL	kgLka..wit	GqRkdQspgt
A. thaliana APSK-15	vevqAlvrnk	glfsFYE...	.DghqeCcri	rKVrPLeacv	eg.....	.....
E. coli CysH	.esaAwgear	ygk.LWE.qg	vEgiekYndi	nKVePMnRAL	keLnaqtwfa	GlRreQs.gs
S. typhimurium CysH	.espAwgear	ygk.LWE.qg	vEgiekYnei	nKVePMnRAL	keLkaqtwfa	GlRreQs.gs
T. roseopersicina CysH	.lspAwqear	lgr.LWE.qg	aDgierYnrl	nKIdPMeRAL	rdLdagtwfa	GlRrqQa.ns
Synechococcus Par	.ispArmeal	ygr.LWEses	vEdfnrYdqm	rKVePMnRAL	qeLgatawls	GvRrqQt.ah
S. cerevisiae MET16	ceseAdfask	ygdfLWEkdd	....dkYdyl	aKVePahRAY	keLhisavft	GrRksQg.sa
Consensus	---A---	---LWE---	---Y---	---KV-PM-RA---	---L---	G-R--Q---
	241					300
A. thaliana APSK-14	RseipVvqv	pvfegldggv	gslvkwnPva	nvegndvwnF	lrtmdvpvnt	LhpaGYisIG
A. thaliana APSK-15	.....	.....	.....	.....	.....	.....
E. coli CysH	RanlpVlaiq	.....	rgvfkvlPii	dwdnrtiyqY	lqkhgklyhp	LwdeGYlsVG
S. typhimurium CysH	RahlpVlaiq	.....	rgvfkvlPii	dwdnrtvyqY	lqkhgklyhp	LwdqGYlsVG
T. roseopersicina CysH	RaelpVlrrq	.....	dgrikfhPii	dwhrprrarY	lrrhdldpdp	LrdqGYvsIG
Synechococcus Par	Rqsmelvelk	.....	rdryairPil	gwhsrdrvqY	ltahdlpyhp	LfdqGYvtVG
S. cerevisiae MET16	RsqIsIieid	.....el	ngilkinPli	nwtfeqvky	idannvpyne	LldlGYrsIG
Consensus	<u>R---V---</u>	-----	-----P	-----Y	-----	L---GY---G
	301					342
A. thaliana APSK-14	cepctk....	.....	.....	.....	.....	.....
A. thaliana APSK-15	.....	.....	.....	.....	.....	.....
E. coli CysH	dthttrkwep	gmaee.etr	fglkrecglh	eg.....	.....	.....
S. typhimurium CysH	dthttrkwep	gmaee.etr	fglkrecglh	eg.....	.....	.....
T. roseopersicina CysH	dvhttvpllp	gmlee.etr	fgikrecglh	r.....	.....	.....
Synechococcus Par	dwhssrplqa	ddsdertrf	rglkqecglh	l.....	.....	.....
S. cerevisiae MET16	dyhstqpvke	gederagrwk	gkaktecgh	easrfaqlkq	da	.....
Consensus	-----	-----	-----	-----	-----	-----

**Figure 4.6** Amino acid sequence alignment of APSK-14, APSK-15 and PAPS reductases

The *Arabidopsis thaliana* incomplete polypeptides APSK-14 and APSK-15 were aligned with PAPS reductases from *Escherichia coli* (CysH - Krone et al. 1991; SWISS-PROT entry P17854), *Salmonella typhimurium* (CysH - Ostrowski et al. 1989b; SWISS-PROT entry P17853), *Thiocapsa roseopersicina* (CysH - Haverkamp, T., Gisselmann, G., Schwenn, J.D. 1993, unpublished SWISS-PROT entry P52672), *Synechococcus* sp. (Par - Niehaus et al. 1992; SWISS-PROT entry S28609) and *Saccharomyces cerevisiae* (MET16 - Thomas et al. 1990, with sequence corrections as described in Berendt et al. 1995; SWISS-PROT entry P18408) using the GCG PileUp program. A consensus sequence shows amino acid residues common to at least six of the seven aligned polypeptides. Two functional motifs identified by Bork and Koonin (1994) are indicated in the consensus sequence: the modified pyrophosphate-binding (PP) motif is single underlined, while the DT motif is double underlined. The eponymous aspartate and threonine amino acids of the DT motif (residues 146 and 147 of alignment) are shown in bold.

in both APSK-14 and APSK-15 (residue 296 in Figure 4.6). The APSK-15 N-terminal extension of about 60 residues beyond the PAPS reductases aligned in Figure 4.6 provides further evidence that APSK-15 has an organellar targeting peptide.

Identity of the (incomplete) polypeptides APSK-14 and APSK-15 with PAPS reductases was unexpected given that clones Apsk-14 and Apsk-15 were obtained by functional rescue of an *E. coli* APS kinase mutant. APSK-14 was found to have only 15 % identity to an APS kinase clone from *A. thaliana* which was published subsequent to the initiation of this work (Arz et al. 1994), and low identity with the APS kinases from *E. coli* (Leyh et al. 1992) and *Saccharomyces cerevisiae* (Korch et al. 1991) (not shown). Although sequence comparisons suggested that the partial APSK-14 and APSK-15 polypeptides were analogous to PAPS reductases rather than APS kinases, genetic evidence from the functional complementation experiment suggested that APSK-14 and APSK-15 either were capable of APS kinase activity, or catalysed another reaction which would circumvent the requirement for APS kinase activity during assimilation of sulphate in *E. coli*.

The BLAST search using APSK-14 as a query sequence also identified matches of low statistical significance with the ATP sulphurylase small subunit proteins CysD from *E. coli* (Leyh et al. 1992) and NodP from *Rhizobium meliloti* (Cervantes et al. 1989). Based on sequence similarity of two functional motifs, Bork and Koonin (1994) have proposed that PAPS reductases may have evolved from ATP sulphurylases. A pyrophosphate-binding (PP) motif of 28 residues found in groups of  $\alpha$ - $\beta$  phosphate bond-hydrolysing ATPases (including ATP sulphurylases, GMP synthetases, asparagine synthases and argininosuccinate synthetases) is present in a modified form in PAPS reductases (Bork & Koonin 1994). The modified PP motif found in PAPS reductases, of 27 or 28 residues, has the same predicted secondary structure as the PP motif, ie. a  $\beta$ -sheet followed by a turn and an  $\alpha$ -helix. In the modified PP

motif region (residues 111 to 137 of the alignment in Figure 4.6), secondary structure prediction using both Chou-Fasman (Chou & Fasman 1978) and Garnier-Osguthorpe-Robson (Garnier et al. 1978) algorithms in the GCG PeptideStructure program reveals that APSK-14 and APSK-15 maintain this  $\beta$ -sheet, turn,  $\alpha$ -helix feature. However, the most conserved region between ATP sulphurylases and PAPS reductases lies downstream from the modified PP motif after an invariable aspartate-threonine (DT) dipeptide (Bork & Koonin 1994). The "DT motif" has 16 conserved amino acids and is composed of three  $\beta$ -strands and seven  $\alpha$ -helices over approximately 140 residues. The APSK-14 polypeptide was found to have 12 (out of a possible 15) of these DT motif conserved residues in the region of overlap with PAPS reductase (residues 141 - 165, 205 - 250 and 261 - 330 of alignment in Figure 4.6; ATP sulphurylase sequences not shown), and a predicted five (out of a possible eight) of the conserved secondary structures (not shown).

What is the functional or structural significance of the PP, modified PP and DT motifs that are found with some variation in APSK-14 and APSK-15, PAPS reductases and ATP sulphurylases? Bork and Koonin (1994) deduced that in ATP sulphurylases the PP motif is involved in binding the phosphate moiety of ATP while the DT motif is required for sulphate binding and transfer to AMP. While the DT motif would serve an analogous sulphate-binding role in PAPS reductases, the modified PP motif in these enzymes has been altered to accommodate a different but related substrate, PAPS (Bork & Koonin 1994). During this divergence of function, or perhaps subsequently, it is feasible that the PP motif could have been adapted to accept APS as substrate. It can be speculated that APSK-14 and APSK-15 polypeptides have a form of the PP motif that will accept APS as substrate and the enzymes would thus be capable either of reducing the sulpho group of APS to sulphite directly (an APS reductase) or of transferring the sulpho group of APS to a carrier molecule subsequent to reduction (an APS sulphotransferase).

Theoretically, it could be argued that both APS reductase and APS sulphotransferase enzymes would be capable of conferring prototrophy to *E. coli* strain JM81A deficient in APS kinase: the former enzyme by reducing the sulpho group of APS directly to sulphite (bypassing reactions 1 and 2 of the free intermediate pathway outlined in Chapter 1.2.3.2) and the latter enzyme by transfer of the sulpho group of APS to a thiol molecule followed by formation of free sulphite by non-enzymatic sulphitolysis (reactions 3 and 4 of the bound intermediate pathway described in Chapter 1.2.3.1).

#### 4.2.2.2 Are other enzyme types present amongst the Apsk clones?

*A. thaliana* cDNAs encoding APS kinase have been isolated (Arz et al. 1994; Jain & Leustek 1994), and one of these has been shown to complement the APS kinase deficiency in *E. coli* strain JM81A (Arz et al. 1994). Hence it is feasible that cDNA clones encoding APS kinase are amongst the 200 Apsk clones isolated in this study. *A. thaliana* APS kinase cDNA clones range in size from 1.1 - 1.2 kb (Arz et al. 1994; Jain & Leustek 1994), smaller than any of the nine Apsk clones sized here (Figure 4.2). As stated previously, these nine clones were isolated from a class of colonies that grew to a large size on minimal selection medium, and this phenotypic selection may have favoured PAPS reductase-like Apsk clones. Clearly, analysis of the remaining Apsk clones is required to investigate whether both PAPS reductase-like and APS kinase-like enzymes complemented *E. coli* strain JM81A in this study.

#### 4.2.3 Complementation of the *Escherichia coli* *cysC* mutant JM96 to prototrophy with the $\lambda$ YES *Arabidopsis thaliana* cDNA library

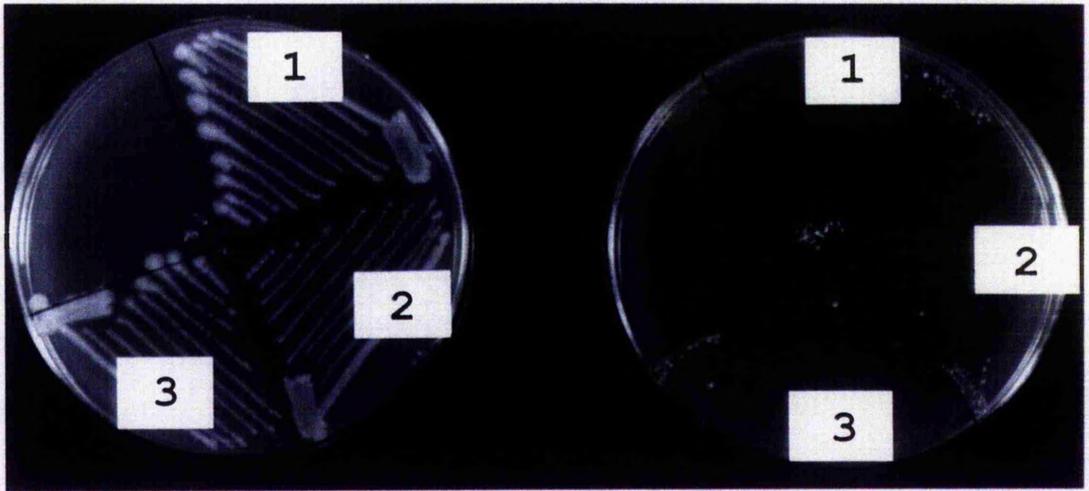
In addition to functional rescue of an *E. coli* APS kinase

mutant by *A. thaliana* cDNAs which encode enzymes that have similarity to PAPS reductase, a further complementation experiment in which an *E. coli* PAPS reductase mutant was transfected with the *A. thaliana*  $\lambda$ YES cDNA expression library was performed. The results of the latter experiment are described next.

#### *Growth characteristics of Escherichia coli strain JM96 and preparation of strain JM96/ $\lambda$ KC lysogens*

*E. coli* strain JM96 (*thr-1*, *leuB6*, *fhuA2*, *lacY1*, *supE44*, *gal-6*,  $\lambda^-$ , *trp-1*, *hisG1*, *cysH56*, *galP63*, *gltB31*, *rpsL9*, *malT1*( $\lambda^R$ ), *xyl-7*, *mtl-2*, *argH1*, *thi-1*), obtained from Dr B. Bachmann, *E. coli* Genetic Stock Centre, Yale University, USA, is defective in PAPS reductase and shows no significant growth with sulphate as sole sulphur source (Jones-Mortimer 1968, 1973). Strain JM96 was plated onto M9 minimal medium supplemented with 0.2 % ( $w/v$ ) mannitol, 40  $\mu$ g/ml of the amino acids threonine, leucine, tryptophan, histidine and arginine, 4  $\mu$ g/ml of the amino acid thiamine and with or without 0.5 mM cysteine to examine the strain's growth characteristics (Figure 4.7). A small number of colonies grew on medium containing sulphate as sole sulphur source, indicating that strain JM96 had a low level of spontaneous reversion of the *cysH* mutation. The "leaky" phenotype of strain JM96 was observed also in independent stock cultures obtained subsequently from the *E. coli* Genetic Stock Center. As with strain JM81A, it was anticipated that ampicillin selection could prevent prototrophs arising from spontaneous reversion of strain JM96 during functional complementation experiments.

Strain JM96 was lysogenised with  $\lambda$ KC (Chapter 2.4.4), and kanamycin-resistant lysogens were selected on lambda agar medium supplemented with 40  $\mu$ g/ml kanamycin. At an infection rate of  $2.6 \times 10^5$  pfu/0.4 ml cells, a yield of approximately 1 lysogen/ $10^5$  pfu was obtained, which is comparable to the lysogeny rate observed for JM15 (Chapter 3.2.1). Two strain



**Figure 4.7** Growth characteristics of *Escherichia coli* cysteine auxotrophic strain JM96

Each of three colonies (1, 2 & 3) of *E. coli* strain JM96, deficient in PAPS reductase, was plated onto M9 minimal medium supplemented with 0.5 mM cysteine, required non-sulphur amino acids and 0.2 % (v/v) mannitol (left) and M9 minimal medium containing sulphate as sole sulphur source, required non-sulphur amino acids and 0.2 % (v/v) mannitol (right). Plates were incubated at 37°C for four days.

JM96/ $\lambda$ KC kanamycin-resistant lysogens were replated onto M9 minimal medium supplemented with 40  $\mu$ g/ml kanamycin and with or without 0.5 mM cysteine. After a period of three days, small colonies of both JM96/ $\lambda$ KC lysogens grew in the presence of cysteine and kanamycin, but no growth was observed on medium with sulphate as sole sulphur source (not shown). Glycerol stocks of these JM96/ $\lambda$ KC lysogens were prepared for subsequent transfection with the  $\lambda$ YES cDNA library.

#### *Complementation of Escherichia coli strain JM96/ $\lambda$ KC with the $\lambda$ YES cDNA library*

An *E. coli* strain JM96/ $\lambda$ KC lysogen was infected with the  $\lambda$ YES *A. thaliana* cDNA library at a rate of  $10^8$  pfu/ $10^9$  cells. After four days of growth on M9 minimal medium with sulphate as sole sulphur source and supplemented with appropriate (non-sulphur) amino acids, 50  $\mu$ g/ml ampicillin and 1 mM IPTG, a total of 70 colonies were detected from two independent transfections. Some of these colonies had the appearance of small "satellite" colonies that grow adjacent to ampicillin-metabolising cells but which are themselves not ampicillin-resistant (Sambrook et al. 1989). From 63 colonies which were replated onto fresh selection medium, 57 colonies that manifested growth after four days were stored as glycerol stocks and designated PAPSR clones. The 57 PAPSR strains which appeared to be both ampicillin-resistant and cysteine-independent were categorised into three classes according to their growth characteristics on selection medium: 16 strains had large colonies (PAPSR1, 6, 7, 19, 20, 26, 27, 29, 34, 37, 39, 41, 42, 43, 44, 49); 8 strains had small colonies (PAPSR9, 10, 12, 28, 30, 31, 36, 52); and the 33 remaining strains had very small colonies which did not grow beyond the initial inoculum. The efficiency of complementation of strain JM96 by the  $\lambda$ YES library can be estimated at between  $2.2 \times 10^{-7}$  and  $4.5 \times 10^{-7}$ , which is lower than that obtained for both strain JM15 (Chapter 3.2.1) and strain JM81A (Section 4.2.1).

It was suspected that the third class of PAPS<sub>R</sub> strains that did not grow well on minimal medium supplemented with ampicillin may have been false positives. Longer incubations on selection medium allow ampicillin degradation and increase the possibility of obtaining revertant colonies that lack plasmids (Senecoff & Meagher 1993). For this reason, several researchers have picked clones after only 12 - 24 h of growth on selection medium (Delauney & Verma 1990; Senecoff & Meagher 1993), but others found it necessary to screen for at least three days (Snustad et al. 1988). Clearly, growth characteristics of *E. coli* strains employed will determine the minimum duration of the selection screen. However, it has been reported that even when complementing clones were selected after an incubation of 24 h and were shown to contain a plasmid, some of the clones were false positives with their plasmids unable to retransform the mutant to prototrophy (Van Camp et al. 1990). Such false positives could arise by simultaneous spontaneous reversion and transfection/transformation, especially with mutant strains that have high spontaneous reversion rates.

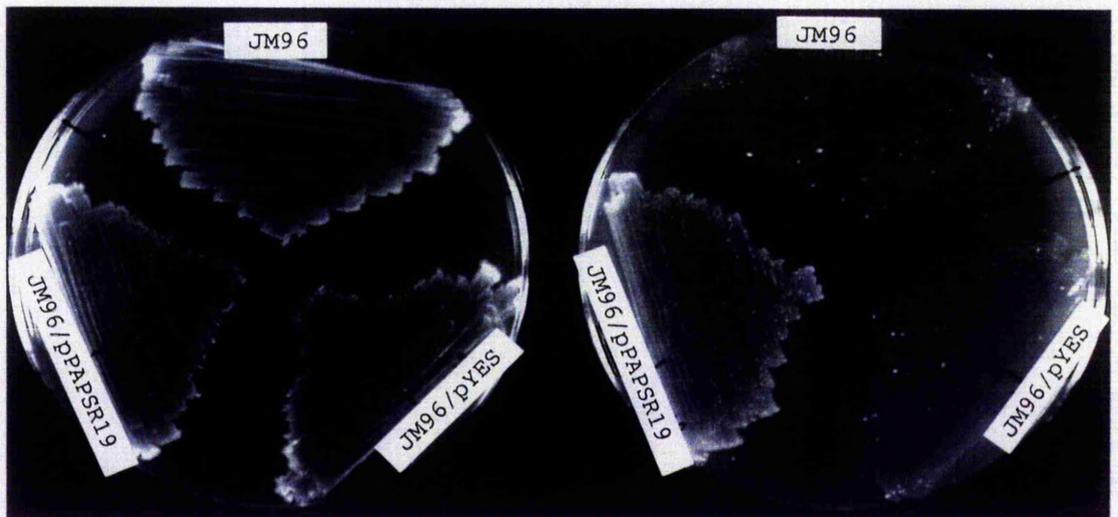
*Retransformation of pPAPS<sub>R</sub>19 and pPAPS<sub>R</sub>26 into Escherichia coli strain JM96*

Plasmids pPAPS<sub>R</sub>19 and pPAPS<sub>R</sub>26 were isolated from strains PAPS<sub>R</sub>19 and PAPS<sub>R</sub>26, respectively, which grew well on M9 minimal medium containing sulphate as sole sulphur source and supplemented with ampicillin and appropriate additives. These plasmids were tested for their ability to confer cysteine-independence when transformed into the mutant strain JM96. In the first experiment, strain JM96 cells transformed with 50 ng of plasmids pPAPS<sub>R</sub>19, pPAPS<sub>R</sub>26 and pYES failed to grow after four days on M9 minimal medium supplemented with 100 µg/ml ampicillin, 0.5 mM IPTG, appropriate non-sulphur amino acids and either with or without 0.5 mM cysteine. As with strain JM81A (Section 4.2.1), it appeared that it was difficult to obtain strain JM96 transformants with initial

selection on minimal medium. Therefore, transformation and complementation experiments were separated, with primary selection of plasmid-containing cells on complete LB medium supplemented with 100  $\mu\text{g/ml}$  ampicillin, followed by replating transformants onto M9 minimal medium supplemented with 0.2 % ( $\text{w/v}$ ) glucose, appropriate non-sulphur amino acids and either with or without 0.5 mM cysteine. Transformation of strain JM96 cells with 500 ng of pPAPSR19, pPAPSR26 or pYES resulted in the appearance of ampicillin-resistant colonies following an overnight incubation on LB medium supplemented with 100  $\mu\text{g/ml}$  ampicillin at a rate of  $1.6 \times 10^3$  -  $1.3 \times 10^4$  transformants per  $\mu\text{g}$  DNA. This low rate of transformation of strain JM96 by pPAPSR plasmids was comparable to that obtained for strain JM81A by pAPSK plasmids (Section 4.2.1). Four transformants from each transformation experiment were then replated onto minimal selection medium. After four days, strain JM96/pPAPSR19 grew well either in the presence of 0.5 mM cysteine or with sulphate as the sole sulphur source (Figure 4.8). Strain JM96 and strain JM96/pYES could not grow on minimal medium containing sulphate as the sole sulphur source, although revertant colonies were observed (Figure 4.8). Similar results were observed for pPAPSR26 (not shown), which verified that both pPAPSR19 and pPAPSR26 conferred restoration of prototrophy to the PAPS reductase-deficient *E. coli* strain JM96.

#### **4.2.4 Analysis of Papsr clones and cDNA sequence determination**

Plasmids derived from pYES were extracted from 47 PAPSR strains grown overnight in liquid LB medium supplemented with 50  $\mu\text{g/ml}$  ampicillin, whereas growth of the remaining ten PAPSR strains was inhibited by ampicillin and they were assumed to contain no plasmids. Papsr cDNA clones were excised from the plasmids by digestion with the restriction endonuclease *EcoRI* and approximate size of the clones was



**Figure 4.8** Growth characteristics of *Escherichia coli* strains JM96, JM96/pYES and JM96/pPAPSR19

*E. coli* strain JM96, deficient in PAPS reductase, strain JM96 transformed with the empty vector pYES (JM96/pYES) and strain JM96 transformed with pPAPSR19 (JM96/pPAPSR19) were plated onto M9 minimal medium containing sulphate as sole sulphur source, required non-sulphur amino acids and 0.2 % (w/v) mannitol (right) and similar medium supplemented with 0.5 mM cysteine (left). Plates were incubated at 37°C for four days.

calculated by comparison with molecular weight markers following fractionation by agarose gel electrophoresis (Table 4.1). An agarose gel showing *EcoRI* restriction products of nine pYES-derived plasmids isolated from the large colony class of PAPSR strains is shown in Figure 4.9. Papsr clones isolated from the larger strains ranged in size from 1.3 - 4.0 kb, although most (75 %) of the clones were 1.8 - 1.9 kb (Table 4.1). This isometry of clones implied non-random selection of cDNAs from the same gene or gene family. In contrast, Papsr clones derived from the small and very small strains were heterogenous in size: five plasmids appeared to have no cDNA insert, while the remaining plasmids contained cDNAs ranging in size from 0.4 - 2.2 kb. These observations invited the conclusion that only the large class of PAPSR strains were derived by functional rescue, while the remainder were transfected revertants. Retransformation of all pPAPSR plasmids into the *E. coli* mutant strain JM96 would be required to test the latter conclusion.

Two Papsr clones derived from PAPS strains which had large colonies were selected for further analysis: Papsr19, which was able to retransform mutant strain JM96 (Section 4.2.3), and Papsr20, which was the shortest clone.

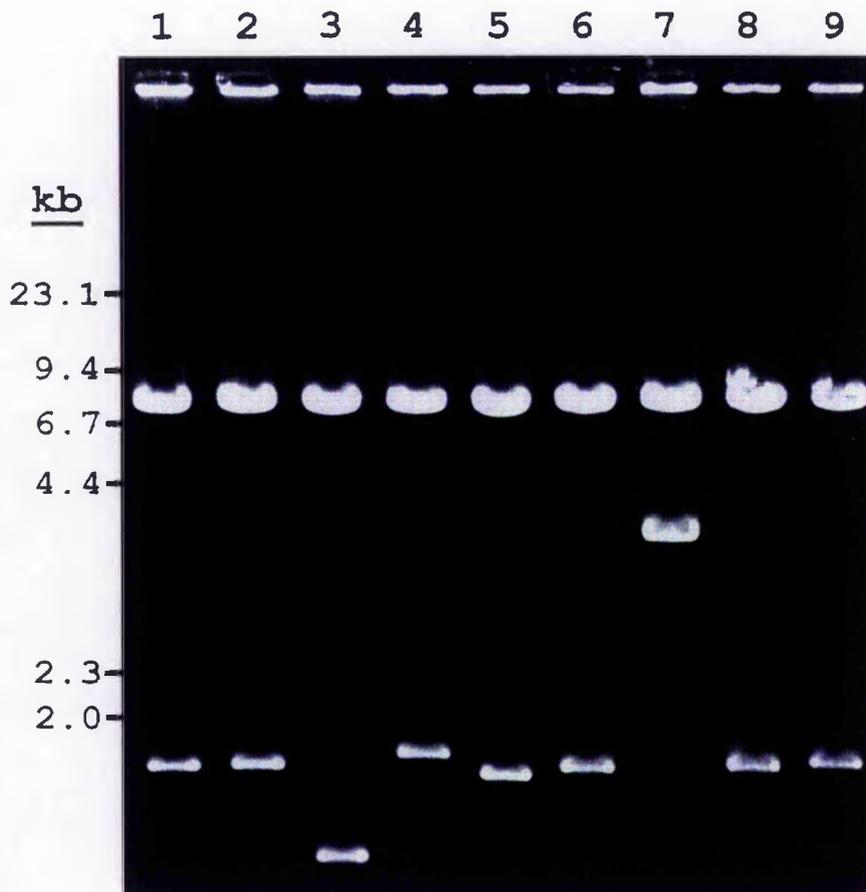
#### 4.2.4.1 Analysis of clones Papsr-19 and Papsr-20

Plasmids pPAPSR19 and pPAPSR20 isolated by Qiagen purification columns from strains PAPSR19 and PAPSR20, respectively, were used as template for DNA sequencing reactions with Sequenase Version 2.0 DNA polymerase (Chapter 2.8.3). Initial reactions with the pYES-derived forward (P2) and reverse (P1) primers did not yield clear results as sequence bands were faint, there was high background, and several sites of banding across all four lanes were evident. Alteration in template concentrations from 2 - 12.5  $\mu$ g plasmid DNA did not substantially alter the pattern obtained

**Table 4.1** Papsr clone size determined by agarose gel fractionation of *EcoRI*-digested plasmids

<u>Phenotypic class</u>	<u>Papsr clone</u>	<u>Insert size (kb)</u>	<u>Phenotypic class</u>	<u>Papsr clone</u>	<u>Insert size (kb)</u>	
Large colonies:	1	1.8	Very small colonies:	2	1.5	
	6	1.6		3	1.0	
	7	1.9		4	1.4	
	19	1.8		5	0.7	
	20	1.3		8	0	
	26	1.9		11	1.0	
	27	1.9		14	2.2	
	29	1.8		16	0	
	34	1.6		17	1.2	
	37	1.8		18	1.2	
	39	4.0		21	0.4	
	41	1.8		23	1.7	
	42	1.9		25	1.0	
	43	1.8		32	1.1	
	44	1.9		33	0.7	
	49	1.9		38	0.7	
	Small colonies:	9		0	40	1.1
		10		1.8	45	0
		12		1.1	46	0.9
28		0.6	47	1.0		
30		0.6	51	1.5		
31		1.7	54	1.0		
36		1.2	56	0		
			57	0.5		

PAPSR strains, characterised into three classes according to growth on minimal selection medium (Section 4.2.3), were grown overnight in LB medium supplemented with 50  $\mu$ g/ml ampicillin. Ten strains (PAPSR13, 15, 22, 24, 35, 48, 50, 52, 53 and 55) did not grow as liquid cultures in two separate attempts. Plasmids isolated from PAPSR cultures were digested with the restriction endonuclease *EcoRI* to excise Papsr cDNA clones. Digestion products were fractionated by agarose gel electrophoresis and Papsr clone sizes estimated by comparison with  $\lambda$ /*HindIII* molecular weight markers.



**Figure 4.9** Electrophoretic analysis of pYES-derived plasmids extracted from complemented *Escherichia coli* JM96 strains

The pYES-derived plasmids pPAPSR1 (lane 1), pPAPSR19 (lane 2), pPAPSR20 (lane 3), pPAPSR26 (lane 4), pPAPSR29 (lane 5), pPAPSR37 (lane 6), pPAPSR39 (lane 7), pPAPSR42 (lane 8) and pPAPSR49 (lane 9), extracted from corresponding complemented *E. coli* JM96 strains, were digested with the restriction enzyme *Eco*RI and fractionated through a 1.2 % (w/v) agarose gel. Linearised empty vector pYES (7.8 kb) and a cDNA insert can be seen in each lane. Band positions of  $\lambda$ /*Hind*III molecular weight markers are shown.

(results not shown). Control reactions confirmed that the sequencing reagents were functioning well, implicating purity of the plasmid templates as a likely cause of the poor sequencing results. Quality of isolated plasmids is dependent largely on *E. coli* host strain (Engelbrecht et al. 1991), and inability to obtain pPAPSR19 and pPAPSR20 plasmids amenable to sequencing may have been due to high nuclease activity or carbohydrate content of the strain JM96 progenitor, strain PA309 (Jones-Mortimer 1968). Therefore, transformation of pPAPSR19 and pPAPSR20 into a preferred *E. coli* strain such as DH5 $\alpha$  or XL1-Blue may have allowed preparation of better quality template for sequencing (Engelbrecht et al. 1991). However, the sequence of clones Papsr-19 and Papsr-20 that was obtained provided sufficient information to establish preliminary identity of the clones.

#### *Characterisation of clones Papsr-19 and Papsr-20*

Single DNA sequencing reactions using forward (P2) and reverse (P1) primers yielded 241 bp of the 5' end (Figure 4.10a) and 150 nucleotides of the 3' end (Figure 4.10b) of clone Papsr-19. The 3' end sequence of Papsr-19 does not contain a poly(A) site, polyadenylation signals, or substantial open reading frames. The longest open reading frame at the 5' end of Papsr-19 encodes a polypeptide of 72 amino acids (termed PAPSR-19; Figure 4.10a), with the putative initiator methionine being in a favourable context for translation and having six of the nine nucleotides which define the plant translation start site consensus sequence (Joshi 1987; Lütcke et al. 1987). No in-frame stop codons were detected upstream from the putative translation start site, so it cannot be assumed that the longest open reading frame of Papsr-19 is full length. However, the first 60 amino acids of the PAPSR-19 polypeptide are enriched in serine (26.7 %), leucine (11.7 %) and valine (11.7 %) residues, suggesting an N-terminal organellar targeting peptide. This high proportion of leucine residues is characteristic of

(a)

```
AAAAATCCAATTTTGGCTGTGAAGATGGCAATGTCTGTAATGTTTCTTCTTCTTCGTCCTTCTGGGATCATAAACTCTCG
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
tttttaggttaaaaacgacacttctacogttacagacatttacaagaagaagaagcagaagaccctagtagtatttgagagc
      M A M S V N V S S S S S S G I I N S R

TTTCGGTGTTCATTGGAGCCAAAAGTTTCGCAAATTGGTTTCGTTGAGGTTATTGGATCGTGTTCATGTTGCTCCTGAGT
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 160
aaagccaaaagtaacctcggttttcaagcgtttaaccaagcaactccaataacctagcacaagtacacaggaggactca
      F G V S L E P K V S Q I G S L R L L D R V H V A P E S

CTCTGAATCTATCTGGGAAGCCATCATCATCTGTTAAACCTTTAAACGCTGAACCAAGACAAAGGATTCAATGATTCTC
161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
gagacttagatagacccttcggttagtagtagacaatttggaaatttgcgacttggtttctgtttcctaagttactaagag
      L N L S G K P S S S V K P L N A E P K T K D S M I L

T
241 - 241
a
```

(b)

```
agaatcaattcaaagccttggatcacaagtcagagataagctctgctacaagccttgagtaattgcagtagtattgtaagat
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
TCTTAGTTAAGTTTCGAAACCTAGTGTTCAGTCTCTATTTCGAGACGATGTTTCGAACTCATTAAACGTCATAACATTCTA

ttataaagttgcgatagtggtgattcctcatgtaaagtgtggtggtctattaaaatcaagcctctggtgt
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 150
AATATTTCAACGCTATCACAACTAAGAAGTACATTTACACACCACAGATAATTTTAGTTTCGGAGACAACA
```

**Figure 4.10** Nucleotide sequence of the 5' and 3' end of clone Papsr-19 and deduced amino acid sequence

*Arabidopsis thaliana* clone Papsr-19 was sequenced manually using the plasmid pPAPSR19 as template and pYES-derived primers P1 or P2. Single strand DNA sequence obtained from the 5' end (a) and 3' end (b) of clone Papsr-19 is shown in capitals, while the complementary sequence is given in small letters. The longest forward open reading frame, a polypeptide of 72 amino acids designated PAPSR-19, is shown below the encoding 5' end nucleotide sequence of clone Papsr-19 (a).

mitochondrial rather than chloroplast targeting peptides (von Heijne et al. 1989), yet the N-terminal domain of PAPSR-19 is not enriched in arginine residues and the relative proportion of serine and arginine residues predicts a chloroplastic transit peptide (von Heijne et al. 1989). Consensus organellar peptide cleavage sites (Gavel & von Heijne 1990; von Heijne 1992) could not be detected. The PSORT program for protein localisation (Nakai & Kanehisa 1992) predicted the following targeting locations for PAPSR-19: chloroplast thylakoid space (certainty = 0.96); mitochondrial intermembrane space (certainty = 0.86); chloroplast stroma (certainty = 0.78); or mitochondrial matrix space (certainty = 0.74). Assignment of a specific destination for the putative organellar targeting peptide of PAPSR-19 thus remains equivocal.

DNA sequencing reactions using the pYES-derived primers P2 and P1 generated 212 nucleotides at the 5' end (Figure 4.11a) and 159 nucleotides at the 3' end (Figure 4.11b) of the cDNA sequence of clone Papsr-20. The longest open reading frame encoded by the 5' end sequence of Papsr-20 is a polypeptide of 68 residues (termed PAPSR-20; Figure 4.11a). The first methionine of PAPSR-20 is not in a favourable context for translation (Joshi 1987; Lütcke et al. 1987), suggesting that the polypeptide is truncated at the N-terminus. The 3' end sequence of Papsr-20 contains a reading frame of 48 residues (termed PAPSR-20C) than is terminated by an in-frame stop-codon, but it could not be ascertained from the sequence information available whether PAPSR-20 and PAPSR-20C were the N- and C-terminal residues of the protein encoded by clone Papsr-20.

Comparisons of clones Papsr-19 and Papsr-20 and of PAPSR-19 and PAPSR-20 using the GCG Gap program did not reveal significant identities (not shown). Lack of identity between Papsr-19 and Papsr-20 and between their encoded polypeptides indicated either that the clones encoded different genes, or

(a)

```
GTTACTCTATGAAACCTTTAAACGCTGAGTCACATTCACGAAGCGAATCTTGGGTTACTCGTGCTTCAACTCTAATTGCT
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
caatgagatactttggaatttgcgactcagtgtaagtgcctcgcttagaacccaatgagcacgaagttgagattaacga
      M K P L N A E S H S R S E S W V T R A S T L I A

CCTGAAGTTGAAGAGAAAGGAGGAGAAGTTGAAGACTTTGAGCAACTTGCTAAAAAGCTTGAAGATGCTTCTCCACTTGA
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 160
ggacttcaacttctcttctcctccttcaacttctgaactcggtgaacgatttttcgaacttctacgaagaggtgaact
      P E V E E K G G E V E D F E Q L A K K L E D A S P L E

AATCATGGATAAAGCTCTTGAGAGATTCGGAGACCAAATCGTAATTGCTTT
161 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 212
ttagtacctatttcgagaactctctaagcctctggtttagcgattaacgaaa
      I M D K A L E R F G D Q I A N C F
```

(b)

```
cagaaggcgcttgctaagcaagagcatcagttagggagcaagccgaaaatacttctctttccgaaaagagctccacgagc
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 80
GTCTTCCGCGAACGATTGCTTCTCGTAGTCAATCCCTCGTTCGGCTTTTATGAAGAGAAAGGCTTTTCTCGAGGTGCTCG
      Q K A L A K Q E H Q L G S K P K I L L F P K R A P R A

tattaagtacccttcagagcatagagatggttgattcactcatgctggtttgtgaatcttcttcgggtgagttatgagataa
81 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 159
ATAATTTCATGGGAAGTCTCGTATCTTACAACFAAGTGAGTACAGCAAACACTTAGAAGAAGCCACTCAATACTCTATT
      I K Y P S E H R D V D S L M S F V N L L R *
```

**Figure 4.11** Nucleotide sequence of the 5' and 3' end of clone Papsr-20 and deduced amino acid sequence

*Arabidopsis thaliana* clone Papsr-20 was sequenced manually using the plasmid pPAPSR20 as template and pYES-derived primers P1 and P2. Single strand DNA sequence obtained from the 5' end (a) and 3' end (b) of clone Papsr-20 is shown in capitals, with the complementary sequence given in small letters. Putative forward reading frames, an open reading frame of 68 residues designated PAPSR-20 at the 5' end (a) and a polypeptide of 48 residues designated PAPSR-20C at the 3' end (b), are shown below their encoding nucleotide sequences. The stop codon in (b) is indicated by an asterisk.

that the available sequence information was derived from disparate regions of similar genes. As the clones appeared to be truncated differentially at both 5' and 3' ends, the latter possibility was not excluded.

#### 4.2.5 Comparison of Apsk and Papsr clones

A similarity between the N-terminal domains of APSK-15 and PAPSR-19 was noticed. Using the GCG Gap program, identity between the two polypeptides was shown to be 57 % (74 % homology). APSK-15 also had 52.2 % identity (68.6 % homology) with PAPSR-20, while APSK-14 had 59.2 % identity (77.8 % homology) with PAPSR-20. An alignment of the regions of overlap between these four polypeptides is shown in Figure 4.12. The region of identity between APSK-15 and PAPSR-19 spans their putative organellar targeting domains and part of the putative mature proteins. PAPSR-20 is truncated at the N-terminus compared with APSK-15 and PAPSR-19, but strong identity was observed between PAPSR-20 and the two APSK polypeptides further into the putative mature proteins. The alignment and identity of APSK and PAPSR sequences implied that these polypeptides were isoforms of the same enzyme, and therefore that the cDNA clones Apsk-14, Apsk-15, Papsr-19 and Papsr-20 were derived from members of the same gene family.

That Apsk and Papsr clones encoded similar polypeptides was not predicted. APSK-14 and APSK-15 (and thus by analogy, PAPSR-19 and PAPSR-20) have strong homologies to bacterial and yeast PAPS reductases (Section 4.2.2.1), which may explain the ability of PAPSR-19 and PAPSR-20 to complement *E. coli* strain JM96 defective in PAPS reductase. Retransformation experiments confirmed that each of the plasmids pAPSK14, pAPSK15, pPAPSR19, pPAPSR20 and pPAPSR26 could complement both *E. coli* mutant strains JM81A and JM96 (not shown). The results suggested that a single *A. thaliana* enzyme could catalyse both APS kinase and PAPS reductase

```

1
APSK-14 .....
APSK-15 MALaINVSSS SSSaIssSsF psS.DlKVtk IGSLRLLnRt nVsaaSLsLS GK.rSSVKaL
PAPSR-20 .....Mk pl.naeshsr
PAPSR-19 MAMsVNVSSS SSSgIinSrF gvSlEpKVsq IGSLRLLdRv hVapeSLnLS GKpsSSVKpL
CON MA---NVSSS SSS-I--S-F --S---KV-- IGSLRLL-R- -V---SL-L- -----

61
APSK-14 ...mIplAaT mVAEIaEEvE VVEIEDFEEL AKKLENASPL EIMDKALEKY GNDIAIAFSG
APSK-15 NvqsITKeSi vaseVtEkLD VVEVEDFEEL AKrLENASPL EIMDKALEKF GNDIAIAFSG
PAPSR-20 seswVTrAST lIApeveEE.k ggEVEDFEqL AKKLEdASPL EIMDKALErF GdqIancF..
PAPSR-19 NaepkTKdSm il.....
CON -----E--- --EVEDFE-L AK-LE-ASPL EIMDKALE-F G--IA--FSG

```

**Figure 4.12** Amino acid comparison of APSK-14, APSK-15, PAPSR-19 and PAPSR-20 polypeptides

A GCG PileUp alignment of the amino acid sequences APSK-14 (first 57 residues only), APSK-15 (first 117 residues only), PAPSR-19 and PAPSR-20, partial polypeptides deduced from the 5' end of clones Apsk-14, Apsk-15, Papsr-19 and Papsr-20, respectively, is shown. Amino acids common to the aligned sequences are given in a consensus sequence (CON) beneath the polypeptides.

activity in *E. coli*, or that the plant enzyme bypassed these two steps and allowed assimilation of sulphate in *E. coli* via a different pathway.

#### **4.2.6 Analysis and sequence comparisons of full-length APSK/PAPSR isoforms**

The sixteen Papsr clones derived from large JM96/ $\lambda$ KC/ $\lambda$ YES strains (Section 4.2.3) were characterised further in this laboratory by a postdoctoral research assistant, Dr J. Gutierrez-Marcos. Three Papsr clones, Papsr-19, Papsr-26 and Papsr-43, which encoded different isoforms of the same enzyme, were sequenced completely by Dr Gutierrez-Marcos and Mr E. Campbell. The deduced polypeptides from these Papsr clones were designated "PAPS Reductase Homologues" (PRHs) (Gutierrez-Marcos et al. 1996). The PRH-19, PRH-26 and PRH-43 deduced sequences obtained by Dr Gutierrez-Marcos and Mr Campbell were used in this study for further analysis.

#### *Characterisation of PRH-19, PRH-26 and PRH-43 isoforms and comparison to APSK-14, APSK-15, PAPSR-19 and PAPSR-20*

The PRH-19, PRH-26 and PRH-43 isoforms were 465, 458 and 453 residues in length, with calculated molecular masses of 51.7, 50.8 and 50.5 kDa, respectively. When compared using the GCG Gap program, PRH-19 showed 84.2 % identity (92.3 % homology) to PRH-26 and 77.9 % identity (87.2 % homology) to PRH-43, while PRH-26 had 78.4 % identity (89.4 % homology) to PRH-43. An alignment of these isoforms illustrated extensive similarity throughout their length, except in two short domains (residues 40 - 89 and 333 - 347 in Figure 4.13). GCG Gap comparisons revealed which APSK and PAPSR polypeptides were most similar to the PRH isoforms: APSK-14 was 99.6 % identical (100 % homologous) to PRH-19, APSK-15 was 96.6 % identical (97.1 % homologous) to PRH-26, and PAPSR-20 was 94.1 % identical (97.1 % homologous) and PAPSR-20C 89.6 %

	1						60
PRH-19	MAMsVnvSSs	ssSgiInSrf	gvSlEpKvsq	IgSlRLldRV	hvapvslnls	gkrsssvKpL	
PRH-26	MALaInvSSs	ssSaiIsSsf	psS.DlKvtk	IgSlRLlnRV	nvsaaslsls	gkr.ssvKaL	
PRH-43	MALaVtsSSt	aiSgsIfSrs	gaSsEsKalq	IcSiRLsdRV	<i>hla.....</i>	<i>..qrrpmKpL</i>	
CON	MAL-V--SS-	--S--I-S--	--S-E-K---	I-S-RL--RV	-----	-----K-L	
	61						120
PRH-19	Naepktk.DS	mIplaatmva	eiaEevevve	IEDFEeLAKk	LEnASPLEIM	DKALEkYGNd	
PRH-26	Nvqsitk.ES	iVa.....s	evtEkldvve	VEDFEeLAKr	LEnASPLEIM	DKALEkFGNd	
PRH-43	NaeshsrsES	wVtrastlia	pevEe.kggE	IEDFEqLAKk	LEdASPLEIM	DKALerFGNq	
CON	N-----ES	-V-----	---E-----E	IEDFE-LAK-	LE-ASPLEIM	DKALE-FGN-	
	121						180
PRH-19	IAIAFSGAED	VALIEYAhLT	GrPFRVFSLD	TGRLNPETRY	FFDaVEKhYG	IRIEYMFpDs	
PRH-26	IAIAFSGAED	VALIEYAhLT	GrPYRVFSLD	TGRLNPETRY	LFDtVEKhYG	IRIEYMFpDa	
PRH-43	IAIAFSGAED	VALIEYArLT	GkPFRVFSLD	TGRLNPETRY	LFDaVEKqYG	IRIEYMFpDa	
CON	IAIAFSGAED	VALIEYA-LT	G-PFRVFSLD	TGRLNPETRY	LFD-VEK-YG	IRIEYMFpD-	
	181						240
PRH-19	VEVQgLVrSk	GLFSFYEDGH	QECCRVrKVR	PLRRALKGLk	AWITGQRKDQ	SPGTRSEIPV	
PRH-26	VEVQaLVRnK	GLFSFYEDGH	QECCRIrKVR	PLRRALKGLr	AWITGQRKDQ	SPGTRSEIPV	
PRH-43	VEVQaLVRnK	GLFSFYEDGH	<u>QECCRVrKVR</u>	PLRRALKGLk	<u>AWITGQRKDQ</u>	SPGTRSEIPV	
CON	VEVQ-LVR-K	GLFSFYEDGH	QECCRVrKVR	PLRRALKGL-	AWITGQRKDQ	SPGTRSEIPV	
	241						300
PRH-19	VQVDPVFEGL	DGGVGSLVKW	NPvANVEGnD	VWNFLRTMDV	PVntLHAaGY	ISIGCEPCTk	
PRH-26	VQVDPVFEGL	DGGVGSLVKW	NPvANVEGnD	VWNFLRTMDV	PVntLHAaGY	VSIGCEPCTr	
PRH-43	VQVDPVFEGL	DGGVGSLVKW	NpLANVEGaD	VWNFLRTMDV	PVNaLHAqGY	VSIGCEPCTr	
CON	VQVDPVFEGL	DGGVGSLVKW	NP-ANVEG-D	VWNFLRTMDV	PVN-LHA-GY	VSIGCEPCT-	
	301						360
PRH-19	aVLPgQHERE	GRWWWEDAKA	KEC <u>GLHKG</u> NV	KE.n.sddakv	ngesksaVaD	IFkSeNlVtL	
PRH-26	aVLPgQHERE	GRWWWEDAKA	KEC <u>GLHKG</u> NI	KEntngnata	nvngtasVaD	IFnSeNvVnL	
PRH-43	pVLPgQHERE	GRWWWEDAKA	KEC <u>GLHKG</u> NI	KE...edga	adskpaaVqE	IFeSnNvVaL	
CON	-VLPgQHERE	GRWWWEDAKA	KEC <u>GLHKG</u> NI	KE-----	-----V-D	IF-S-N-V-L	
	361						420
PRH-19	SrqGIENLMK	LEfrKEpWiV	VLYa <u>PWC</u> PFC	QAMEASYdEL	AaKLaGsGdK	VAKFRADGDQ	
PRH-26	SrqGIENLMK	LEnrKEaWiV	VLYr <u>PWC</u> PFC	QAMEASfDEL	AdKLgGsGvK	VAKFRADGDQ	
PRH-43	SkgGVENLLK	LEnrKEaWlV	VLYa <u>PWC</u> PFC	QAMEASYiEL	AeKLaGkGvK	VAKFRADGEO	
CON	S--GIENLMK	LE-RKE-W-V	VLY- <u>PWC</u> PFC	QAMEASY-EL	A-KL-G-G-K	VAKFRADGDQ	
	421				467		
PRH-19	KEFAKqELQL	GSFPTILvFP	KnssrpIKYP	SEkRDVESLl	SFlnLvR		
PRH-26	KDFAKkELQL	<u>GSFPTILvFP</u>	KnssipIKYP	SEkRDVDSLt	SFlnLvR		
PRH-43	KEFAKqELQL	<u>GSFPTILlFP</u>	KrapraIKYP	SEhRDVDSLm	SFvNlLr		
CON	KEFAK-ELQL	GSFPTIL-FP	K-----IKYP	SE-RDVDSL-	SF-NL-R		

**Figure 4.13** Amino acid alignment of PRH-19, PRH-26 and PRH-43 proteins

The *Arabidopsis thaliana* PAPS reductase homologue (PRH) proteins PRH-19, PRH-26 and PRH-43 (full-length sequence obtained from Gutierrez-Marcos et al. 1996) deduced from clones Papsr-19, Papsr-26 and Papsr-43, respectively, were aligned using the GCG PileUp program. A consensus sequence (CON) shows amino acids conserved between the aligned sequences. Putative N-terminal transit peptide cleavage sites are given in italics. A conserved cysteine residue essential for PAPS reductase catalytic activity in *Escherichia coli* (Berendt et al. 1995) is single underlined, while a WCPFC motif with homology to the active site of thioredoxin and identical to the active site of glutaredoxin (Holmgren 1989) is double underlined.

identical (91.7 % homologous) to PRH-43. The partial polypeptide PAPSR-19 showed 94.1 % identity (97.1 % homology) to the full length sequence of PRH-19. Residues which differ between partial polypeptides and full-length proteins was attributed to DNA sequencing errors, particularly within the Papsr clones which did not produce clear sequence information, and the identity between the partial and corresponding full-length polypeptides is still greater than that between the PRH isoforms. It was concluded that three isoforms were present amongst the PRH, APSK and PAPSR deduced polypeptides analysed: (i) PRH-19 (= PAPSR-19) and APSK-14; (ii) PRH-26 and APSK-15; and (iii) PRH-43 and PAPSR-20.

The three PRH isoforms have similar N-terminal domains which show characteristics of organellar targeting peptides (Figure 4.13). With PAPSR-20 truncated by 47 residues compared with its full-length homologue PRH-43, and APSK-14 shorter by 69 residues than its homologue PRH-19, it is clear that the N-terminal extensions of PRH-43 and PRH-19 are not required for the enzyme activity that restores cysteine prototrophy to *E. coli* strains JM81A and JM96. The putative SPP semi-conserved chloroplast cleavage site motif [VI]-X-[CA]↓A (Gavel & von Heijne 1990) found in the N-terminal domain of PRH-26 (residue 42 - 45 in Figure 4.13) was described for APSK-15 (Section 4.2.2.1), while the putative transit peptide of PRH-19 was analysed for PAPSR-19 (Section 4.2.4.1). The PRH-19 sequence VHV↓A (residue 40 - 43 in Figure 4.13) has been suggested as a putative cleavage site based on partial identity with the SPP cleavage site motif (Gutierrez-Marcos et al. 1996). However, site-directed mutagenesis experiments have shown that a valine residue at position -1 relative to the cleavage site does not allow cleavage (von Heijne 1992), so unequivocal location of a cleavage site in PRH-19 requires further investigation. The motif VHL↓A in PRH-43 (residue 40 - 43 in Figure 4.13) has some identity to the SPP cleavage site, and is in an analogous position to the suggested PRH-19 cleavage site. The PRH N-terminal extensions have

characteristics which indicate that each isoform may be targeted to the chloroplast (Gutierrez-Marcos et al. 1996).

#### *Sequence comparisons using PRH-19*

PRH-19 was used as a GCG BLASTP program query sequence to search for matching proteins. Two recently published *A. thaliana* sequences, APR3 and APR2 (Setya et al. 1996, GenBank entries U56922 and U56291, respectively), showed most significant identity to PRH-19 (high scores 1289 and 1220, with the probability P that these high scores arose by chance in a similar size search being  $1.3 \times 10^{-237}$  and  $1.3 \times 10^{-228}$ , respectively). APR2 has 98.8 % identity with the PRH-43 isoform but is truncated by 47 residues at the N-terminus compared with PRH-43. Differences in the APR2 and PRH-43 sequences may be ascribed to cDNA sequencing errors or the two proteins may be encoded by disparate genes. APR3 shows greatest identity (92.3 %) with PRH-26; most of the divergence of sequence is at their C-termini following residue 408 of PRH-26 and is caused possibly by a frame-shift reading error. The three PRH isoforms have identity at their C-termini (Figure 4.13), which implies that the APR3 C-terminal sequence is incorrect. APR3 is truncated by 27 residues at the N-terminus compared with PRH-26.

Since residues 69 - 299 of PRH-19 are equivalent to residues 1 - 230 of the truncated polypeptide APSK-14, the PAPS reductase enzymes from *Thiocapsa roseopersicina* (Haverkamp, T., Gisselmann, G., Schwenn, J.D. 1993, unpublished SWISS-PROT entry P52672), *Synechococcus* sp. (Niehaus et al. 1992), *Salmonella typhimurium* (Ostrowski et al. 1989b), *E. coli* (Krone et al. 1991) and *Saccharomyces cerevisiae* (Thomas et al. 1990) that showed significant homology to APSK-14 (Section 4.2.2.1; Figure 4.6) were detected also by PRH-19. The BLASTP high scores between PRH-19 and these PAPS reductases were between 61 and 96, with P values ranging from  $3.6 \times 10^{-12}$  to  $2.1 \times 10^{-4}$ . A recently submitted PAPS reductase

sequence from the fission yeast *Schizosaccharomyces pombe* (Connor, R., Churcher, C.M., Barrell, B.G., Rajandream, M.A., Walsh, S.V. 1996, unpublished SWISS-PROT entry Q10270) also showed homology to PRH-19 (high score 114, with an associated P value of  $3.6 \times 10^{-11}$ ). The single cysteine residue essential for catalytic activity in *E. coli* PAPS reductase and conserved in all identified PAPS reductases (Berendt et al. 1995) was found to be present also in the PRH sequences (residue 323 in Figure 4.13).

The complete sequence of PRH-19 detected additional significant matches which were not found with the truncated APSK-14 polypeptide. The C-terminal region of PRH-19 showed greatest homology to a protein disulphide isomerase (PDI; EC 5.3.4.1) from the blood fluke *Schistosoma mansoni* (Finken et al. 1994). The primary role of PDIs is the formation and rearrangement of disulphide bonds required for folding and stability of secretory and cell surface proteins (Bardwell 1994; Freedman et al. 1994). Several PDI-like sequences which define a family of PDI proteins differing in sequence homology and specific cellular function (Freedman et al. 1994) were detected by PRH-19. Highest scoring members of each of the six PDI family groups and corresponding BLASTP values were: (i) PDI from *S. mansoni* (Finken et al. 1994; high score 90, P value  $6.6 \times 10^{-8}$ ); (ii) ERp72 from *Mus musculus* (Mazzarella et al. 1990; high score 98, P value  $1.5 \times 10^{-4}$ ); (iii) ERp60 from *S. mansoni* (Finken, M., Kunz, W. 1993, unpublished SWISS-PROT entry P38658; high score 87, P value  $1.5 \times 10^{-3}$ ); (iv) PDI from *Saccharomyces cerevisiae* (Farquhar et al. 1991; high score 77, P value  $1.9 \times 10^{-3}$ ); (v) P5 from *Medicago sativa* (Shorrosh & Dixon 1992; high score 71, P value  $2.9 \times 10^{-2}$ ); and (vi) BS2 from *Trypanosoma brucei* (Hsu et al. 1989; high score 62, P value 0.72). The C-terminus of PRH-19 detected several thioredoxin proteins, with the best match to thioredoxin from *Haemophilus influenzae* (Fleischmann et al. 1995) having a high score of 62 with an associated P value of 0.45. Thioredoxins are small

ubiquitous enzymes that catalyse reduction of disulphide bonds to dithiol groups (Holmgren 1985, 1989). An alignment of the C-terminal region of PRH-19 with the corresponding domains from optimal matching examples from each group of the PDI family and *H. influenzae* thioredoxin is shown in Figure 4.14.

PDI consists of the structural domains a, e, b, b', a' and c, where a and a' are duplicated domains which are functionally and structurally homologous to thioredoxin, e has similarity with oestrogen receptors, b and b' are duplicated domains of unknown function, and c is a putative Ca<sup>2+</sup>-binding region (Holmgren 1989; Freedman et al. 1994; Lyles & Gilbert 1994; Darby & Creighton 1995). Significant homology between PRH-19 and members of the PDI family was confined to the thioredoxin-like domains a and a'. Mutagenesis studies have revealed that the active site of PDI lies within these thioredoxin-like domains and consists of two redox-active cysteines that are separated by glycine and histidine residues and preceded by tryptophan (WCGHC) (Vuori et al. 1992). The active site of PDI thus corresponds with the active site of thioredoxin, WCXZC, where X and Z are usually glycine and proline, respectively (Holmgren 1985, 1989; Schürmann 1993). In PRH-19, the residues WCPFC align with the active sites of PDI-like proteins and thioredoxin (Figure 4.14). The WCPFC motif in PRH-19 is identical with the active site of several glutaredoxin proteins, which are similar in structure and function to thioredoxin (Holmgren 1989; Minakuchi et al. 1994), whereas the domains flanking the PRH-19 WCPFC site show greatest identity with *Schistosoma mansoni* PDI and related members of the PDI family (Figure 4.14).

Can the structural similarity between the PRHs, PDI-like proteins and thioredoxin/glutaredoxin tell us anything about the possible activity of the C-terminal domain of the PRHs? The two redox-active cysteine residues in the active sites of PDI and thioredoxin/glutaredoxin contain dithiol groups that

```

PRH-19          369 lefrkepiwV vLYAPWCpfc qAMEasYDEL aaklagS.g dkvAKfrADg
S. mansoni PDI 375 vkdkskdvFV kLYAPWCqhc kALapvWDEL getfknS..d tviAKmDatv
Consensus 1    -----V -LYAPWC--C -A-----DEL -----S--- ---AK--A--
ERp72          183 vvnadiiLV eFYAPWCqhc kkLapeYeka akelskrspplakvDate
ERp60          30  elksipvaLV kFYAPWCqhc kkLapeFtsa aqiisgktnvklvkvDctt
S. cerevisiae PDI 390 vndpkkdvLV LYYAPWCqhc krLaptYqel adtyanatsd vliaklDhte
P5             161 vldgtkdvLV eFYAPWCqhc ksLapiYekv aavfkse.dd vvianlDadk
BS2            362 hltsgkdmLI lFFAPWCqhc knFaptFdki akef..datd livaelData
TRX            38  sqykgkpvYV kmWAsWCpiC laglaeidl saekdrnfev itivspDhkg
Consensus 2    -----LV -FYAPWC--C --L---Y--- -----D---

PRH-19          dqkEfak.qe LqlgSFPTil vFPKNSsr.p IkYpsekRdv EsLtsF
S. mansoni PDI  nevE....d LkvtSFPTlk fYPKNSee.v IdYtgd.Rsf EaLkkF
Consensus 1    ---E----- L---SFPT-- --PKNS---- I-Y----R-- E-L--F
ERp72          qtdla...kr FdvsgYPTlk iFrkgr...p fdyngp.Rek ygivdY
ERp60          qesic...se FgvsgYPTlk iFrngd..ld geyngp.Rna ngianY
S. cerevisiae PDI  ndvrg.... vviEGYPTiv lYpggkkses vvyqgs.Rsl dslfdF
P5             yrdla...ek YdvsgFPTlk fFpkgnk.ag edyggg.Rdl ddfvaF
BS2            nyvns...st FtvtaFPTvf fvpnggk..p vvfege.Rsf envyeF
TRX            ekdtadfiew YkgleYknit vLldek...g eiidka.Rvr gypfnL
Consensus 2    -----R-- F-----PT-- -F-----R-- -----F

```

**Figure 4.14** Amino acid alignment of the C-terminus of PRH-19 with thioredoxin and thioredoxin-like domains of representative members of the PDI family

The C-terminus of *Arabidopsis thaliana* PRH-19 (Gutierrez-Marcos et al. 1996) was aligned with highest BLASTP scoring sequences from members of each of the six PDI family groups, viz. *Schistosoma mansoni* PDI (Finken et al. 1994; GenBank translation PID:g312018 from Z22933), *Mus musculus* ERp72 (Mazzarella et al. 1990; SWISS-PROT entry P08003), *S. mansoni* ERp60 (Finken, M., Kunz, W. 1993; unpublished SWISS-PROT entry P38658), *Saccharomyces cerevisiae* PDI (Farquhar et al. 1991; SWISS-PROT entry P17967), *Medicago sativa* P5 (Shorrosh & Dixon 1992; SWISS-PROT entry P38661) and *Trypanosoma brucei* BS2 (Hsu et al. 1989; SWISS-PROT entry P12865), and from *Haemophilus influenzae* thioredoxin (TRX - Fleischmann et al. 1995; GenBank translation PID:g1221596 from U32769) using the GCG PileUp and LineUp programs. Consensus 1 reveals residues conserved between PRH-19 and *Schistosoma mansoni* PDI, while Consensus 2 shows residues common to at least seven of the eight aligned sequences. The conserved active site WCXXC is single underlined.

can react with each other to form a disulphide bond (Holmgren 1989; Bardwell 1994; Freedman et al. 1994). The redox potential of dithiol/disulphide interchange at the WCG[PF]C active site in thioredoxin and glutaredoxin is highly reducing, comparable to that of reduced DTT, and the enzymes therefore preferentially catalyse the reduction of disulphide bonds to disulphide groups in substrate proteins (Holmgren 1985; Krause et al. 1991). Conversely, the thioredoxin-like domains of PDI catalyse disulphide bond formation in substrate proteins, with dithiol/disulphide interchange at the active site WCGHC having a more oxidising redox potential (Hawkins et al. 1991; Freedman et al. 1994; Darby & Creighton 1995). Mutagenesis studies with *E. coli* thioredoxin indicate that the residues X and Z in the active site motif WCXZC are critical in determining the redox potential of the active site disulphide, and hence whether the enzyme catalyses disulphide bond formation or breakage (Krause et al. 1991; Lundström et al. 1992; Freedman et al. 1994). Since the X and Z residues in the corresponding region of the PRHs are identical with the active site of glutaredoxin, it can be reasoned that the conserved putative active site in the C-terminal domain of the PRHs is highly reducing rather than oxidising in thiol/disulphide interchange reactions. The functional role of domains which flank the thioredoxin/glutaredoxin/PDI active site, which in the case of PRHs show greater identity to PDI than thioredoxin or glutaredoxin, is not understood clearly (Freedman et al. 1994). High levels of similarity between the PRHs and PDI in these domains flanking the WCXZC site may indicate that the C-terminal domain of PRHs was derived from the thioredoxin-like domain of a PDI-like protein, perhaps by DNA splicing onto a modified PAPS reductase gene capable of reducing the sulpho group of APS. It is possible that these flanking domains perform an analogous role in PRHs and PDIs not required by thioredoxin or glutaredoxin.

The proposed reaction mechanism for *E. coli* (Berendt et al.

1995) and *Saccharomyces cerevisiae* (Schwenn et al. 1988) PAPS reductase activity involves transfer of two electrons from the dithiol group of reduced thioredoxin to the enzyme, followed by enzymatic reduction of PAPS to PAP and sulphite. Since the N-terminal region of PRHs is homologous to PAPS reductase and structural evidence implies that the C-terminal region will exhibit thioredoxin activity, it can be hypothesised that the PRHs achieve reduction of PAPS (and/or APS; Section 4.2.2.1) by interaction of reduced thioredoxin-like C-terminal domain with the active site for PAPS (or APS) reduction in the N-terminal region on the same protein, followed by reduction of the sulpho group of PAPS (or APS) to sulphite. In this scheme, initial reduction of the thioredoxin-like domain of PRHs might be accomplished by reduced ferredoxin or NADPH, as for thioredoxin in plants (Schürmann 1993). A known thioredoxin system provides a molecular mechanism for this hypothesis. In *Mycobacterium leprae*, thioredoxin and thioredoxin reductase are fused on a single protein, enabling oxidised thioredoxin to swing into the active site of thioredoxin reductase and become reduced by NADPH (Wieles et al. 1995). The 20 residue polypeptide linker between thioredoxin reductase and thioredoxin on the mycobacterial molecule is thought to allow conformational bending (Wieles et al. 1995) and is paralleled by the non-conserved domain which is 11 to 15 residues in length that links the PAPS reductase-like and thioredoxin/PDI-like domains of the PRH isoforms (residue 333 - 347 in Figure 4.13).

Two proteins containing PDI- or thioredoxin-like motifs illustrate another possible catalytic mechanism for the PRHs. Firstly, the *E. coli* DipZ protein has a C-terminal domain with a PDI-like motif that is thought to maintain the appropriate pairs of cysteine residues of cytochrome *c* apoproteins in a reduced state and in correct confirmation to allow attachment of haem groups (Crooke & Cole 1995). Secondly, the C-terminus of the self-incompatibility protein

S<sub>1</sub> from *Phalaris coerulescens* contains a thioredoxin-like motif which is thought to modify dithiol/disulphide bonds in newly synthesised proteins that effect self-incompatibility (Li et al. 1995). The mechanism that DipZ and S<sub>1</sub> have in common is that catalytic activity requires PDI/thioredoxin-like domains of the proteins to react with dithiol/disulphide groups of substrate proteins. It is feasible that the C-terminus of the PRHs may have a similar role, ie. to achieve modification of a carrier protein molecule which could then accept the sulpho group of APS or PAPS for further reduction. In this way, the PRHs would catalyse an APS sulphotransferase reaction that has been advanced for the bound intermediate pathway of higher plant sulphate assimilation (Schmidt 1972, 1973; Urlaub & Jankowski 1982; Li & Schiff 1992; Schiff et al. 1993; Chapter 1.2.3.1). However, the existence of a carrier molecule which is required for this proposed APS sulphotransferase mechanism of PRH activity has not been demonstrated in higher plants (Schmidt & Jäger 1992).

Clearly, further studies are required to determine the catalytic mechanism and *in vivo* role of the novel PRH proteins that contain domains structurally similar to PAPS reductase and thioredoxin, and which are capable of functional complementation of APS kinase- and PAPSR reductase-deficient *E. coli* mutant strains JM81A and JM96, respectively.

#### **4.2.7 Enzymatic analysis of JM96, wild-type and pPAPSR-complemented *Escherichia coli* strains, *Arabidopsis thaliana* and *Spinacia oleracea***

##### *PAPS reductase activity*

PAPS reductase assays were performed on *E. coli* strain JM96, which is defective in PAPS reductase activity (Jones-Mortimer 1968, 1973), strain TB1, which is wild-type with respect to

sulphate assimilation (Johnston et al. 1986), and complemented strains derived from JM96 to ascertain whether the ability of strains PAPSR19, PAPSR26 and PAPSR43 to grow well on minimal medium containing sulphate as sole sulphur source (Section 4.2.3) was associated with restoration of PAPS reductase activity. Enzyme activity was assayed by measuring acid-volatile activity produced with <sup>35</sup>S-PAPS as substrate and in the presence of reduced ferredoxin, DTT, carrier sulphite and with or without purified recombinant *E. coli* thioredoxin (Schriek & Schwenn 1986; Schwenn & Shriek 1987; Schwenn et al. 1988; Schwenn 1989; Krone et al. 1991). As discussed in Chapter 1.2.3.3, this *in vitro* PAPS reductase assay is useful for determining whether a sample contains an enzyme capable of reducing the sulpho group of PAPS to compound(s) which can form or exchange with acid-volatile product(s), but enzyme reaction product(s) cannot be identified because of the possibility of cross-reactions with DTT and/or carrier sulphite. Conventional use of the term "PAPS reductase activity" will be adopted here, although *sensu strictu* the enzyme assay does not measure the ability to reduce PAPS to sulphite.

Preliminary investigations were carried out to optimise conditions for disruption of bacterial extracts by probe sonication (data not shown). Sonication is an effective method for disintegration of *E. coli* cells, but requires empirical determination of parameters such as viscosity (dependent on resuspension volume of washed pellet), oscillation amplitude and frequency, and pulse duration (Hughes et al. 1971). Cell-free bacterial extracts obtained after centrifugation of sonicated cells were assayed for PAPS reductase activity (Chapter 2.17.3). An incubation period of 5 min was used routinely in PAPS reductase assays as initial studies with wild-type extracts indicated that after longer time periods the reaction velocity decreased (data not shown). No activity could be detected in extracts from the mutant strain JM96, or in strain JM96 transformed with the

empty plasmid pYES (Table 4.2). This complete deficiency of PAPS reductase activity in strain JM96 conflicts with the original characterisation of the mutant strain where 3 % activity compared with wild-type was reported (Jones-Mortimer 1968). This discrepancy may be accounted for by differences in the assay systems employed: Jones-Mortimer (1968) utilised an assay in which reducing power was provided by NADPH regenerated by glucose-6-phosphate and glucose-6-phosphate dehydrogenase, and where reaction products from <sup>35</sup>S-PAPS substrate were quantified following paper chromatography of an *N*-ethylmaleimide derivative of <sup>35</sup>S-sulphite (Pasternak et al. 1965). In contrast with the mutant strain JM96, extracts from wild-type strain TB1 demonstrated PAPS reductase activity that was thioredoxin-dependent (Table 4.2); no activity was detected in boiled enzyme extract. PAPS reductase activity values for strain TB1 in the presence of exogenous thioredoxin were similar to those reported for wild-type *E. coli* using an analogous assay (Krone et al. 1990b), although strain TB1 had lower activities without exogenous thioredoxin. Endogenous thioredoxins present in crude extracts may account for some variation in assay results (Schwenn et al. 1988). Activity data obtained from this study and that of Krone et al. (1990b) imply an error in units of activity in a report giving PAPS reductase activity for wild-type *E. coli* as 48 mmol/mg/min (Niehaus et al. 1992), approximately 10<sup>6</sup>-fold greater than that obtained for strain TB1.

The complemented strains PAPSR19, PAPSR26 and PAPSR43 exhibited low PAPS reductase activity that in contrast to the wild-type *E. coli* activity was thioredoxin-independent (Table 4.2). Extracts from strain JM96 cells retransformed with plasmids pPAPSR19, pPAPSR26 or pPAPSR43 had up to ten-fold higher levels of PAPS reductase activity than the original complemented strains, and activity again was not thioredoxin-dependent. Lower PAPS reductase activity in extracts from the original complemented strains may have been caused by

**Table 4.2** PAPS reductase activity in cell-free extracts of *Escherichia coli* JM96 (*cysH*), TB1 (wild-type) and pPAPSR-complemented JM96 strains, and in leaf extracts of *Arabidopsis thaliana* and *Spinacia oleracea*

<u>Extract Source</u>	<u>PAPS reductase activity (pmol/mg/min)</u>	
	<u>+ thioredoxin</u>	<u>-thioredoxin</u>

*E. coli* strain:

JM96 ( <i>cysH</i> )	0.000	0.070
JM96/pYES	0.000	0.000
TB1 (wild-type)	72.11	0.398
PAPSR19	1.278	1.348
JM96/pPAPSR19	12.92	14.08
PAPSR26	0.966	0.663
JM96/pPAPSR26	5.430	4.736
PAPSR43	1.648	1.396
JM96/pPAPSR43	6.426	6.227

Plant leaf:

<i>A. thaliana</i>	0.000	0.000
<i>S. oleracea</i>	0.104	0.064

PAPS reductase activity was assayed in the presence or absence of 4.5  $\mu$ g purified *E. coli* thioredoxin in cell-free extracts of the *E. coli* *cysH* mutant strain JM96, strain JM96 transformed with the empty vector pYES, wild-type strain TB1, complemented strains PAPSR19, PAPSR26 and PAPSR43, and strain JM96 retransformed with pPAPSR19, pPAPSR26 or pPAPSR43, and in crude extracts from leaves of *A. thaliana* and *S. oleracea*. An incubation period of 5 min was used. Data are representative of at least two separate experiments, with 1 or 2 replicates per sample.

diminished expression levels of active enzyme through extended cleavage of exogenous plasmid DNA by the wild-type restriction system of strain JM96. Unlike SAT-complemented strains SAT1 and SAT2 which had restored serine acetyltransferase activity that was about 20-fold higher than wild-type levels (Table 3.4), PAPS reductase activity in extracts from the original complemented JM96 strains was only 1.3 - 2.3 % of wild-type activity, and 7.5 - 17.9 % of wild-type activity in extracts from strain JM96 retransformed with pPAPSR plasmids (Table 4.2). Due to limited availability of <sup>35</sup>S-PAPS, it was not established whether the concentration of this substrate was saturating for kinetic activity in the complemented or retransformed JM96 strains. In addition, partially purified PAPS reductase from *S. cerevisiae* exhibited only 25 % initial velocity and 6 - 10 % maximum velocity in the presence of *E. coli* thioredoxin compared with homologous thioredoxin (Schwenn et al. 1988), so it is feasible that the plant PRHs cannot use heterologous *E. coli* thioredoxin efficiently for catalysis. Structural dissimilarity between bacterial PAPS reductases and the PRH proteins suggest differences in kinetic mechanism of the enzymes (Section 4.2.6), so the possibility that the assay for PAPS reductase activity employed in this study may not provide necessary but uncharacterised requirements for optimal reduction of PAPS by the PRHs cannot be excluded.

The data in Table 4.2 suggest that PAPS reductase activity conferred by the plasmids pPAPSR19, pPAPSR26 and pPAPSR43 to strain JM96 is thioredoxin-independent. This result could be explained by the presence of a thioredoxin-like C-terminal domain in the PRHs encoded by these plasmids that may, unlike *E. coli* and *S. cerevisiae* PAPS reductase catalytic activity, obviate the need for thioredoxin (Section 4.2.6). Schwenn (1989) has reported thioredoxin-dependent PAPS reductase activity in a partially purified extract from *Spinacia oleracea*. In the present study, very low levels of PAPS reductase activity were detected in total leaf extracts from

*S. oleracea*, and no PAPS reductase activity could be detected in *A. thaliana* leaf extracts (even after prolonged incubation periods of one hour) (Table 4.2). No PAPS reductase activity was detected in extracts from *S. oleracea* in the absence of DTT or with boiled extract. Activity in *S. oleracea* leaf extracts in the presence of thioredoxin was slightly higher than without thioredoxin, but the values obtained were close to the lower detection limit of the assay and are not conclusive. It remains unclear whether higher plants have PRH-like enzymes capable of reducing PAPS in the absence of thioredoxin in addition to thioredoxin-dependent PAPS reductases reported by Schwenn (1989).

Evidence presented here for three complemented PAPS<sub>R</sub> strains and strain JM96 retransformed with pPAPS<sub>R</sub> plasmids extracted from these strains suggested that restoration of cysteine prototrophy in the strains was associated with the presence of PAPS reductase activity. PAPS reductase activity in these complemented and retransformed strains was low, however, and activity differed from previously described PAPS reductases in that it was thioredoxin-independent.

#### *APS reductase activity*

Demonstration of PAPS reductase activity in JM96-derived *E. coli* strains expressing the PRH-19, PRH-26 and PRH-43 proteins did not provide an explanation of why expression of APSK-14 (= PRH-19) and APSK-15 (= PRH-26) was associated also with restoration of cysteine prototrophy in *E. coli* strain JM81A which is deficient in APS kinase (Section 4.2.1). The presence of wild-type levels of thioredoxin-dependent PAPS reductase activity in strain JM81A (data not shown) implied that the PRHs were capable of catalysing a reaction other than reduction of the sulpho group of PAPS which would allow growth of complemented JM81A strains on minimal medium containing sulphate as sole sulphur source. Extracts from strain JM96 retransformed with pPAPS<sub>R</sub>19, pPAPS<sub>R</sub>26 or pPAPS<sub>R</sub>43

were tested in this laboratory by Dr J. Gutierrez-Marcos for ability to reduce the sulpho group of  $^{35}\text{S}$ -APS to acid-volatile radioactivity using an "APS reductase" assay similar to the PAPS reductase assay. Extracts from JM96/pPAPSR19, JM96/pPAPSR26 and JM96/pPAPSR43 strains reduced the sulpho group of  $^{35}\text{S}$ -APS to acid-volatile radioactivity at rates that were 175- to 666-fold greater than reduction of the sulpho group of  $^{35}\text{S}$ -PAPS, whereas extracts from wild-type, JM96 and JM96/pYES strains did not show significant APS reductase activity (Gutierrez-Marcos et al. 1996). The APS reductase activity in JM96/pPAPSR19, JM96/pPAPSR26 and JM96/pPAPSR43 extracts was thioredoxin-independent (Gutierrez-Marcos et al. 1996).

The enzyme assays performed on extracts from JM96 strains expressing the pPAPSR19, pPAPSR26 and pPAPSR43 plasmids suggested that APS was preferred to PAPS as substrate for the PRHs. A recent report describing the cloning of *A. thaliana* cDNAs which encode PRH-like proteins APR1, -2 and -3 has provided similar evidence that PRH activity appears to be specific for APS rather than PAPS (Setya et al. 1996). As described in Section 4.2.2.1 for APSK-14 and APSK-15, the modified PP motif present at the N-terminus of PAPS reductases and conserved in the PRHs could have been adapted in the PRHs to accept APS as substrate rather than PAPS. This putative APS binding site would then conceivably accept PAPS only with much reduced affinity, as demonstrated by the ratio of APS reductase to PAPS reductase activity with these substrates. APS reductase activity exhibited by strains expressing the PRHs would allow functional rescue of strain JM81A defective in APS kinase, but it is not clear whether APS reductase activity or the low PAPS reductase activity of the PRHs allowed functional rescue of strain JM96 defective in PAPS reductase.

The APS reductase assay employed by Dr Gutierrez-Marcos in this laboratory and by Setya et al. (1996) does not allow

description of the *in vivo* kinetic mechanism of the PRHs, but shows that these enzymes are capable of reducing the sulpho group of  $^{35}\text{S}$ -APS to acid-volatile radioactivity in the presence of a reductant (DTT) and carrier sulphite. This assay has been used also to measure APS sulphotransferase activity in plant tissue (Schiff & Levinthal 1968; Brunold 1990; Kanno et al. 1996), illustrating that interpretation of assay results can be ambiguous. Whether the PRHs act as reductases by reducing APS to free sulphite and AMP, as sulphotransferases by transferring the sulpho group of APS to a carrier molecule or in yet a different manner cannot be determined from either structural analysis (Section 4.2.6) or available enzyme assay data. Considering that expression of *A. thaliana* APS kinase in *E. coli* confers both APS kinase and APS sulphotransferase/reductase activity (Arz et al. 1994), it remains a formal possibility that the PRHs may be APS kinases, but that low PAPS reductase activity levels (vestigial from a hypothetical PAPS reductase progenitor enzyme) allowed for functional rescue of strain JM96. Clearly, it is critical that the *in vivo* kinetic mechanism of the PRHs be determined before the role of these enzymes in plant sulphate assimilation can be understood.

#### **4.2.8 Expression and purification of a GST-PRH-26 fusion protein and preparation of rabbit anti-GST-PRH-26 polyclonal antibodies**

In order to understand further the catalytic role of the PRHs during plant sulphate assimilation, it was necessary to obtain purified PRH protein for use in enzyme activity assays and for the generation of immunogens for expression studies. The pGEX system (Smith & Johnson 1988) for expression in *E. coli* and purification of PRH-26 as a fusion protein with glutathione S-transferase (GST) from *Schistosoma japonicum* was selected. GST fusion proteins have been widely used as they can allow high levels of IPTG-inducible expression in *E.*

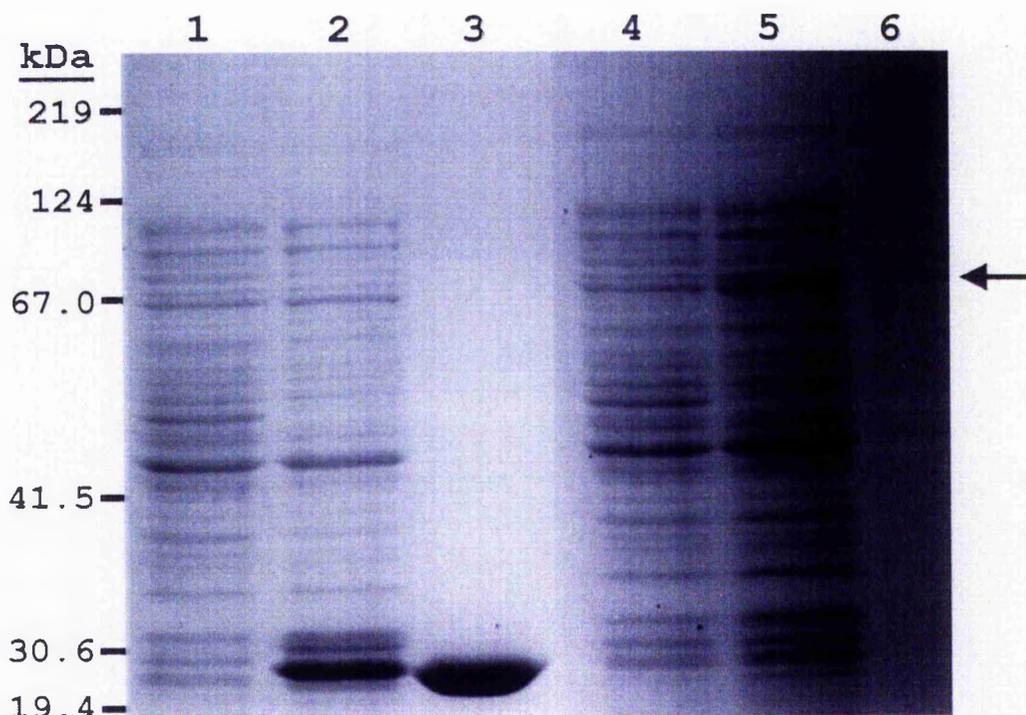
*coli* and provide simple and rapid purification in nondenaturing conditions by affinity chromatography of the GST moiety with glutathione-agarose beads (Smith & Johnson 1988; Smith 1993; Smith & Corcoran 1994). Agarose-bound fusion proteins can be eluted with reduced glutathione, or the study protein cleaved from agarose-immobilised GST by the site-specific proteases thrombin (eg. pGEX-4T vectors) or factor Xa (eg. pGEX-5X vectors) (Smith 1993). Purified uncleaved fusion proteins are utilised frequently for further applications as the GST moiety is only weakly immunogenic, and in some cases does not interfere with biochemical properties of the study protein (Smith 1993).

#### *Formation of a GST-PRH-26 fusion protein and optimisation of expression*

The pGEX-4T vectors are designed with unique restriction endonuclease recognition sites each available in three forward reading frames at the C-terminus end of the GST gene and allow thrombin cleavage at a LVPR↓GS motif situated between the GST moiety and the cloning site (Smith 1993). As none of the required restriction endonuclease cleavage sites were available at the 5' end of clone Papsr-26 for subcloning into pGEX-4T vectors, a blunt 5' end was created by filling a *Bgl*III-digested restriction site approximately 120 bp from the 5' end of clone Papsr-26 using the Klenow fragment of *E. coli* DNA polymerase, as described in Chapter 2.13.1, while the 3' end of clone Papsr-26 was excised from pPAPSR-26 by digestion with the restriction endonuclease *Xho*I. The 1.7 kb Papsr-26/*Bgl*III/Klenow/*Xho*I fragment could then be ligated directionally and in the correct reading frame into a pGEX-4T-2 vector digested with the restriction endonucleases *Sma*I (which creates blunt ends) and *Xho*I. The predicted 75 kDa fusion protein, GST-PRH-26, contains a 435 residue PRH moiety that lacks 23 residues from the N-terminal end of the complete PRH-26 protein. While some workers have deemed it necessary to remove all or part of plant transit peptides to

obtain recombinant proteins that will simulate the processed mature protein (Arz et al. 1994), others have shown that complete purified recombinant plant proteins maintain enzymatic activity (Masuta et al. 1992). Since *E. coli* can process some plant transit peptides, resulting in excretion of mature protein into the periplasm (Overbeeke & Verrips 1993), it may be advisable to remove all or part of the transit peptide if soluble recombinant proteins are required. The ability of the truncated clones Apsk-14 (Section 4.2.2.1) and Papsr-20 (Section 4.2.4.1) to complement the phenotype of *E. coli* strains JM81A and JM96, respectively, confirms that the N-terminal putative transit peptides of the PRHs are not required for catalytic activity.

Ligated pGST-PRH-26 fusion protein plasmids and a control pGEX-4T-2 vector were transformed into *E. coli* strain TB1, and transformants selected by ampicillin resistance conferred by the pGEX-4T-2 vector. Initial screening using a small-scale induction and purification method (Smith & Corcoran 1994; Chapter 2.13.2) showed that a strain TB1 transformant containing pGEX-4T-2 had high levels of IPTG-inducible expression of the 27.5 kD GST protein that could be detected following fractionation by SDS-PAGE (Chapter 2.15.1) both in total cell-free extract and in purified protein immobilised to glutathione-agarose (Figure 4.15, lanes 1-3). Fractionated glutathione-agarose immobilised protein from transformed strain TB1/pGST-PRH-26 yielded a faint band at approximately 76 kD, the expected size of the GST-PRH-26 fusion protein (Figure 4.15, lane 6). Although the binding capacity of glutathione-agarose is 5-fold less for a 76 kDa than a 27 kDa GST fusion protein (Frangioni & Neel 1993), inability to detect a strong IPTG-induced band at 76 kDa in total cell-free extracts also suggested that expression levels of the GST-PRH-26 fusion protein were insubstantial (Figure 4.15, lane 5). Similar tests confirmed that the fusion protein GST-PRH-26 was being expressed in three of the *E. coli* strain TB1/pGST-PRH-26 transformants at similarly low levels



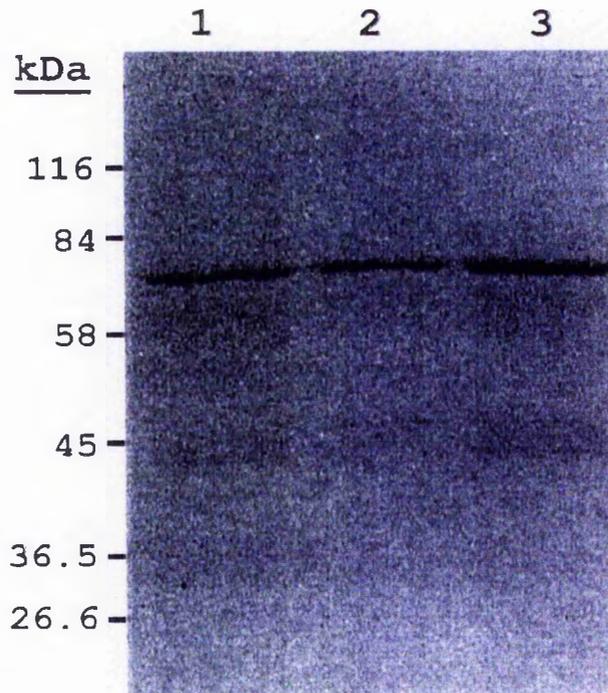
**Figure 4.15** SDS-PAGE analysis of expression and purification of GST and the fusion protein GST-PRH-26 in transformed *Escherichia coli* strain TB1 cells

*E. coli* strain TB1 cells transformed with a control vector pGEX-4T-2 or fusion protein plasmid pGST-PRH-26 were analysed for expression of the GST protein (lanes 1 - 3) or fusion protein GST-PRH-26 (lanes 4 - 6), respectively, using a small-scale screening protocol. 10  $\mu$ l cell-free extract from cells grown in LB for 4.25 h at 37°C (lanes 1 & 4), 10  $\mu$ l cell-free extract from cells grown in LB for 4.25 h at 37°C and then supplemented with 0.4 mM IPTG for 1.5 h at 30°C (lanes 2 & 5) and 10  $\mu$ l of a 50 % (w/v) glutathione-agarose bead suspension which was incubated with IPTG-induced cell-free extracts and washed extensively to allow purification (lanes 3 & 6) were fractionated by SDS-PAGE through a 10 % (w/v) polyacrylamide gel. Bands were visualised following Coomassie Brilliant Blue staining. Band positions of Kaleidoscope Prestained Standard molecular weight standards (Biorad, UK) are shown.

(results shown for a single transformant only).

Low yields of fusion protein expressed in *E. coli* can be attributed to a variety of causes, including insolubility and the formation of inclusion bodies, mechanical damage during cell disruption, degradation by proteases, and toxicity effects of expressed proteins on the host (di Guan et al. 1988; Frangioni & Neel 1993; Smith 1993; Riggs 1994; Smith & Corcoran 1994). Experiments were undertaken with a strain TB1/pGST-PRH-26 transformant to attempt to increase the yield of the GST-PRH-26. Using first a small-scale method of Smith and Corcoran (1994), parameters altered to increase solubility of GST-PRH-26 included: reducing incubation temperature from 37°C to 30°C; and altering IPTG induction time (0, 1.5, 2, 4 or 6 h), concentration (0.2 - 0.4 mM) and induction temperature (30°C or 37°C). Highest yields of the 76 kD fusion protein were obtained at 37°C after 6 h of induction with 0.4 mM IPTG, but expression levels were still low (results not shown).

Subsequently, an alternative protocol developed by Frangioni and Neel (1993) to purify insoluble GST fusion proteins was employed. Optimisation of the concentration of additives such as lysozyme for partial disruption of cells, the ionic detergent N-lauroylsarcosine for solubilisation, and DTT and the nonionic detergent Triton X-100 for binding to glutathione-agarose (Chapter 2.13.2) produced higher levels of purified GST-PRH-26 fusion protein that could be eluted from glutathione-agarose beads (Figure 4.16, lanes 1 & 2). After two elutions using a nondenaturing buffer containing reduced glutathione and Triton X-100, about 50 % of the fusion protein remained attached to the glutathione-agarose beads (Figure 4.16, lane 3). In the absence of Triton X-100 in the elution buffer, only 10 - 20 % of GST fusion proteins would be expected to elute under nondenaturing conditions (Frangioni & Neel 1993). The yield of purified GST-PRH-26 obtained from a 500 ml culture was approximately 0.1 mg/l,



**Figure 4.16** SDS-PAGE analysis of purified and eluted fusion protein GST-PRH-26

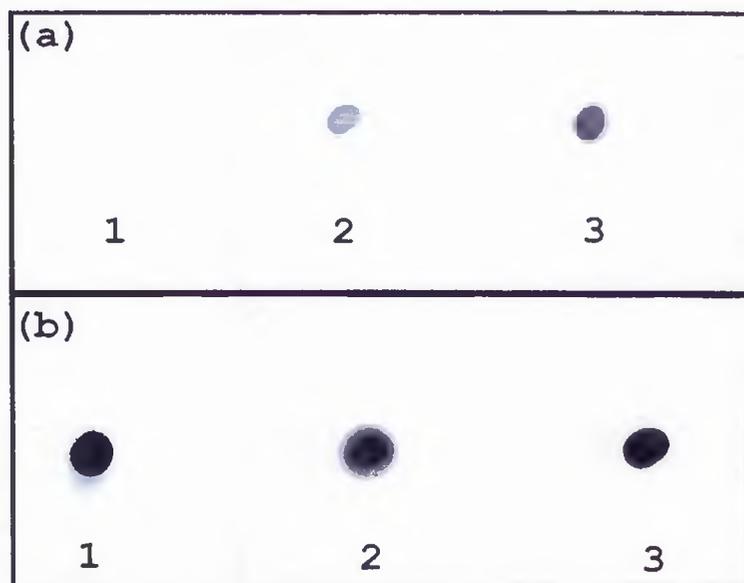
The 76 kDa fusion protein GST-PRH-26 was purified using 2.5 ml of a 50 % (w/v) glutathione-agarose bead slurry from a 500 ml culture of *Escherichia coli* strain TB1/pGST-PRH-26 cells which were induced in LB medium with 0.4 mM IPTG for 5 h at 37°C. The fusion protein was eluted from the glutathione-agarose beads using a non-denaturing buffer containing reduced glutathione. 10 µl of each of the following were fractionated by SDS-PAGE through a 10 % (w/v) polyacrylamide gel: first elution (total volume 0.6 ml) (lane 1); second elution (total volume 0.3 ml) (lane 2); and bead suspension following the two elutions (lane 3). Bands were visualised following Coomassie Brilliant Blue staining. Band positions of Prestained SDS-PAGE Standard Solution markers (Sigma, UK) are shown.

which is low compared with published yields for GST fusion proteins of 0.2 - 50 mg/l (Smith 1993).

Improvements in GST-PRH-26 fusion protein yield achieved using the solubilisation and purification methods of Frangioni and Neel (1993) suggested that initial low yields were due at least partially to insolubility of GST-PRH-26. However, SDS-PAGE analysis of whole cells and of pelleted insoluble material following cell disruption did not reveal high levels of IPTG-induced GST-PRH-26 fusion protein (results not shown). Although some degradation products of GST-PRH-26 could be detected after purification, the protease inhibitors phenylmethanesulfonyl fluoride (PMSF, 1 mM) and leupeptin (10  $\mu$ M) did not increase yield (results not shown) and instability of the fusion protein was excluded as a primary reason for low levels of expression. It is feasible that higher expression of GST-PRH-26 is toxic to *E. coli* train TB1.

#### *Production of rabbit anti-GST-PRH-26 polyclonal antibodies*

A New Zealand White rabbit was injected initially with 27.5  $\mu$ g of purified GST-PRH-26 as antigen to obtain polyclonal antibodies against the fusion protein (Chapter 2.14). Injection of up to 100  $\mu$ g of protein is recommended for antibody production in rabbits, but as little as 100 ng protein has been used successfully (Sambrook et al. 1989). Dot-blot western analysis was carried out with pre-immune serum and first antiserum collected seven weeks after the primary injection to determine cross-reactivity to GST-PRH-26. Using sera dilutions of 1/200, it was found that both pre-immune serum and antiserum cross-reacted with extract from *E. coli* strains JM96 and JM96/pPAPSR26, but that only the first bleed serum cross-reacted with purified GST-PRH-26 (Figure 4.17). These results suggested that antibodies to GST-PRH-26 were present in the first bleed serum but not in pre-immune serum, and that at a 1/200 dilution, antibodies



**Figure 4.17** Dot-blot western analysis using rabbit pre-immune serum and primary antiserum raised against the GST-PRH-26 fusion protein

Two nitrocellulose filters were dotted in duplicate with 2.5  $\mu$ l each of purified GST-PRH-26 fusion protein (1), cell-free extract from *Escherichia coli* strain JM96 (2) and cell-free extract from *E. coli* strain JM96/pPAPSR26 (3). Filters were exposed to a 1/200 dilution of either rabbit pre-immune serum (a) or rabbit antiserum extracted seven weeks of immunisation of the GST-PRH-26 fusion protein (b) as primary antibody for western analysis. Alkaline phosphatase-linked goat anti-rabbit IgG diluted 1/3000 was used as secondary antibody. Secondary antibody was detected by incubation for 1 min in a chromogenic substrate solution containing NBT and BCIP (Chapter 2.15.4).

present in both pre-immune and first antiserum detected other bacterial proteins. A boost injection of 21  $\mu$ g GST-PRH-26 was given to the rabbit to enhance anti-GST-PRH-26 antibody, and a final bleed taken after another two weeks. Preliminary western analysis of the final antiserum revealed the presence of several cross-reacting bands in fractionated strain JM96 and strain JM96/pPAPSR-26 extracts at 1/500 and 1/2000 serum dilutions, but no bands could be detected in SDS-PAGE fractionated *A. thaliana* leaf extracts at 1/200, 1/1000, 1/2500 and 1/5000 serum dilutions (results not shown). Further investigations are required to determine whether specific polyclonal antibodies to GST-PRH-26 are present in the final antiserum.

### 4.3 Conclusions

Functional rescue of the *Escherichia coli* strain JM81A (defective in APS kinase) to cysteine prototrophy has resulted in the isolation of 200 Apsk putative complementing *A. thaliana* cDNA clones, while functional rescue of *E. coli* strain JM96 (defective in PAPS reductase) to cysteine prototrophy has resulted in the isolation of 16 Papsr putative complementing *A. thaliana* cDNA clones. Initial characterisation of two Apsk clones, Apsk-14 and Apsk-15 (Section 4.2.2.1), and two Papsr clones, Papsr-19 and Papsr-20 (Section 4.2.4.1), revealed that these Apsk and Papsr clones are members of the same gene family that encode enzymes with structural similarity to bacterial and yeast PAPS reductase. Retransformation experiments confirmed that each of the plasmids pAPSK-14, pAPSK-15, pPAPSR-19, pPAPSR-20 and pPAPSR-26 could complement both *E. coli* mutant strains JM81A and JM96 (Section 4.2.5). Cysteine prototrophy in *E. coli* strain JM96 transformed with pPAPSR-19, pPAPSR-26 or pPAPSR43 was associated with restoration of PAPS reductase activity that, unlike wild-type bacterial PAPS reductase activity, was thioredoxin-independent (Table 4.2).

Additional sequence information from clones Papsr-19, Papsr-26 and Papsr-43 was obtained by Dr J. Gutierrez-Marcos and Mr E. Campbell in this laboratory, and the clones were renamed PRHs (PAPS reductase homologues) (Gutierrez-Marcos et al. 1996). PRH-19, PRH-26 and PRH-43 are different members of a small gene family that encode enzymes which in addition to structural similarity with PAPS reductase also have a C-terminal domain that resembles the thioredoxin-like moiety of PDI (Section 4.2.6). The three PRHs, which may all be targeted to the chloroplast, were capable of reducing the sulpho group of both APS and PAPS to acid-volatile radioactivity, with APS the preferred substrate (Gutierrez-Marcos et al. 1996). The *in vivo* catalytic activity of the PRHs is not clear.

Sulphate assimilation enzymes with structural similarity to the *A. thaliana* PRH family have not been reported in *E. coli* or in other organisms. *In vitro* enzymatic assays (Gutierrez-Marcos et al. 1996) suggest that the PRHs behave as APS reductases, which can explain the observed functional complementation of both APS kinase- and PAPS reductase-deficient *E. coli* strains JM81A and JM96, respectively. Therefore a novel mechanism for sulphate assimilation in plants can be proposed, whereby a modified PP motif of PAPS reductase evolved to metabolise APS, so that the sulpho group of APS, rather than PAPS, could be converted directly to sulphite (Section 4.2.2.1). The fusion of a thioredoxin-like moiety to the APS-reducing enzyme allowed thioredoxin-independent catalysis of APS. Independent researchers describing PRH-like *A. thaliana* proteins APR1, -2 and -3 have proposed similarly that the APRs are APS reductases (Setya et al. 1996). An alternative pathway, whereby the PRHs catalyse an APS sulphotransferase reaction, transferring the sulpho group of APS to an acceptor molecule that is maintained in correct conformational state by the C-terminal thioredoxin-like domain, cannot be excluded from the available data (Section 4.2.7).

Isolation of the PRHs provides evidence for the existence of a pathway of sulphate assimilation in plants which does not occur in bacteria, fungi or yeasts. Whether the free intermediate pathway of sulphate assimilation, requiring APS kinase and "conventional" PAPS reductase activity, also operates in plants is not known. It will be possible now to attempt to modify expression of genes encoding the PRHs genes in order to gain further insight into their role and achieve a greater understanding of plant sulphate assimilation pathways.

## CHAPTER 5: CONCLUDING REMARKS

Three decades ago, Thompson (1967) observed that knowledge of higher plant sulphur metabolism was deficient in comparison with carbon, nitrogen and phosphorus metabolism. Major limitations preventing biochemical delineation of the pathway of reductive sulphate assimilation have been the "promiscuous reactions of cox intermediates" (Schiff & Hodson 1973), the lability of enzyme preparations and the questionable physiological significance of *in vitro* enzyme reactions (Brunold 1990). It was envisaged that the isolation of genes encoding plant sulphate assimilation enzymes would be required to understand the metabolic steps involved (Schiff & Hodson 1973; Brunold 1992; Schmidt & Jäger 1992). The present study was undertaken to clone *Arabidopsis thaliana* sulphate assimilation genes in order to facilitate attainment of this understanding.

### *Serine acetyltransferase*

Functional complementation of *Escherichia coli* strain JM15 (deficient in serine acetyltransferase) with the *A. thaliana* cDNA expression library  $\lambda$ YES resulted in the isolation of 110 putative serine acetyltransferase Sat clones (Chapter 3). Identity of the SAT-1 protein deduced from clone Sat-1a was confirmed using an *in vitro* serine acetyltransferase activity assay and by sequence similarity with bacterial and plant serine acetyltransferases. At least three members of a small multigene family (clones Sat-1, Sat-52 and Sat-53) are present amongst the isolated Sat clones, and sequence analysis suggests that the deduced isoforms encoded by clones Sat-1, Sat-52 and Sat-53 are targeted to separate subcellular compartments (Ruffet et al. 1995; Roberts & Wray 1996; Howarth et al. 1997).

Availability of new plant serine acetyltransferase genes

provides considerable scope for future investigations that will yield useful and interesting information. Research priorities are to determine empirically the subcellular location of the three isoforms so that roles within their respective compartments can be investigated; to examine the multienzyme complex between serine acetyltransferase and *O*-acetylserine (thiol)-lyase, which has been described *in vitro* for *Citrullus vulgaris* (Saito et al. 1995) and which will provide an understanding of the catalytic mechanism of cysteine formation; and to study the *in vivo* transcriptional activity of the genes and *in vitro* catalytic responses of the serine acetyltransferase isoforms in the presence of different sulphur and nitrogen compounds so that involvement of the enzyme in regulation of the converging pathways of sulphur and nitrogen metabolism can be assessed. Transgenic plants containing sense or antisense-DNA constructs of the serine acetyltransferase genes may prove useful in addressing these aims.

*APS reductase and a new hypothesis for a pathway of higher plant reductive sulphate assimilation*

Functional complementation of the *E. coli* cysteine auxotrophic strains JM81A (deficient in APS kinase) and JM96 (deficient in PAPS reductase) with the  $\lambda$ YES *A. thaliana* cDNA library resulted in the isolation of 200 Apsk and 16 Papsr *A. thaliana* cDNA clones, respectively (Chapter 4). Rather than encoding anticipated APS kinase and PAPS reductase proteins, it was discovered that the two Apsk clones that were chosen for analysis and the 16 Papsr clones are members of a single multigene family that encode novel PRH enzymes (Gutierrez-Marcos et al. 1996). Three deduced isoforms (PRH-19 [= PAPSR-19, APSK-14], PRH-26 [= PAPSR-26, APSK-15] and PRH-43 [= PAPSR-43, PAPSR-20]) show preference for APS over PAPS as substrate when using an *in vitro* reductase assay in which <sup>35</sup>S-APS or <sup>35</sup>S-PAPS is converted to acid-volatile radioactivity (Gutierrez-Marcos et al. 1996). Each of the PRHs have a C-

terminal thioredoxin-like domain, and current evidence indicates that the PRHs are thioredoxin-independent APS reductases (Gutierrez-Marcos et al. 1996; Setya et al. 1996).

With the isolation and preliminary characterisation of *A. thaliana* putative APS reductases, an hypothesis can be proposed for a novel enzymatic pathway of higher plant reductive sulphate assimilation (Gutierrez-Marcos et al. 1996; Setya et al. 1996): activation of intracellular sulphate is catalysed by ATP sulphurylase, forming APS; the sulpho group of APS is reduced by APS reductase directly to free sulphite; free sulphite is reduced further to free sulphide by sulphite reductase; and finally, sulphide is incorporated into cysteine by *O*-acetylserine (thiol)-lyase (possibly in a complex with serine acetyltransferase). This hypothesis can serve as a working model for further investigations. It will be important to determine the kinetic mechanism of the putative APS reductases and the function of their C-terminal thioredoxin-like domain. As biochemical studies have failed to isolate stable PRH-like enzymes, purification of functional enzyme using fusion protein expression vectors such as pGEX (Chapter 4.2.8) will be necessary. It has been demonstrated that *Spinacia oleracea* leaf chloroplast extract has the ability to form <sup>35</sup>S-cysteine directly from <sup>35</sup>S-APS (Schürmann & Brunold 1980), and N-terminal sequence analysis of each of the three PRH isoforms suggest a chloroplastic location (Gutierrez-Marcos et al. 1996). The subcellular location of the PRHs, and other sulphate assimilation enzymes, requires further study so that compartmentalisation of reductive sulphate assimilation and cysteine biosynthesis can be ascertained. It is not clear if plants also possess "conventional" PAPS reductase enzymes which would allow reductive sulphate assimilation to proceed also via a pathway analogous to that in bacteria. Specific down-regulation of genes encoding APS kinase and the PRHs using antisense-RNA techniques may be useful for determining physiological functions of these enzymes, although the

presence of multigene families may complicate such experiments (Saito et al. 1994a). Finally, if PRH-26 polyclonal antibodies (Chapter 4.2.8) are found to bind PRHs specifically, such antibodies will be useful in conjunction with PRH clones as probes to study the translational and transcriptional regulation of the putative APS reductases under different regimes of nutrient supply and environmental conditions.

#### *Productive years in plant sulphate assimilation research*

At the inception of the work described here, only two plant sulphate assimilation genes, both encoding *O*-acetylserine (thiol)-lyase, had been reported in the literature (Römer et al. 1992; Saito et al. 1992). In the present study, functional complementation of *E. coli* cysteine auxotrophic strains has been an effective method for isolating higher plant genes encoding a novel putative APS reductase, which has no known homologue in *E. coli*, and serine acetyltransferase, which is present in *E. coli*. These *A. thaliana* serine acetyltransferase and putative APS reductase genes augment the list of published reports of plant sulphate assimilation genes isolated in recent years (Römer et al. 1992; Saito et al. 1992; Hell et al. 1993; Hesse & Altmann 1993; Rolland et al. 1993a; Saito et al. 1993b; Arz et al. 1994; Hell et al. 1994; Jain & Leustek 1994; Klonus et al. 1994; Leustek et al. 1994; Noji et al. 1994; Saito et al. 1994b; Barroso et al. 1995; Bogdanova et al. 1995; Brander et al. 1995; Klonus et al. 1995; Murillo & Leustek 1995; Ruffet et al. 1995; Saito et al. 1995; Smith et al. 1995; Gutierrez-Marcos et al. 1996; Logan et al. 1996; Roberts & Wray 1996; Setya et al. 1996; Takahashi et al. 1996; Youssefian et al. 1993; Howarth et al. 1997). The surge of interest in this field has provided an abundance of molecular tools that can be employed to find answers to unsolved questions about higher plant reductive sulphate assimilation.

## REFERENCES

- Abrams, W.R., Schiff, J.A. (1973) Studies of sulfate utilization by algae. 11. An enzyme-bound intermediate in the reduction of adenosine-5'-phosphosulfate (APS) by cell-free extracts of wild-type *Chlorella* and mutants blocked for sulfate reduction. *Arch. Microbiol.* 94: 1-10.
- Adams, C.A., Sheard, R.W. (1966) Alterations in the nitrogen metabolism of *Medicago sativa* and *Dactylis glomerata* as influenced by potassium and sulfur nutrition. *Can. J. Plant Sci.* 46: 671-680.
- Aketagawa, J., Tamura, G. (1980) Ferredoxin-sulfite reductase from spinach. *Agric. Biol. Chem.* 44: 2371-2378.
- Allen, A., Maskell, D. (1996) The identification, cloning and mutagenesis of a genetic locus required for lipopolysaccharide biosynthesis in *Bordetella pertussis*. *Mol. Microbiol.* 19: 37-52.
- Allignet, J., Loncle, V., Simenel, C., Delepierre, M., El Solh, N. (1993) Sequence of a staphylococcal gene, *vat*, encoding an acetyltransferase inactivating the A-type compounds of viginiamycin-like antibiotics. *Gene* 130: 91-98.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment tool. *J. Mol. Biol.* 215: 403-410.
- Anderson, J.W. (1980) Assimilation of inorganic sulfate into cysteine. In: *The biochemistry of plants*, Vol. 5, Mifflin, B.J. (ed.), Academic Press, New York, pp 203-223.
- Annunziato, P.W., Wright, L.F., Vann, W.F., Silver, R.P. (1995) Nucleotide sequence and genetic analysis of the *neuD* and *neuB* genes in region 2 of the polysialic acid gene cluster of *Escherichia coli* K1. *J. Bacteriol.* 177: 312-319.
- Arber, W., Enquist, L., Hohn, B., Murray, N., Murray, K. (1983) Experimental methods for use with lambda. In: *Lambda II*, Hendrix, R.W., Roberts, J.W., Stahl, F.W., Weisberg, R.A. (eds), Cold Spring Harbor Laboratory, Cold Spring Harbor, pp 433-466.
- Arndt, K.T., Styles, C.A., Fink, G.R. (1989) A suppressor of a *HIS4* transcription defect encodes a protein with homology to the

- catalytic subunit of protein phosphatases. *Cell* 56: 527-537.
- Arz, H.E., Gisselmann, G., Schiffmann, S., Schwenn, J.D. (1994) A cDNA for adenyllyl sulphate (APS)-kinase from *Arabidopsis thaliana*. *Biochim. Biophys. Acta* 1218: 447-452.
- Asada, K. (1967) Purification and properties of a sulfite reductase from leaf tissue. *J. Biol. Chem.* 242: 3646-3654.
- Asada, K., Tamura, G., Bandurski, R.S. (1968) Methyl viologen-linked reductase from spinach leaves: A hemoprotein. *Biochem. Biophys. Res. Comm.* 30: 554-559.
- Asahi, T., Bandurski, R.S., Wilson, L.G. (1961) Yeast sulfate-reducing system. II. Enzymatic reduction of protein disulfide. *J. Biol. Chem.* 236: 1830-1835.
- Ascaño, A., Nicholas, D.J.D. (1977) Purification and properties of O-acetyl-L-serine sulphydrylase from wheat leaves. *Phytochemistry* 16: 889-893.
- Baecker, P.A., Wedding, R.T. (1980) Purification of serine acetyltransferase, a component of a multienzyme complex, by immunoadsorption and selective dissociation of the complex. *Anal. Biochem.* 102: 16-21.
- Baev, N., Kondorosi, A. (1992) Nucleotide sequence of the *Rhizobium meliloti nodL* gene located in locus n5 of the *nod* regulon. *Plant Mol. Biol.* 18: 843-846.
- Bairoch, A. (1993) The PROSITE dictionary of sites and patterns in proteins, its current status. *Nucleic Acids Res.* 21: 3097-3103.
- Bairoch, A., Boeckmann, B. (1994) The SWISS-PROT protein sequence data bank: current status. *Nucleic Acids Res.* 22: 3578-3580.
- Bairoch, A., Bucher, P. (1994) PROSITE: recent developments. *Nucleic Acids Res.* 22: 3583-3589.
- Balharay, G.J.E., Nicholas, D.J.D. (1970) ATP-sulphurylase in spinach leaves. *Biochim. Biophys. Acta* 220: 513-524.
- Bandurski, R.S., Wilson, L.G., Squires, C.L. (1956) The mechanism of "active sulfate" formation. *J. Am. Chem. Soc.* 78: 6408-6409.

- Bardwell, J.C.A. (1994) Building bridges: disulphide bond formation in the cell. *Mol. Microbiol.* 14: 199-205.
- Barnes, W.M. (1977) Plasmid detection and sizing in single colony lysates. *Science* 195: 393-394.
- Barney, P.E., Bush, L.P. (1985) Interaction of nitrate and sulfate reduction in tobacco. 1. Influence of availability of nitrate and sulfate. *J. Plant Nutrition* 8: 505-515.
- Barroso, C., Vega, J.M., Gotor, C. (1995) A new member of the cytosolic O-acetylserine(thiol)lyase gene family in *Arabidopsis thaliana*. *FEBS Lett.* 363: 1-5.
- Bell, C.I., Cram, W.J., Clarkson, D.T. (1994) Compartmental analysis of  $^{35}\text{SO}_4^{2-}$  exchange kinetics in roots and leaves of a tropical legume *Macroptilium atropurpureum* cv. Siratro. *J. Exp. Bot.* 45: 879-886.
- Belyavsky, A., Vinogradova, T., Rajewsky, K. (1989) PCR-based cDNA library construction: general cDNA libraries at the level of a few cells. *Nucleic Acids Res.* 17: 2919-2932.
- Bennett, R.N., Wallsgrave, R.M. (1994) Secondary metabolites in plant defence mechanisms. *New Phytol.* 127: 617-633.
- Benson, D.A., Boguski, M., Lipman, D.J., Ostell, J. (1994) GenBank. *Nucleic Acids Res.* 22: 3441-3444.
- Berendt, U., Haverkamp, T., Prior, A., Schwenn, J.D. (1995) Reaction mechanism of thioredoxin:3'-phospho-adenylylsulfate reductase investigated by site-directed mutagenesis. *Eur. J. Biochem.* 233: 347-356.
- Bethesda Research Laboratories (1986) BRL pUC host: *E. coli* DH5 $\alpha$ <sup>TM</sup> competent cells. *Bethesda Res. Lab. Focus* 8(2): 9.
- Bloemberg, G.V., Thomas-Oates, J.E., Lugtenberg, B.J.J., Spaink, H.P. (1994) Nodulation protein NodL of *Rhizobium leguminosarum* O-acetylates lipo-oligosaccharides, chitin fragments and N-acetylglucosamine *in vitro*. *Mol. Microbiol.* 11: 793-804.
- Bogdanova, N., Bork, C., Hell, R. (1995) Cysteine biosynthesis in plants: isolation and functional identification of a cDNA encoding a serine acetyltransferase from *Arabidopsis thaliana*. *FEBS Lett.* 358: 43-47.

- Borck, K., Beggs, J.D., Brammar, W.J., Hopkins, A.S., Murray, N.E. (1976) The construction in vitro of transducing derivatives of phage lambda. *Mol. Gen. Genet.* 146: 199-207.
- Borek, E., Ryan, A. (1958) The transfer of irradiation-elicited induction in a lysogenic organism. *Proc. Natl Acad. Sci. USA* 44: 374-377.
- Borek, E., Ryan, A. (1973) Lysogenic induction. In: *Progress in nucleic acid research and molecular biology*, Vol. 13, Davidson, J.N., Cohn, W.E. (eds), Academic Press, New York, pp 249-293.
- Bork, P., Koonin, E.V. (1994) A P-loop-like motif in a widespread ATP pyrophosphatase domain: Implications for the evolution of sequence motifs and enzyme activity. *Proteins* 20: 347-355.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.
- Brander, K.A., Owttrim, G.W., Brunold, C. (1995) Isolation of a cDNA (EMBL X85803) encoding a putative chloroplastic isoform of cysteine synthase from maize. *Plant Physiol.* 108: 1748.
- Breton, A., Surdin-Kerjan, Y. (1977) Sulfate uptake in *Saccharomyces cerevisiae*: Biochemical and genetics study. *J. Bacteriol.* 132: 224-232.
- Brunold, C. (1990) Reduction of sulfate to sulfide. In: *Sulfur nutrition and sulfur assimilation in higher plants*, Rennenberg, H., Brunold, C., De Kok, L.J., Stulen, I. (eds), SPB Academic, The Hague, pp 13-31.
- Brunold, C. (1993) Regulatory interactions between sulfate and nitrate assimilation. In: *Sulfur nutrition and assimilation in higher plants*, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Rauser, W.E. (eds), SPB Academic, The Hague, pp 61-75.
- Brunold, C., Schiff, J.A. (1976) Studies of sulfate utilization by algae. 15. Enzymes of assimilatory sulfate reduction in *Euglena* and their cellular localization. *Plant. Physiol.* 57: 430-436.
- Brunold, C., Suter, M. (1982) Intracellular localization of serine acetyltransferase in spinach leaves. *Planta* 155: 321-327.
- Brunold, C., Suter, M. (1984) Regulation of sulfate assimilation by

- nitrogen-nutrition in the duckweed *Lemna minor* L. *Plant Physiol.* 76: 579-583.
- Brunold, C., Suter, M., Lavanchy, P. (1987) Effect of high and low sulfate concentrations on adenosine 5'-phosphosulphate sulfotransferase activity from *Lemna minor*. *Physiol. Plant.* 70: 168-174.
- Bullock, W.O., Fernandez, J.M., Short, J.M. (1987) XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *BioTechniques* 5: 376-379.
- Burnell, J.N. (1984) Sulfate assimilation in C<sub>4</sub> plants. *Plant Physiol.* 75: 873-875.
- Burnell, J.N., Anderson, J.W. (1973) Adenosine 5'-sulphatophosphate kinase activity in spinach leaf tissue. *Biochem. J.* 134: 565-579.
- Burnell, J.N., Whatley, F.R. (1975) A new, rapid and sensitive assay for adenosine 5'-phosphosulphate (APS) kinase. *Anal. Biochem.* 68: 281-288.
- Buwalda, F., De Kok, L.J., Stulen, I. (1992) Flux of atmospheric H<sub>2</sub>S to spinach leaves can be affected by the supply of O-acetylserine. *Phyton* 32: 13-16.
- Cacco, G., Saccomani, M., Ferrari, G. (1983) Changes in the uptake and assimilation efficiency for sulfate and nitrate in maize hybrids selected during the period 1930 through 1975. *Physiol. Plant.* 58: 171-174.
- Canter Cremers, H.C.J., Spaik, H.P., Wijfjes, A.H.M., Pees, E., Wijffelman, C.A., Okker, R.J.H., Lugtenberg, B.J.J. (1989) Additional nodulation genes on the Sym plasmid of *Rhizobium leguminosarum* biovar *viciae*. *Plant Mol. Biol.* 13: 163-174.
- Castellazzi, M., George, J., Buttin, G. (1972a) Prophage induction and cell division in *E. coli*. I. Further characterization of the thermosensitive mutation *tif-1* whose expression mimics the effect of UV irradiation. *Mol. Gen. genet.* 119: 139-152.
- Castellazzi, M., George, J., Buttin, G. (1972b) Prophage induction and cell division in *E. coli*. II. Linked (*recA*, *zab*) and unlinked (*lex*) suppressors of *tif-1*-mediated induction and filamentation. *Mol. Gen. Genet.* 119: 153-174.

- Cervantes, E., Sharma, S.B., Maillet, F., Vasse, J., Truchet, G., Rosenberg, C. (1989) The *Rhizobium meliloti* host range *nodQ* gene encodes a protein which shares homology with translation elongation and initiation factors. *Mol. Microbiol.* 3: 745-755.
- Chee, P.P., Drong, R.F., Slightom, J.L. (1991) Using polymerase chain reaction to identify transgenic plants. In: *Plant molecular biology manual*, Gelvin, S.B., Schilperoort, R.A. (eds), Kluwer Academic, Dordrecht, pp C3.1-28.
- Chen, E.Y., Seeburg, P.H. (1985) Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. *DNA* 4: 165-170.
- Cherest, H., Kerjan, P., Surdin-Kerjan, Y. (1987) The *Saccharomyces cerevisiae* *MET3* gene: Nucleotide sequence and relationship of the 5' non-coding region to that of *MET25*. *Mol. Gen. Genet.* 210: 307-313.
- Chou, P.Y., Fasman, G.D. (1978) Empirical predictions of protein conformations. *Ann. Rev. Biochem.* 47: 251-276.
- Chung, C.T., Niemela, S.L., Miller, R.H. (1989) One-step preparation of competent *Escherichia coli*: Transformation and storage of bacterial cells in the gene solution. *Proc. Natl Acad. Sci. USA* 86: 2172-2175.
- Clarkson, D.T., Hawkesford, M.J., Davidian, J.-C. (1993) Membrane and long-distance transport of sulfate. In: *Sulfur nutrition and assimilation in higher plants*, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Rauser, W.E. (eds), SPB Academic, The Hague, pp 3-19.
- Clarkson, D.T., Saker, L.R., Purves, J.V. (1989) Depression of nitrate and ammonium transport in barley plants with diminished sulphate status. Evidence of co-regulation of nitrogen and sulphate intake. *J. Exp. Bot.* 40: 953-963.
- Coleman, J., Raetz, C.R.H. (1988) First committed step of lipid A biosynthesis in *Escherichia coli*: Sequence of the *lpxA* gene. *J. Bacteriol.* 170: 1268-1274.
- Collins, J.F., Coulson, A.F.W. (1987) Molecular sequence comparison and alignment. In: *Nucleic acid and protein sequence analysis: A practical approach*, Bishop, M.J., Rawlings, C.J. (eds), IRL Press, Oxford, pp 323-358.

- Contesse, G., Crépin, M., Gros, F., Ullmann, A. Monod, J. (1970) On the mechanism of catabolite repression. In: The lactose operon, Beckwith, J.R., Zipser, D. (eds), Cold Spring Harbor Laboratory, Cold Spring Harbor, pp 401-415.
- Correa-Rotter, R., Mariash, C.N., Rosenberg, M.E. (1992) Loading and transfer control for northern hybridization. *BioTechniques* 12: 154-158.
- Cram, W.J. (1990) Uptake and transport of sulfate. In: Sulfur nutrition and sulfur assimilation in higher plants, Rennenberg, H., Brunold, C., De Kok, L.J., Stulen, I. (eds), SPB Academic, The Hague, pp 3-11.
- Crooke, H., Cole, J. (1995) The biogenesis of *c*-type cytochromes in *Escherichia coli* requires a membrane-bound protein, DipZ, with a protein disulphide isomerase-like domain. *Mol. Microbiol.* 15: 1139-1150.
- Croy, E.J., Ikemura, T., Shirsat, A., Croy, R.R.D. (1993) Plant nucleic acids. In: Plant molecular biology Labfax, Croy, R.R.D. (ed.), BIOS Scientific, Oxford, pp 21-48.
- Dallam, R.D. (1987) Sulfate metabolism. I. Sulfate uptake and redistribution of acid rain sulfate by edible plants. *Life Sciences* 40: 1147-1154.
- D'Aquila, R.T., Bechtel, L.J., Videler, J.A., Eron, J.J., Gorczyca, P., Kaplan, J.C. (1991) Maximizing sensitivity and specificity of PCR by pre-amplification heating. *Nucleic acids Res.* 19: 3749.
- Darby, N.J., Creighton, T.E. (1995) Functional properties of the individual thioredoxin-like domains of protein disulfide isomerase. *Biochemistry* 34: 11725-11735.
- DasSarma, S., Tischer, E., Goodman, H.M. (1986) Plant glutamine synthetase complements a *glnA* mutation in *Escherichia coli*. *Science* 232: 1242-1244.
- Davies, J.P., Yildiz, F., Grossman, A.R. (1994) Mutants of *Chlamydomonas* with aberrant responses to sulfur deprivation. *Plant Cell* 6: 53-63.
- Davis, L.G.D., Dibner, M.D., Batley, J.F. (1986) Basic methods in molecular biology, Elsevier, New York.

- Dean, C., Sjodin, C., Page, T., Jones, J., Lister, C. (1992) Behaviour of the maize transposable element *Ac* in *Arabidopsis thaliana*. *Plant J.* 2: 69-81.
- Dean, C., Tamaki, S., Dunsmuir, P., Favreau, M., Katayama, C., Dooner, H., Bedbrook, J. (1986) Messenger-RNA transcripts of several plant genes are polyadenylated at multiple sites in vivo. *Nucleic Acids Res.* 14: 2229-2240.
- DeBoer, D.L., Duke, S.H. (1982) Effects of sulphur nutrition on nitrogen and carbon metabolism in lucerne (*Medicago sativa* L.). *Physiol. Plant.* 54: 343-350.
- Delauney, A.J., Verma, D.P.S. (1990) A soybean gene encoding  $\Delta^1$ -pyrroline-5-carboxylate reductase was isolated by functional complementation in *Escherichia coli* and is found to be osmoregulated. *Mol. Gen. Genet.* 221: 299-305.
- Denk, D., Böck, A. (1987) L-Cysteine biosynthesis in *Escherichia coli*: Nucleotide sequence and expression of the serine acetyltransferase (*cysE*) gene from the wild-type and a cysteine-excreting mutant. *J. Gen. Microbiol.* 133: 515-525.
- Devereux, J., Haeberli, P., Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12: 387-395.
- Dicker, I.B., Seetharam, S. (1991) Cloning and nucleotide sequence of the *firA* gene and the *firA200(Ts)* allele from *Escherichia coli*. *J. Bacteriol.* 173: 334-344.
- Dicker, I.B., Seetharam, S. (1992) What is known about the structure and function of the *Escherichia coli* protein FirA? *Mol. Microbiol.* 6: 817-823.
- Diessner, W., Schmidt, A. (1981) Isoenzymes of cysteine synthase in the cyanobacterium *Synechococcus* 6301. *Z. Pflanzenphysiol.* 102: 57-68.
- di Guan, C., Li, P., Riggs, P.D., Inouye, H. (1988) Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* 67: 21-30.
- Dijkshoorn, W., Van Wyk, A.L. (1967) The sulphur requirements of plants as evidenced by the sulphur-nitrogen ratio in the organic matter:

A review of published data. *Plant & Soil* 26: 129-157.

- Dittrich, A.P.M., Pfanz, H., Heber, U. (1992) Oxidation and reduction of sulfite by chloroplasts and formation of sulfite addition compounds. *Plant Physiol.* 98: 738-744.
- Dobrogosz, W.J. (1981) Enzymatic activity. In: Manual of methods for general bacteriology, Gerhardt, P. (ed.), American Society for Microbiology, Washington DC, pp 365-392.
- Doolittle, R.F. (1994) Convergent evolution: the need to be explicit. *Trends Biochem. Sci.* 19: 15-18.
- Dorit, R.L., Ohara, O., Hwang, C.B.-C. (1991) Direct DNA sequencing of PCR products. In: Current protocols in molecular biology, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (eds), John Wiley & Sons, USA, pp 15.2.1-15.2.11.
- Downie, J.A. (1989) The *nodL* gene from *Rhizobium leguminosarum* is homologous to the acetyl transferases encoded by *lacA* and *cySE*. *Mol. Microbiol.* 3: 1649-1651.
- Dreyfuss, J., Monty, K.J. (1963) The biochemical characterization of cysteine-requiring mutants of *Salmonella typhimurium*. *J. Biol. Chem.* 238: 1019-1024.
- Droux, M., Martin, J., Sajus, P., Douce, R. (1992) Purification and characterization of O-acetylserine (thiol) lyase from spinach chloroplasts. *Arch. Biochem. Biophys.* 295: 379-390.
- Eaton, S.V. (1935) Influence of sulphur deficiency on the metabolism of the soy bean. *Bot. Gaz.* 97: 68-100.
- Eaton, S.V. (1941) Influence of sulphur deficiency on metabolism of the sunflower. *Bot. Gaz.* 102: 536-556.
- Eaton, S.V. (1942) Influence of sulphur deficiency on metabolism of black mustard. *Bot. Gaz.* 104: 306-315.
- Eaton, S.V. (1951) Effects of sulphur deficiency on growth and metabolism of tomato. *Bot. Gaz.* 112: 300-307.
- Edwards, J.B.D.M., Delort, J., Mallet, J. (1991) Oligodeoxyribonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by *in vitro*

- amplification. *Nucleic Acids Res.* 19: 5227-5232.
- Elledge, S.J., Mulligan, J.T., Ramer, S.W., Spottswood, M., Davis, R.W. (1991)  $\lambda$ YES: A multifunctional cDNA expression vector for the isolation of genes by complementation of yeast and *Escherichia coli* mutations. *Proc. Natl Acad. Sci. USA* 88: 1731-1735.
- Ellis, R.J. (1969) Sulphate activation in higher plants. *Planta* 88: 34-42.
- Emmert, D.B., Stoehr, P.J., Stoesser, G., Cameron, G.N. (1994) The European Bioinformatics Institute (EBI) databases. *Nucleic Acids Res.* 22: 3445-3449.
- Engbrecht, J., Brent, R., Kaderbhai, M.A. (1991) Minipreps of plasmid DNA. In: *Current protocols in molecular biology*, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (eds), John Wiley & Sons, USA, pp 1.6.1-1.6.10.
- Eppendorfer, W. (1968) The effect of nitrogen and sulphur on changes in nitrogen fractions of barley plants at various early stages of growth and on yield and amino acid composition of grain. *Plant & Soil* 29: 424-438.
- Evans, D.J., Jones, R., Woodley, P.R., Wilborn, J.R., Robson, R.L. (1991) Nucleotide sequence and genetic analysis of the *Azotobacter chroococcum nifUSVWZM* gene cluster, including a new gene (*nifP*) which encodes a serine acetyltransferase. *J. Bacteriol.* 173: 5457-5469.
- Fankhauser, H., Brunold, C. (1978) Localization of adenosine 5'-phosphosulfate sulfotransferase in spinach leaves. *Planta* 143: 285-289.
- Farquhar, R., Honey, N., Murant, S.J., Bossier, P., Schultz, L., Montgomery, D., Ellis, R.W., Freedman, R.B., Tuite, M.F. (1991) Protein disulfide isomerase is essential for viability in *Saccharomyces cerevisiae*. *Gene* 108: 81-89.
- Feinberg, A.P., Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6-13.
- Feinberg, A.P., Vogelstein, B. (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity:

Addendum. Anal. Biochem. 137: 266-267.

- Feng, D.-F., Doolittle, R.F. (1990) Progressive alignment and phylogenetic tree construction of protein sequences. *Methods Enzymol.* 183: 375-387.
- Ferrari, G., Cacco, G., Bottacin, A., Saccomani, M. (1990) Energetics of sulfate uptake by plants as related to previous sulfate availability. In: *Sulfur nutrition and sulfur assimilation in higher plants*, Rennenberg, H., Brunold, C., De Kok, L.J., Stulen, I. (eds), SPB Academic, The Hague, pp 181-183.
- Ferretti, M., Merlo, L., Passera, C., Ghisi, R. (1995) Influence of irradiance on enzymes of sulphate and nitrate assimilation pathways in maize leaves. *Plant Physiol. Biochem.* 33: 111-114.
- Fiedler, G., Pajatsch, M., Böck, A. (1996) Genetics of a novel starch utilisation pathway present in *Klebsiella oxytoca*. *J. Mol. Biol.* 256: 279-291.
- Finken, M., Sobek, A., Symmons, P., Kunz, W. (1994) Characterization of the complete protein disulfide-isomerase gene of *Schistosoma mansoni* and identification of the tissues of its expression. *Mol. Biochem. Parasitol.* 64: 135-144.
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O., Venter, J.C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269: 496-512.
- Frangioni, J.V., Neel, B.G. (1993) Solubilization and purification of enzymatically active glutathione *S*-transferase (pGEX) fusion proteins. *Anal. Biochem.* 210: 179-187.
- Freedman, R.B., Hirst, T.R., Tuite, M.F. (1994) Protein disulphide isomerase: building bridges in protein folding. *Trends Biochem. Sci.* 19: 331-336.

- Friedrich, J.W., Schrader, L.E. (1978) Sulfur deprivation and nitrogen metabolism in maize seedlings. *Plant Physiol.* 61: 900-903.
- Frisch, D.A., Tommey, A.M., Gengenbach, B.G., Somers, D.A. (1991) Direct genetic selection of a maize cDNA for dihydrodipicolinate synthase in an *Escherichia coli* *dapA*<sup>-</sup> auxotroph. *Mol. Gen. Genet.* 228: 287-293.
- Frohman, M.A., Dush, M.K., Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA* 85: 8998-9002.
- Frommer, W.B., Ninnemann, O. (1995) Heterologous expression of genes in bacterial, fungal, animal, and plant cells. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 46: 419-444.
- Fry, B.A. (1963) The infection of *Escherichia coli* with lambda ( $\lambda_{22}$ ) phage and the establishment of lysogeny. *J. Gen. Microbiol.* 31: 297-309.
- Fukuchi, J., Kashiwagi, K., Takio, K., Igarashi, K. (1994) Properties and structure of spermidine acetyltransferase in *Escherichia coli*. *J. Biol. Chem.* 269: 22581-22585.
- Furlong, C.E. (1987) Osmotic-shock-sensitive transport systems. In: *Escherichia coli* and *Salmonella typhimurium*. Cellular and molecular biology, Vol. 1, Neidhardt, F.C. (ed.), American Society for Microbiology, Washington DC, pp 768-796.
- Gagnon, Y., Breton, R., Putzer, H., Pelchat, M., Grunberg-Manago, M., Lapointe, J. (1994) Clustering and co-transcription of the *Bacillus subtilis* genes encoding the aminoacyl-tRNA synthetases specific for glutamate and for cysteine and the first enzyme for cysteine biosynthesis. *J. Biol. Chem.* 269: 7473-7482.
- Garnier, J., Osguthorpe, D.J., Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120: 97-120.
- Gavel, Y., von Heijne, G. (1990) A conserved cleavage-site motif in chloroplast transit peptides. *FEBS Lett.* 261: 455-458.
- George, D.G., Barker, W.C., Mewes, H.-W., Pfeiffer, F., Tsugita, A. (1994) The PIR-International protein sequence database. *Nucleic Acids Res.* 22: 3569-3573.

- Ghisi, R., Anaclerio, F., Bassi, R., Passera, C. (1987) Light-induced formation of enzymes of the sulphate assimilation pathway in greening maize leaves. *Phytosynthetica* 21: 253-260.
- Gilmore, D.F., Godchaux, W., Leadbetter, E.R. (1989) Regulation of sulfate assimilation in *Cytophaga johnsonae*. *Arch. Microbiol.* 152: 387-392.
- Giovanelli, J. (1990) Regulatory aspects of cysteine and methionine biosynthesis. In: Sulfur nutrition and sulfur assimilation in higher plants, Rennenberg, H., Brunold, C., De Kok, L.J., Stulen, I. (eds), SPB Academic, The Hague, pp 33-48.
- Giovanelli, J., Mudd, S.H. (1967) Synthesis of homocysteine and cysteine by enzyme extracts of spinach. *Biochem. Biophys. Res. Comm.* 27: 150-156.
- Giovanelli, J., Mudd, S.H., Datko, A.H. (1980) Sulfur amino acids in plants. In: *The biochemistry of plants*, Vol. 5, Mifflin, B.J. (ed.), Academic Press, New York, pp 453-505.
- Gisselmann, G., Klausmeier, P., Schwenn, J.D. (1993) The ferredoxin: sulphite reductase gene from *Synechococcus* PCC7942. *Biochem. Biophys. Acta* 1144: 102-106.
- Gisselmann, G., Niehaus, A., Schwenn, J.D. (1992) Homologies in the structural genes coding for sulphate reducing enzymes from higher plants and prokaryotes. *Bot. Acta* 105: 213-218.
- Goldschmidt, E.E., Tsang, M. L.-S., Schiff, J.A. (1975) Studies of sulfate utilization by algae. 13. Adenosine 5'-phosphosulfate (APS) as an intermediate in the conversion of adenosine 3'-phosphate 5'-phosphosulfate (PAPS) to acid-volatile radioactivity. *Plant Sci. Lett.* 4: 293-299.
- Gutierrez-Marcos, J.F., Roberts, M.A., Campbell, E.I., Wray, J.L. (1996) Three members of a novel small gene-family from *Arabidopsis thaliana* able to complement functionally an *Escherichia coli* mutant defective in PAPS reductase activity encode proteins with a thioredoxin-like domain and "APS reductase" activity. *Proc. Natl Acad. Sci. USA* 93: 13377-13382.
- Haller, E., Suter, M., Brunold, C. (1986) Regulation of ATP-sulfurylase and adenosine 5'-phosphosulfate sulfotransferase by the sulfur and the nitrogen source in heterotrophic cell suspension cultures of

- Paul's Scarlet rose. *J. Plant Physiol.* 125: 275-283.
- Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166: 557-580.
- Harvey, R.J., Darlison, M.G. (1991) Random-primed cDNA synthesis facilitates the isolation of multiple 5'-cDNA ends by RACE. *Nucleic Acids Res.* 19: 4002.
- Hattori, M., Sakaki, Y. (1986) Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* 152: 232-238.
- Hawkins, H.C., de Nardi, M., Freedman, R.B. (1991) Redox properties and cross-linking of the dithiol/disulphide active sites of mammalian protein disulphide-isomerase. *Biochem. J.* 275: 341-348.
- Hediger, M.A., Johnson, D.F., Nierlich, D.P., Zabin, I. (1985) DNA sequence of the lactose operon: The *lacA* gene and the transcription termination region. *Proc. Natl Acad. Sci. USA* 82: 6414-6418.
- Hell, R., Bork, C., Bogdanova, N., Frolov, I., Hauschild, R. (1994) Isolation and characterization of two cDNAs encoding for compartment specific isoforms of *O*-acetylserine (thiol) lyase from *Arabidopsis thaliana*. *FEBS Lett.* 351: 257-262.
- Hell, R., Schuster, G., Gruissem, W. (1993) An *O*-acetylserine (thiol) lyase cDNA from spinach. *Plant Physiol.* 102: 1057-1058
- Henikoff, S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28: 351-359.
- Henikoff, S. (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods Enzymol.* 155: 156-165.
- Henikoff, S. (1993) Sequence analysis by electronic mail server. *Trends Biochem. Sci.* 18: 267-268.
- Hennies, H.H. (1975) A ferredoxin-linked sulfite reductase from *Spinacia oleracea*. *Z. Naturforsch.* 30c: 359-362.
- Herman, E.M. (1988) Immunocytochemical localization of macromolecules with the electron microscope. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 39: 139-155.
- Hesse, H., Altmann, T. (1995) Molecular cloning of a cysteine synthase

- cDNA from *Arabidopsis thaliana*. *Plant Physiol.* 108: 851-852.
- Hirvas, L., Koski, P., Vaara, M. (1991) Identification and sequence analysis of the gene mutated in the conditionally lethal outer membrane permeability mutant SS-C of *Salmonella typhimurium*. *EMBO J.* 10: 1017-1023.
- Hodson, R.C., Schiff, J.A. (1971) Studies of sulfate utilization by algae. 9. Fractionation of a cell-free system from *Chlorella* into two activities necessary for the reduction of adenosine 3'-phosphate 5'-phosphosulfate to acid-volatile radioactivity. *Plant Physiol.* 47: 300-305.
- Holmgren, A. (1985) Thioredoxin. *Ann. Rev. Biochem.* 54: 237-271.
- Holmgren, A. (1989) Thioredoxin and glutaredoxin systems. *J. Biol. Chem.* 264: 13963-13966.
- Howarth, J.R., Roberts, M.A., Wray, J.L. (1997) Cysteine biosynthesis in higher plants: A new member of the *Arabidopsis thaliana* serine acetyltransferase small gene-family obtained by functional complementation of an *Escherichia coli* cysteine auxotroph. *Biochim. Biophys. Acta* (In press).
- Hryniewicz, M., Sirko, A., Hulanicka, D. (1989) Identification and mapping of the sulphate permease promoter region in *Escherichia coli*. *Acta Biochim. Polon.* 36: 353-363.
- Hsu, L.-C., Chiou, T.-J., Chen, L., Bush, D.R. (1993) Cloning a plant amino acid transporter by functional complementation of a yeast amino acid transport mutant. *Proc. Natl Acad. Sci. USA* 90: 7441-7445.
- Hsu, M.P., Muhich, M.L., Boothroyd, J.C. (1989) A developmentally regulated gene of trypanosomes encodes a homologue of rat protein-disulfide isomerase and phosphoinositol-phospholipase C. *Biochemistry* 28: 6440-6446.
- Hughes, D.E., Wimpenny, J.W.T., Lloyd, D. (1971) The disintegration of micro-organisms. In: *Methods in microbiology*, Vol. 5B, Norris, J.R., Ribbons, D.W. (eds), Academic Press, London, pp 1-54.
- Hunt, A.G. (1994) Messenger RNA 3' end formation in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 45: 47-60.

- Hunt, C.L., Colless, V., Smith, M.T., Molasky, D.O., Malo, M.S., Loughlin, R.E. (1987) Lambda transducing phage and clones carrying genes of the *cysJIHDC* gene cluster of *Escherichia coli* K12. *J. Gen. Microbiol.* 133: 2707-2712.
- Ikegami, F., Horiuchi, S., Kobori, M., Morishige, I., Murakoshi, I. (1992) Biosynthesis of neuroactive amino acids by cysteine synthases in *Lathyrus latifolius*. *Phytochemistry* 31: 1991-1996.
- Ikegami, F., Itagaki, S., Murakoshi, I. (1993) Purification and characterization of two forms of cysteine synthase from *Allium tuberosum*. *Phytochemistry* 32: 31-34.
- Ikegami, F., Kaneko, M., Kobori, M., Murakoshi, I. (1988) Purification and characterization of cysteine synthase from *Brassica juncea*. *Phytochemistry* 27: 3379-3383.
- Ikegami, F., Mizuno, M., Kihara, M., Murakoshi, I. (1990) Enzymatic synthesis of the thyrotoxic amino acid mimosine by cysteine synthase. *Phytochemistry* 29: 3461-3465.
- Imhoff, J.F. (1982) Occurrence and evolutionary significance of two sulfate assimilation pathways in the Rhodospirillaceae. *Arch. Microbiol.* 132: 197-203.
- Innis, M.A., Myambo, K.B., Gelfand, D.H., Brow, M.A.D. (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl Acad. Sci. USA* 85: 9436-9440.
- Izui, K., Ishijima, S., Yamaguchi, Y., Katagiri, F., Murata, T., Shigesada, K., Sugiyama, T., Katsuki, H. (1986) Cloning and sequence analysis of cDNA encoding active phosphoenolpyruvate carboxylase of the C<sub>4</sub>-pathway from maize. *Nucleic Acids Res.* 14: 1615-1628.
- Jackson, R.H., Cornish-Bowden, A., Cole, J.A. (1981) Prosthetic groups of the NADH-dependent nitrite reductase from *Escherichia coli* K12. *Biochem. J.* 193: 861-867.
- Jacobsen, K.S., Breivold, E., Hornes, E. (1990) Purification of mRNA directly from crude plant tissues in 15 minutes using magnetic oligo dT microspheres. *Nucleic Acids Res.* 18: 3669.
- Jacobson, M.R., Brigle, K.E., Bennett, L.T., Setterquist, R.A., Wilson,

- M.S., Cash, V.L., Beynon, J., Newton, W.E., Dean, D.R. (1989) Physical and genetic map of the major *nif* gene cluster from *Azotobacter vinelandii*. J. Bacteriol. 171: 1017-1027.
- Jain, A., Leustek, T. (1994) A cDNA clone for 5'-adenylylphosphosulfate kinase from *Arabidopsis thaliana*. Plant Physiol. 105: 771-772.
- Jender, H.G., Schwenn, J.D. (1984) Purification and properties of the ATP:adenylylsulphate 3'-phosphotransferase from *Chlamydomonas reinhardtii*. Arch. Microbiol. 138: 9-14.
- Jenni, B.E., Brunold, C., Zrýd, J.-P., Lavanchy, P. (1980) Properties and regulation of adenosine 5'-phosphosulfate sulfotransferase from suspension cultures of *Nicotiana sylvestris*. Planta 150: 140-143.
- Jitrapakdee, S., Walker, M.E., Wallace, J.C. (1996) Identification of novel alternatively spliced pyruvate carboxylase mRNAs with divergent 5'-untranslated regions which are expressed in a tissue-specific manner. Biochem. Biophys. Res. Comm. 223: 695-700.
- Johnston, T.C., Thompson, R.B., Baldwin, T.O. (1986) Nucleotide sequence of the *luxB* gene of *Vibrio harveyi* and the complete amino acid sequence of the  $\beta$  subunit of bacterial luciferase. J. Biol. Chem. 261: 4805-4811.
- Jolivet, P. (1993) Elemental sulfur in agriculture. In: Sulfur nutrition and assimilation in higher plants, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Raiser, W.E. (eds), SPB Academic, The Hague, pp 193-206.
- Jones-Mortimer, M.C. (1968) Positive control of sulphate reduction in *Escherichia coli*. Biochem. J. 110: 589-595.
- Jones-Mortimer, M.C. (1973) Mapping the structural genes for the enzymes of cysteine biosynthesis in *Escherichia coli* K12 and *Salmonella typhimurium* LT2. Heredity 31: 213-221.
- Jones-Mortimer, M.C., Wheldrake, J.F., Pasternak, C.A. (1968) The control of sulphate reduction in *Escherichia coli* by *O*-acetyl-L-serine. Biochem. J. 107: 51-53.
- Joshi, C.P. (1987) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Res. 15: 6643-6653.

- Journet, E.P., Pichon, M., Dedieu, A., de Billy, F., Truchet, G., Barker, D.G. (1994) *Rhizobium meliloti* Nod factors elicit cell-specific transcription of the *ENOD12* gene in transgenic alfalfa. *Plant J.* 6: 241-249.
- Kaldenhoff, R., Kölling, A., Richter, G. (1993) A novel blue light- and abscisic acid-inducible gene of *Arabidopsis thaliana* encoding an intrinsic membrane protein. *Plant Mol. Biol.* 23: 1187-1198.
- Kanno, N., Nagahisa, E., Sato, M., Sato, Y. (1996) Adenosine 5'-phosphosulfate sulfotransferase from the marine macroalga *Porphyra yezoensis* Ueda (Rhodophyta): stabilization, purification and properties. *Planta* 198: 440-446.
- Kanno, N., Sato, M., Sato, Y. (1990) Purification and properties of adenosine-5'-phosphosulfate kinase from the marine red macroalga *Porphyra yezoensis* Ueda. *Bot. Marina* 33: 369-374.
- Karlin-Neumann, G.A., Tobin, E.M. (1986) Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. *EMBO J.* 5: 9-13.
- Kaufman, D.L., Evans, G.A. (1990) Restriction endonuclease cleavage at the termini of PCR products. *BioTechniques* 9: 304-306.
- Kaufman, J., Spicer, L.D., Siegel, L.M. (1993) Proton NMR of *Escherichia coli* sulfite reductase: The unligated heme protein subunit. *Biochemistry* 32: 2853-2867.
- Keegstra, K., Olsen, L.J., Theg, S.M. (1989) Chloroplastic precursors and their transport across the envelope membranes. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 40: 471-501.
- Ketter, J.S., Jarai, G., Fu, Y.-H., Marzluf, G.A. (1991) Nucleotide sequence, messenger RNA stability, and DNA recognition elements of *cys-14*, the structural gene for sulfate permease II in *Neurospora crassa*. *Biochemistry* 30: 1780-1787.
- Ketter, J.S., Marzluf, G.A. (1988) Molecular cloning and analysis of the regulation of *cys-14'*, a structural gene of the sulfur regulatory circuit of *Neurospora crassa*. *Mol. Cell. Biol.* 8: 1504-1508.
- Kleppinger-Sparace, K.F., Mudd, J.B., Sparace, S.A. (1990a) Biosynthesis of plant sulfolipids. In: Sulfur nutrition and sulfur assimilation

- in higher plants, Rennenberg, H., Brunold, C., De Kok, L.J., Stulen, I. (eds), SPB Academic, The Hague, pp 205-208.
- Kleppinger-Sparace, K.F., Sparace, S.A., Mudd, J.B. (1990b) Plant sulfolipids. In: Sulfur nutrition and sulfur assimilation in higher plants, Rennenberg, H., Brunold, C., De Kok, L.J., Stulen, I. (eds), SPB Academic, The Hague, pp 77-88.
- Klonus, D., Höfgen, R., Willmitzer, L., Riesmeier, J.W. (1994) Isolation and characterization of two cDNA clones encoding ATP-sulfurylase from potato by complementation of a yeast mutant. *Plant J.* 6: 105-112.
- Klonus, D., Riesmeier, J.W., Willmitzer, L. (1995) A cDNA clone for an ATP-sulfurylase from *Arabidopsis thaliana*. *Plant Physiol.* 107: 653-654.
- Koguchi, O., Takahashi, H., Tamura, G. (1988) Purification and partial characterization of ferredoxin-sulfite reductase from rape leaves. *Agric. Biol. Chem.* 52: 1867-1868.
- Korch, C., Mountain, H.A., Byström, A.S. (1991) Cloning, nucleotide sequence, and regulation of *MET14*, the gene encoding the APS kinase of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 229: 96-108.
- Korn, D., Weissbach, A. (1962) Thymineless induction in *Escherichia coli* K12( $\lambda$ ). *Biochim. Biophys. Acta* 61: 775-790.
- Kotler, L.E., Zevin-Sonkin, D., Sobolev, I.A., Beskin, A.D., Ulanovsky, L.E. (1993) DNA sequencing: Modular primers assembled from a library of hexamers or pentamers. *Proc. Natl Acad. Sci. USA* 90: 4241-4245.
- Kourilsky, P. (1973) Lysogenization by bacteriophage lambda: I. Multiple infection and the lysogenic response. *Mol. Gen. Genet.* 122: 183-195.
- Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 12: 857-872.
- Krause, G., Lundström, J., Barea, J.L., de la Cuesta, C.P., Holmgren, A. (1991) Mimicking the active site of protein disulfide-isomerase by substitution of proline 34 in *Escherichia coli* thioredoxin. *J. Biol. Chem.* 266: 9494-9500.

- Kredich, N.M. (1971) Regulation of L-cysteine biosynthesis in *Salmonella typhimurium*. J. Biol. Chem. 246: 3474-3484.
- Kredich, N.M. (1987) Biosynthesis of cysteine. In: *Escherichia coli* and *Salmonella typhimurium*. Cellular and molecular biology, Vol. 1, Neidhardt, F.C. (ed.), American Society for Microbiology, Washington DC, pp 419-428.
- Kredich, N.M. (1993) Gene regulation of sulfur assimilation. In: Sulfur nutrition and assimilation in higher plants, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Rauser, W.E. (eds), SPB Academic, The Hague, pp 37-47.
- Kredich, N.M., Becker, M.A., Tomkins, G.M. (1969) Purification and characterization of cysteine synthetase, a bifunctional protein complex, from *Salmonella typhimurium*. J. Biol. Chem. 244: 2428-2439.
- Kredich, N.M., Tomkins, G.M. (1966) The enzyme synthesis of L-cysteine in *Escherichia coli* and *Salmonella typhimurium*. J. Biol. Chem. 241: 4955-4965.
- Krone, F.A., Mundt, C., Schwenn, J.D. (1990a) Probing for PAPS-reductase in higher organisms by Southern blotting of genomic DNA with the cloned *cysH* gene from *Escherichia coli*. In: Sulfur nutrition and sulfur assimilation in higher plants, Rennenberg, H., Brunold, C., De Kok, L.J., Stulen, I. (eds), SPB Academic, The Hague, pp 209-212.
- Krone, F.A., Westphal, G., Meyer, H.E., Schwenn, J.D. (1990b) PAPS-reductase of *Escherichia coli*. FEBS Lett. 260: 6-9.
- Krone, F.A., Westphal, G., Schwenn, J.D. (1991) Characterisation of the gene *cysH* and of its product phospho-adenylylsulphate reductase from *Escherichia coli*. Mol. Gen. Genet. 225: 314-319.
- Kuske, C.R., Hill, K.K., Guzman, E., Jackson, P.J. (1996) Subcellular location of O-acetylserine sulfhydrylase isoenzymes in cell cultures and plant tissues of *Datura innoxia* Mill. Plant Physiol. 112: 659-667.
- Laemmli, U.K. (1970) Cleavage of the structural protein during the assembly of the head of bacteriophage T4. Nature 227: 680-685.
- Lai, C.-Y., Baumann, P. (1992) Sequence analysis of a DNA fragment from

- Buchnera aphidicola* (an endosymbiont of aphids) containing genes homologous to *dnaG*, *rpoD*, *cysE*, and *secB*. *Gene* 119: 113-118.
- Lal, S.K., Johnson, S., Conway, T., Kelley, P.M. (1991) Characterization of a maize cDNA that complements an enolase-deficient mutant of *Escherichia coli*. *Plant Mol. Biol.* 16: 787-795.
- Lech, K., Brent, R. (1988) Plating lambda phage to generate plaques. In: *Current protocols in molecular biology*, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (eds), John Wiley & Sons, USA, pp 1.11.1-1.11.4.
- Leon, J., Vega, J.M. (1991) Separation and regulatory properties of O-acetyl-L-serine sulfhydrylase isoenzymes from *Chlamydomonas reinhardtii*. *Plant Physiol. Biochem.* 29: 595-599.
- Leu, L.-S., Cook, P.F. (1994) Kinetic mechanism of serine transacetylase from *Salmonella typhimurium*. *Biochemistry* 33: 2667-2671.
- Leustek, T. (1996) Molecular genetics of sulfate assimilation in plants. *Physiol. Plant.* 97: 411-419.
- Leustek, T., Murillo, M., Cervantes, M. (1994) Cloning of a cDNA encoding ATP sulfurylase from *Arabidopsis thaliana* by functional complementation in *Saccharomyces cerevisiae*. *Plant Physiol.* 105: 897-902.
- Levinthal, M., Schiff, J.A. (1968) Studies on sulfate utilization by algae. 5. Identification of thiosulfate as a major acid-volatile product formed by a cell-free sulfate-reducing system from *Chlorella*. *Plant Physiol.* 43: 555-562.
- Leyh, T.S., Taylor, J.C., Markham, G.D. (1988) The sulfate activation locus of *Escherichia coli* K12: Cloning, genetic and enzymatic characterization. *J. Biol. Chem.* 263: 2409-2416.
- Leyh, T.S., Vogt, T.F., Suo, Y. (1992) The DNA sequence of the sulfate activation locus from *Escherichia coli* K-12. *J. Biol. Chem.* 267: 10405-10410.
- Li, C., Peck, H.D., Przybyla, A.E. (1987) Cloning of the 3'-phosphoadenylyl sulfate reductase and sulfite reductase genes from *Escherichia coli* K-12. *Gene* 53: 227-234.
- Li, J., Saidha, T., Schiff, J.A. (1991) Purification and properties of

- two forms of ATP sulfurylase from *Euglena*. *Biochim. Biophys. Acta* 1078: 68-76.
- Li, J., Schiff, J.A. (1991) Purification and properties of adenosine 5'-phosphosulphate sulphotransferase from *Euglena*. *Biochem. J.* 274: 355-360.
- Li, J., Schiff, J.A. (1992) Adenosine 5'-phosphosulfate sulfotransferase from *Euglena*: Enzyme-bound intermediates. *Plant Cell Physiol.* 33: 63-72.
- Li, X., Nield, J., Hayman, D., Langridge, P. (1995) Thioredoxin activity in the C terminus of *Phalaris S* protein. *Plant J.* 8: 133-138.
- Lin, W.S., Cunneen, T., Lee, C.Y. (1994) Sequence analysis and molecular characterization of genes required for the biosynthesis of type 1 capsular polysaccharide in *Staphylococcus aureus*. *J. Bacteriol.* 176: 7005-7106.
- Logan, H.M., Cathala, N., Grignon, C., Davidian, J.-C. (1996) Cloning of a cDNA encoded by a member of the *Arabidopsis thaliana* ATP sulfurylase multigene family. *J. Biol. Chem.* 271: 12227-12233.
- Logemann, J., Schell, J., Willmitzer, L. (1987) Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* 163: 16-20.
- Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L., Davis, M.M. (1989) Polymerase chain reaction with single-sided specificity: Analysis of T cell receptor  $\delta$  chain. *Science* 243: 217-220.
- Lowe, G. (1991) Mechanisms of sulphate activation and transfer. *Phil. Trans. R. Soc. Lond. B* 332: 141-148.
- Lundström, J., Krause, G., Holmgren, A. (1992) A Pro to His mutation in active site of thioredoxin increase its disulfide-isomerase activity 10-fold. *J. Biol. Chem.* 267: 9047-9052.
- Lunn, J.E., Droux, M., Martin, J., Douce, R. (1990) Localization of ATP sulfurylase and *O*-acetylserine(thiol)lyase in spinach leaves. *Plant Physiol.* 94: 1345-1352.
- Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., Scheele, G.A. (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J.* 6: 43-48.

- Lyle, S., Geller, D.H., Ng, K., Westley, J., Schwartz, N.B. (1994) Kinetic mechanism of ATP-sulphurylase from rat chondrosarcoma. *Biochem. J.* 301: 349-354.
- Lyles, M.M., Gilbert, H.F. (1994) Mutation in the thioredoxin sites of protein disulfide isomerase reveal functional nonequivalence of the N- and C-terminal domains. *J. Biol. Chem.* 269: 30946-30952.
- Magasanik, B. (1970) Glucose effects: inducer exclusion and repression. In: *The lactose operon*, Beckwith, J.R., Zipser, D. (eds), Cold Spring Harbor Laboratory, Cold Spring Harbor, pp 189-219.
- Mandelstam, J. (1962) The repression of constitutive  $\beta$ -galactosidase in *Escherichia coli* by glucose and other carbon sources. *Biochem. J.* 82: 489-493.
- Marzluf, G.A. (1970a) Genetic and biochemical studies of distinct sulfate permease species in different developmental stages of *Neurospora crassa*. *Arch. Biochem. Biophys.* 138: 254-263.
- Marzluf, G.A. (1970b) Genetic and metabolic controls for sulfate metabolism in *Neurospora crassa*: Isolation and study of chromate-resistant and sulfate transport-negative mutants. *J. Bacteriol.* 102: 716-721.
- Masada, M., Fukushima, K., Tamura, G. (1975) Cysteine synthase from rape leaves. *J. Biochem.* 77: 1107-1115.
- Masuta, C., Furuno, M., Tanaka, H., Yamada, M., Koiwai, A. (1992) Molecular cloning of a cDNA clone for the tobacco lipid transfer protein and expression of the functional protein in *Escherichia coli*. *FEBS Lett.* 311: 119-123.
- Mazzarella, R.A., Srinivasan, M., Haugejorden, S.M., Green, M. (1990) ERp72, an abundant luminal endoplasmic reticulum protein, contains three copies of the active site sequences of protein disulfide isomerase. *J. Biol. Chem.* 265: 1094-1101.
- Mercer, E.I., Thomas, G. (1969) The occurrence of ATP-adenylsulphate 3'-phosphotransferase in the chloroplasts of higher plants. *Phytochemistry* 8: 2281-2285.
- Metzler, D.E. (1977) *Biochemistry*, Academic Press, New York.
- Mifflin, B.J., Lea, P.J. (1977) Amino acid metabolism. *Ann. Rev. Plant*

Physiol. 28: 299-329.

- Millard, P. Sharp, G.S., Scott, N.M. (1985) The effect of sulphur deficiency on the uptake and incorporation of nitrogen in ryegrass. J. Agric. Sci. 105: 501-504.
- Miller, J.H. (1972) Experiments in molecular genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, pp 352-355.
- Mills, D.R., Kramer, F.R. (1979) Structure-independent nucleotide sequence analysis. Proc. Natl Acad. Sci. USA 76: 2232-2235.
- Minakuchi, K., Yabushita, T., Masumura, T., Ichihara, K., Tanaka, K. (1994) Cloning and sequence analysis of a cDNA encoding rice glutaredoxin. FEBS Lett. 337: 157-160.
- Minet, M., Dufour, M.-E., Lacroute, F. (1992) Complementation of *Saccharomyces cerevisiae* auxotrophic mutants by *Arabidopsis thaliana* cDNAs. Plant J. 2: 417-422.
- Monod, M., Mohan, S., Dubnau, D. (1987) Cloning and analysis of *ermG*, a new macrolide-lincosamide-streptogramin B resistance element from *Bacillus sphaericus*. J. Bacteriol. 169: 340-350.
- Murakoshi, I., Ikegami, F., Kaneko, M. (1985) Purification and properties of cysteine synthase from *Spinacia oleracea*. Phytochemistry 24: 1907-1911.
- Murakoshi, I., Kaneko, M., Koide, C., Ikegami, F. (1986) Enzymatic synthesis of the neuroexcitatory amino acid quisqualic acid by cysteine synthase. Phytochemistry 25: 2759-2763.
- Murillo, M., Leustek, T. (1995) Adenosine-5'-triphosphate-sulfurylase from *Arabidopsis thaliana* and *Escherichia coli* are functionally equivalent but structurally and kinetically divergent: Nucleotide sequence of two adenosine-5'-triphosphate-sulfurylase cDNAs from *Arabidopsis thaliana* and analysis of a recombinant enzyme. Arch. Biochem. Biophys. 323: 195-204.
- Murray, N.E., Brammar, W.J., Murray, K. (1977) Lambdoid phages that simplify the recovery of in vitro recombinants. Mol. Gen. Genet. 150: 53-61.
- Musich, P.R., Chu, W. (1993) A hot alkaline plasmid DNA miniprep method for automated DNA sequencing protocols. BioTechniques 14: 958-960.

- Nakai, K., Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14: 897-911.
- Nakamura, K., Hayama, A., Masada, M., Fukushima, K., Tamura, G. (1987) Measurement of serine acetyltransferase activity in crude plant extracts by a coupled assay system using cysteine synthase. *Plant Cell Physiol.* 28: 885-891.
- Nakamura, K., Hayama, A., Masada, M., Fukushima, K., Tamura, G. (1988) Purification and some properties of plant serine acetyltransferase. *Plant Cell Physiol.* 29: 689-693.
- Nakamura, K., Tamura, G. (1989) Five isoforms of cysteine synthase in rape leaves. *Agric. Biol. Chem.* 53: 2537-2538.
- Nakamura, K., Tamura, G. (1990) Isolation of serine acetyltransferase complexed with cysteine synthase from *Allium tuberosum*. *Agric. Biol. Chem.* 54: 649-656.
- Needleman, S.B., Wunsch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48: 443-453.
- Neuenschwander, U., Suter, M., Brunold, C. (1991) Regulation of sulfate assimilation by light and O-acetyl-L-serine in *Lemna minor* L. *Plant Physiol.* 97: 253-258.
- Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E., Somerville, C. (1994) Genes galore: A summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiol.* 106: 1241-1255.
- Ng, B.H., Anderson, J.W. (1978) Chloroplast cysteine synthases of *Trifolium repens* and *Pisum sativum*. *Phytochemistry* 17: 879-885.
- Ngo, T.T., Shargool, P.D. (1973) The use of a sulfide ion selective electrode to study O-acetylserine sulfhydrylase from germinating rapeseed. *Anal. Biochem.* 54: 247-261.
- Ngo, T.T., Shargool, P.D. (1974) The enzymatic synthesis of L-cysteine in higher plant tissues. *Can. J. Biochem.* 52: 435-440.
- Nicholson, M.L., Gaasenbeek, M., Laudénbach, D.E. (1995) Two enzymes

- together capable of cysteine biosynthesis are encoded on a cyanobacterial plasmid. *Mol. Gen. Genet.* 247: 623-632.
- Niehaus, A., Gisselmann, G., Schwenn, J.D. (1992) Primary structure of the *Synechococcus* PCC 7942 PAPS reductase gene. *Plant Mol. Biol.* 20: 1179-1183.
- Nilsson, D., Hove-Jensen, B., Arnvig, K. (1989) Primary structure of the *tms* and *prs* genes of *Bacillus subtilis*. *Mol. Gen. Genet.* 218: 565-571.
- Nishikawa, K. (1983) Assessment of secondary-structure prediction of proteins. Comparison of computerized Chou-Fasman method with others. *Biochim. Biophys. Acta* 748: 285-299.
- Noji, M., Murakoshi, I., Saito, K. (1994) Molecular cloning of a cysteine synthase cDNA from *Citrullus vulgaris* (watermelon) by genetic complementation in an *Escherichia coli* Cys<sup>-</sup> auxotroph. *Mol. Gen. Genet.* 244: 57-66.
- Ohara, O., Dorit, R.L., Gilbert, W. (1989) One-sided polymerase chain reaction: The amplification of cDNA. *Proc. Natl Acad. Sci. USA* 86: 5673-5677.
- Ostrowski, J., Barber, M.J., Rueger, D.C., Miller, B.E., Siegel, L.M., Kredich, N.M. (1989a) Characterization of the flavoprotein moieties of NADPH-sulfite reductase from *Salmonella typhimurium* and *Escherichia coli*. *J. Biol. Chem.* 264: 15796-15808.
- Ostrowski, J., Wu, J.-Y., Rueger, D.C., Miller, B.E., Siegel, L.M., Kredich, N.M. (1989b) Characterization of the *cysJIH* regions of *Salmonella typhimurium* and *Escherichia coli* B. *J. Biol. Chem.* 264: 15726-15737.
- Otsuji, N., Sekiguchi, M., Iijima, T., Takagi, Y. (1959) Induction of phage formation in the lysogenic *Escherichia coli* K-12 by mitomycin C. *Nature* 184: 1079-1080.
- Overbeeke, N., Verrips, C.T. (1993) Expression of plant genes in yeast and bacteria. In: *Plant protoplast and genetic engineering*, Vol. 4, Bajaj, Y.P.S. (ed.), Springer, Berlin, pp 183-198.
- Pang, Q., Hays, J.B., Rajagopal, I. (1992) A plant cDNA that partially complements *Escherichia coli* *recA* mutations predicts a polypeptide not strongly homologous to RecA proteins. *Proc. Natl Acad. Sci. USA*

89: 8073-8077.

- Parent, R., Roy, P.H. (1992) The chloramphenicol acetyltransferase gene of Tn2424: a new breed of cat. *J. Bacteriol.* 174: 2891-2897.
- Pasternak, C.A., Ellis, R.J., Jones-Mortimer, M.C., Crichton, C.E. (1965) The control of sulphate reduction in bacteria. *Biochem. J.* 96: 270-275.
- Pasternak, J.J. (1993) DNA sequence organization and gene copy number determination. In: *Methods in plant molecular biology and biotechnology*, Glick, B.R., Thompson, J.E. (eds), CRC Press, Boca Raton, pp 29-36.
- Pearson, W.R., Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl Acad. Sci. USA* 85: 2444-2448.
- Peck, H.D., Lissolo, T. (1988) Assimilatory and dissimilatory sulphate reduction: Enzymology and bioenergetics. In: *The nitrogen and sulphur cycles*, Cole, J.A., Ferguson, S.J. (eds), Cambridge University Press, Cambridge, pp 99-132.
- Peng, Z., Verma, D.P.S. (1995) A rice *HAL2*-like gene encodes a Ca<sup>2+</sup>-sensitive 3' (2'), 5'-diphosphonucleoside 3' (2')-phosphohydrolase and complements yeast *met22* and *Escherichia coli cysQ* mutations. *J. Biol. Chem.* 270: 29105-29110.
- Quirk, P.G., Guffanti, A.A., Clejan, S., Cheng, J., Krulwich, T.A. (1994) Isolation of Tn917 insertional mutants of *Bacillus subtilis* that are resistant to the protonophore carbonyl cyanide *m*-chlorophenylhydrazone. *Biochim. Biophys. Acta* 1186: 27-34.
- Raetz, C.R.H., Roderick, S.L. (1995) A left-handed parallel  $\beta$  helix in the structure of UDP-*N*-acetylglucosamine acyltransferase. *Science* 270: 997-1000.
- Rao, V.B. (1994) Direct sequencing of polymerase chain reaction-amplified DNA. *Anal. Biochem.* 216: 1-14.
- Reich, E., Shatkin, A.J., Tatum, E.L. (1961) Bacteriocidal action of mitomycin C. *Biochim. Biophys. Acta* 53: 132-149.
- Rennenberg, H. (1983) Role of *O*-acetylserine in hydrogen sulfide emission from pumpkin leaves in response to sulfate. *Plant Physiol.* 73: 560-565.

- Renosto, F., Martin, R.L., Segel, I.H. (1989) Sulfate-activating enzymes of *Penicillium chrysogenum*. J. Biol. Chem. 264: 9433-9437.
- Renosto, F., Seubert, P.A., Knudson, P., Segel, I.H. (1985) APS kinase from *Penicillium chrysogenum*. J. Biol. Chem. 260: 1535-1544.
- Renosto, F., Seubert, P.A., Segel, I.H. (1984) Adenosine 5'-phosphosulfate kinase from *Penicillium chrysogenum*. J. Biol. Chem. 259: 2113-2123.
- Reuveny, Z., Dougall, D.K., Trinity, P.M. (1980) Regulatory coupling of nitrate and sulfate assimilation pathways in cultured tobacco cells. Proc. Natl Acad. Sci. USA 77: 6670-6672.
- Reuveny, Z., Filner, P. (1977) Regulation of adenosine triphosphate sulfurylase in cultured tobacco cells. J. Biol. Chem. 252: 1858-1864.
- Riggs, P. (1994) Expression and purification of maltose-binding protein fusions. In: Current protocols in molecular biology, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (eds), John Wiley & Sons, USA, pp 16.6.1-16.6.14.
- Robbins, P.W., Lipmann, F. (1956) The enzymatic sequence in the biosynthesis of active sulfate. J. Am. Chem. Soc. 78: 6409-6410.
- Roberts, M.A., Wray, J.L. (1996) Cloning and characterisation of an *Arabidopsis thaliana* cDNA clone encoding an organellar isoform of serine acetyltransferase. Plant Mol. Biol. 30: 1041-1049.
- Rolland, N., Droux, M., Douce, R. (1992) Subcellular distribution of O-acetylserine(thiol)lyase in cauliflower (*Brassica oleracea* L.) inflorescence. Plant Physiol. 98: 927-935.
- Rolland, N., Droux, M., Lebrun, M., Douce, R. (1993a) O-acetylserine(thiol)lyase from spinach (*Spinacia oleracea* L.) leaf: cDNA cloning, characterization, and overexpression in *Escherichia coli* of the chloroplast isoform. Arch. Biochem. Biophys. 300: 213-222.
- Rolland, N., Job, D., Douce, R. (1993b) Common sequence motifs coding for higher-plant and prokaryotic O-acetylserine (thiol)-lyases: bacterial origin of a chloroplastic transit peptide? Biochem J. 293: 829-833.

- Römer, S., d'Harlingue, A., Camara, B., Schantz, R., Kuntz, M. (1992) Cysteine synthase from *Capsicum annuum* chloroplasts. *J. Biol. Chem.* 267: 17966-17970.
- Rosenthal, E., Leustek, T. (1995) A multifunctional *Urechis caupo* protein, PAPS synthetase, has both ATP sulfurylase and APS kinase activities. *Gene* 165: 243-248.
- Ruffet, M.-L., Droux, M., Douce, R. (1994) Purification and kinetic properties of serine acetyltransferase free of O-acetylserine(thiol)lyase from spinach chloroplasts. *Plant Physiol.* 104: 597-604.
- Ruffet, M.-L., Lebrun, M., Droux, M., Douce, R. (1995) Subcellular distribution of serine acetyltransferase from *Pisum sativum* and characterization of an *Arabidopsis thaliana* putative cytosolic isoform. *Eur. J. Biochem.* 227: 500-509.
- Sabelli, P.A., Shewry, P.R. (1993) Nucleic acid blotting and hybridisation. In: *Methods in plant biochemistry: Molecular biology*, Vol. 10, Bryant, J., Dey, P.M., Harborne, J.B. (eds), Academic Press, New York, pp 79-100.
- Saccomani, M., Cacco, G., Ferrari, G. (1984) Effect of nitrogen and/or sulfur deprivation on the uptake and assimilation steps of nitrate and sulfate in maize seedlings. *J. Plant Nutrition* 7: 1043-1057.
- Saidha, T., Na, S.-Q., Li, J., Schiff, J.A. (1988) A sulphate metabolizing centre in *Euglena* mitochondria. *Biochem. J.* 253: 533-539.
- Saidha, T., Schiff, J.A. (1989) The role of mitochondria in sulfolipid biosynthesis by *Euglena* chloroplasts. *Biochim. Biophys. Acta* 1001: 268-273.
- Saito, K., Kurosawa, M., Murakoshi, I. (1993a) Determination of a functional lysine residue of a plant cysteine synthase by site-directed mutagenesis, and the molecular evolutionary implications. *FEBS Lett.* 328: 111-114.
- Saito, K., Kurosawa, M., Tatsuguchi, K., Takagi, Y., Murakoshi, I. (1994a) Modulation of cysteine biosynthesis in chloroplasts of transgenic tobacco overexpressing cysteine synthase [O-acetylserine(thiol)-lyase]. *Plant Physiol.* 106: 887-895.

- Saito, K., Miura, N., Yamazaki, M., Hirano, H., Murakoshi, I. (1992) Molecular cloning and bacterial expression of cDNA encoding a plant cysteine synthase. *Proc. Natl Acad. Sci. USA* 89: 8078-8082.
- Saito, K., Tatsuguchi, K., Murakoshi, I., Hirano, H. (1993b) cDNA cloning and expression of cysteine synthase B localized in chloroplasts of *Spinacia oleracea*. *FEBS Lett.* 324: 247-252.
- Saito, K., Tatsuguchi, K., Takagi, Y., Murakoshi, I. (1994b) Isolation and characterization of cDNA that encodes a putative mitochondrion-localizing isoform of cysteine synthase (O-acetylserine(thiol)-lyase) from *Spinacia oleracea*. *J. Biol. Chem.* 269: 28187-28192.
- Saito, K., Yokoyama, H., Noji, M., Murakoshi, I. (1995) Molecular cloning and characterization of a plant serine acetyltransferase playing a regulatory role in cysteine biosynthesis from watermelon. *J. Biol. Chem.* 270: 16321-16326.
- Sakamoto, T., Murata, N. (1995) Sequence announcement: *cysE* and *pgi*. *Plant. Mol. Biol.* 29: 187.
- Salisbury, F.B., Ross, C.W. (1992) *Plant physiology*, Wadsworth, Belmont, pp 305-307.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) *Molecular cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor.
- Sanger, F., Nicklen, S., Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA* 94: 5463-5467.
- Satishchandran, C., Hickman, Y.N., Markham, G.D. (1992) Characterization of the phosphorylated enzyme intermediate formed in the adenosine 5'-phosphosulfate kinase reaction. *Biochemistry* 31: 11684-11688.
- Satishchandran, C., Markham, G.D. (1989) Adenosine-5'-phosphosulfate kinase from *Escherichia coli* K12. *J. Biol. Chem.* 264: 15012-15021.
- Sauer, B., Henderson, N. (1998) The cyclization of linear DNA in *Escherichia coli* by site-specific recombination. *Gene* 70: 331-341.
- Sawhney, S.K., Nicholas, D.J.D. (1975) Nitrite, hydroxylamine and sulphite reductase in wheat leaves. *Phytochemistry* 14: 1499-1503.
- Schiff, J.A., Hodson, R.C. (1973) The metabolism of sulfate. *Ann. Rev. Plant Physiol.* 24: 381-414.

- Schiff, J.A., Levinthal, M. (1968) Studies of sulfate utilization by algae. 4. Properties of a cell-free sulfate-reducing system from *Chlorella*. *Plant Physiol.* 43: 547-554.
- Schiff, J.A., Saidha, T. (1987) Overview: Inorganic sulfur and sulfate activation. *Methods Enzymol.* 143: 329-334.
- Schiff, J.A., Stern, A.I., Saidha, T., Li, J. (1993) Some molecular aspects of sulfate metabolism in photosynthetic organisms. In: *Sulfur nutrition and assimilation in higher plants*, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Rauser, W.E. (eds), SPB Academic, The Hague, pp 21-35.
- Schiffmann, S., Schwenn, J.D. (1994) APS-sulfotransferase activity is identical to higher plant APS-kinase. *FEBS Lett.* 355: 229-232.
- Schmidt, A. (1972) On the mechanism of photosynthetic sulfate reduction. An APS-sulfotransferase from *Chlorella*. *Arch. Microbiol.* 84: 77-86.
- Schmidt, A. (1973) Sulfate reduction in a cell-free system of *Chlorella*. *Arch. Microbiol.* 93: 29-52.
- Schmidt, A. (1975) Distribution of APS-sulfotransferase activity among higher plants. *Plant Sci. Lett.* 5: 407-415.
- Schmidt, A. (1976) The adenosine-5'-phosphosulfate sulfotransferase from spinach (*Spinacea oleracea* L.). Stabilization, partial purification and properties. *Planta* 130: 257-263.
- Schmidt, A. (1986) Regulation of sulfur metabolism in plants. *Progr. Bot.* 48: 133-150.
- Schmidt, A., Abrams, W.R., Schiff, J.A. (1974) Reduction of adenosine 5'-phosphosulfate to cysteine in extracts from *Chlorella* and mutants blocked for sulfate reduction. *Eur. J. Biochem.* 47: 423-434.
- Schmidt, A., Jäger, K. (1992) Open questions about sulfur metabolism in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 43: 325-349.
- Schmidt, G.W., Mishkind, M.L. (1986) The transport of proteins into chloroplasts. *Ann. Rev. Biochem.* 55: 879-912.
- Schmutz, D., Brunold, C. (1982) Regulation of sulfate assimilation in plants. XIII. Assimilatory sulfate reduction during ontogenesis of primary leaves of *Phaseolus vulgaris* L. *Plant Physiol.* 70: 524-527.

- Schmutz, D., Brunold, C. (1985) Localization of nitrite and sulfite reductase in bundle sheath and mesophyll cells of maize leaves. *Physiol. Plant.* 64: 523-528.
- Schnorr, K.M., Nygaard, P., Laloue, M. (1994) Molecular characterization of *Arabidopsis thaliana* cDNAs encoding three purine biosynthetic enzymes. *Plant J.* 6: 113-121.
- Schnug, E. (1993) Physiological and environmental relevance of sulfur-containing secondary metabolism. In: Sulfur nutrition and assimilation in higher plants, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Rauser, W.E. (eds), SPB Academic, The Hague, pp 179-190.
- Schriek, U., Schwenn, J.D. (1986) Properties of the purified APS-kinase from *Escherichia coli* and *Saccharomyces cerevisiae*. *Arch. Microbiol.* 145: 32-38.
- Schürmann, P. (1993) Plant thioredoxins. In: Sulfur nutrition and assimilation in higher plants, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Rauser, W.E. (eds), SPB Academic, The Hague, pp 153-162.
- Schürmann, P., Brunold, C. (1980) Formation of cysteine from adenosine 5'-phosphosulphate (APS) in extracts from spinach chloroplasts. *Z. Pflanzenphysiol.* 100: 257-268.
- Schwedock, J., Long, S.R. (1989) Nucleotide sequence and protein products of two new nodulation genes of *Rhizobium meliloti*, *nodP* and *nodQ*. *Mol. Plant-Microbe. Interact.* 2: 181-194.
- Schwenn, J.D. (1989) Sulphate assimilation in higher plants: A thioredoxin-dependent PAPS-reductase from spinach leaves. *Z. Naturforsch.* 44c: 504-508.
- Schwenn, J.D. (1994) Photosynthetic sulphate reduction. *Z. Naturforsch.* 49c: 531-539.
- Schwenn, J.D., Jender, H.G. (1981) A kinetic investigation of the APS-kinase from *Chlamydomonas reinhardtii* CW15. *Phytochemistry* 20: 601-604.
- Schwenn, J.D., Kemena, A. (1984) Expression of the plant sulphite reductase in cell suspension cultures from *Catharanthus roseus* L. *Planta* 160: 151-158.

- Schwenn, J.D., Krone, F.A., Husmann, K. (1988) Yeast PAPS reductase: properties and requirements of the purified enzyme. *Arch. Microbiol.* 150: 313-319.
- Schwenn, J.D., Schriek, U. (1987) PAPS-reductase from *Escherichia coli*: Characterization of the enzyme as probe for thioredoxins. *Z. Naturforsch.* 42c: 93-102.
- Segel, I.H., Renosto, F., Seubert, P.A. (1987) Sulfate-activating enzymes. *Methods Enzymol.* 143: 334-349.
- Seidman, C.E., Struhl, K., Sheen, J. (1989) Introduction of plasmid DNA into cells. In: *Current protocols in molecular biology*, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (eds), John Wiley & Sons, USA, pp 1.8.1-1.8.8.
- Senecoff, J.F., Meagher, R.B. (1993) Isolating the *Arabidopsis thaliana* genes for de novo purine synthesis by suppression of *Escherichia coli* mutants. I. 5'-Phosphoribosyl-5'-aminoimidazole synthetase. *Plant Physiol.* 102: 387-399.
- Setya, A., Murillo, M., Leustek, T. (1996) Sulfate reduction in higher plants: Molecular evidence for a novel 5'-adenylylsulfate reductase. *Proc. Natl Acad. Sci. USA* 93: 13383-13388.
- Shorrosh, B.S., Dixon, R.A. (1992) Molecular characterization and expression of an alfalfa protein with sequence similarity to mammalian ERp72, a glucose-regulated endoplasmic reticulum protein containing active site sequences of protein disulphide isomerase. *Plant J.* 2: 51-58.
- Siegel, L.M., Davis, P.S. (1974) Reduced nicotinamide adenine dinucleotide phosphate-sulfite reductase of enterobacteria. IV. The *Escherichia coli* hemoflavoprotein: Subunit structure and the dissociation into hemoprotein and flavoprotein components. *J. Biol. Chem.* 249: 1587-1598.
- Siegel, L.M., Murphy, M.J., Kamin, H. (1973) Reduced nicotinamide adenine dinucleotide phosphate-sulfite reductase of enterobacteria. I. The *Escherichia coli* hemoflavoprotein: Molecular parameters and prosthetic groups. *J. Biol. Chem.* 248: 251-264.
- Slatko, B.E., Eckert, R.L., Albright, L.M., Ausubel, F.M. (1993) DNA sequencing strategies. In: *Current protocols in molecular biology*, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman,

- J.G., Smith, J.A., Struhl, K. (eds), John Wiley & Sons, USA, pp 7.1.1-7.1.7.
- Smith, D.B. (1993) Purification of glutathione S-transferase fusion proteins. *Methods Mol. Cell Biol.* 4: 220-229.
- Smith, D.B., Corcoran, L.M. (1994) Expression and purification of glutathione-S-transferase fusion proteins. In: *Current protocols in molecular biology*, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (eds), John Wiley & Sons, USA, pp 16.7.1-16.7.7.
- Smith, D.B., Johnson, K.S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67: 31-40.
- Smith, F.W., Ealing, P.M., Hawkesford, M.J., Clarkson, D.T. (1995) Plant members of a family of sulfate transporters reveal functional subtypes. *Proc. Natl Acad. Sci. USA* 92: 9373-9377.
- Smith, I.K. (1972) Studies of L-cysteine biosynthetic enzymes in *Phaseolus vulgaris* L. *Plant Physiol.* 50: 477-479.
- Smith, I.K. (1975) Sulfate transport in cultured tobacco cells. *Plant Physiol.* 55: 303-307.
- Smith, I.K. (1980) Regulation of sulfate assimilation in tobacco cells. *Plant Physiol.* 66: 877-883.
- Smith, I.K., Thompson, J.F. (1971) Purification and characterization of L-serine transacetylase and O-acetyl-L-serine sulfhydrylase from kidney bean seedlings (*Phaseolus vulgaris*). *Biochim. Biophys. Acta* 227: 288-295.
- Smith, J.K., Schloss, J.V., Mazur, B.J. (1989) Functional expression of plant acetolactate synthase genes in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 86: 4179-4183.
- Snustad, D.P., Hunsperger, J.P., Chereskin, B.M., Messing, J. (1988) Maize glutamine synthetase cDNAs: Isolation by direct genetic selection in *Escherichia coli*. *Genetics* 120: 1111-1124.
- Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.

- Stallmeyer, B., Nerlich, A., Schiemann, J., Brinkmann, H., Mendel, R.R. (1995) Molybdenum co-factor biosynthesis: the *Arabidopsis thaliana* cDNA *cnx1* encodes a multifunctional two-domain protein homologous to a mammalian neuroprotein, the insect protein Cinnamon and three *Escherichia coli* proteins. *Plant J.* 8: 751-762.
- Stanley, P.E., Kelley, B.C., Tuovinen, O.H., Nicholas, D.J.D. (1975) A bioluminescence method for determining adenosine 3'-phosphate 5'-phosphate (PAP) and adenosine 3'-phosphate 5'-sulfatophosphate (PAPS) in biological materials. *Anal. Biochem.* 67: 540-551.
- Stroehrer, U.H., Karageorgos, L.E., Morona, R., Manning, P.A. (1992) Serotype conversion in *Vibrio cholerae* O1. *Proc. Natl Acad. Sci. USA* 89: 2566-2570.
- Struhl, K., Cameron, J.R., Davis, R.W. (1976) Functional genetic expression of eukaryotic DNA in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 73: 1471-1475.
- Stulen, I., De Kok, L.J. (1993) Whole plant regulation of sulfur metabolism: A theoretical approach and comparison with current ideas on regulation of nitrogen metabolism. In: Sulfur nutrition and assimilation in higher plants, De Kok, L.J., Stulen, I., Rennenberg, H., Brunold, C., Rauser, W.E. (eds), SPB Academic, The Hague, pp 77-91.
- Surin, B.P., Downie, J.A. (1988) Characterization of the *Rhizobium leguminosarum* genes *nodLMN* involved in efficient host-specific nodulation. *Mol. Microbiol.* 2: 173-183.
- Suter, M., Lavanchy, P., von Arb, C., Brunold, C. (1986) Regulation of sulfate assimilation by amino acids in *Lemna minor* L. *Plant Sci.* 44: 125-132.
- Suter, M., Tschanz, A., Brunold, C. (1992) Adenosine 5'-phosphosulfate sulfotransferase from Norway spruce: Biochemical and physiological properties. *Bot. Acta* 105: 190-196.
- Tabor, S., Richardson, C.C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl Acad. Sci. USA* 84: 4767-4771.
- Tabor, S., Richardson, C.C. (1989) Effect of manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. *Proc. Natl Acad.*

Sci. USA 86: 4076-4080.

- Tabor, S., Richardson, C.C. (1990) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *J. Biol. Chem.* 265: 8322-8328.
- Takahashi, H., Sasakura, N., Noji, M., Saito, K. (1996) Isolation and characterization of a cDNA encoding a sulfate transporter from *Arabidopsis thaliana* FEBS Lett. 392: 95-99.
- Tan, J., Cowan, J.A. (1991) Enzymatic redox chemistry: A proposed reaction pathway for the six-electron reduction of  $\text{SO}_3^{2-}$  to  $\text{S}^{2-}$  by the assimilatory-type sulfite reductase from *Desulfovibrio vulgaris* (Hildenborough). *Biochemistry* 30: 8910-8917.
- Tei, H., Murata, K., Kimura, A. (1990a) Molecular cloning of the *cys* genes (*cysC*, *cysD*, *cysH*, *cysI*, *cysJ*, and *cysG*) responsible for cysteine biosynthesis in *Escherichia coli* K-12. *Biotech. Appl. Biochem.* 12: 212-216.
- Tei, H., Murata, K., Kimura, A. (1990b) Structure and expression of *cysX*, the second gene in the *Escherichia coli* K-12 *cysE* locus. *Biochem. Biophys. Res. Comm.* 167: 948-955.
- Tennigkeit, J., Matzura, H. (1991) Nucleotide sequence analysis of a chloramphenicol-resistance determinant from *Agrobacterium tumefaciens* and identification of its gene product. *Gene* 98: 113-116.
- Thomas, D., Barbey, R., Surdin-Kerjan, Y. (1990) Gene-enzyme relationship in the sulfate assimilation pathway of *Saccharomyces cerevisiae*. Study of the 3'-phosphoadenylylsulfate reductase structural gene. *J. Biol. Chem.* 265: 15518-15524.
- Thomas, D., Barbey, R., Sufdin-Kerjan, Y. (1990) Gene-enzyme relationship in the sulfate assimilation pathway of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 265: 15518-15524.
- Thompson, J.F. (1967) Sulfur metabolism in plants. *Ann. Rev. Plant Physiol.* 18: 59-84.
- Thompson, J.F., Moore, D.P. (1968) Enzymatic synthesis of cysteine and S-methylcysteine in plant extracts. *Biochem. Biophys. Res. Comm.* 31: 281-286.
- Tsang, M.L.-S. (1981) Assimilatory sulfate reduction in *Escherichia coli*:

- Identification of the alternate cofactor for adenosine 3'-phosphate 5'-phosphosulfate reductase as glutaredoxin. *J. Bacteriol.* 146: 1059-1066.
- Tsang, M.L.-S., Schiff, J.A. (1975) Studies of sulfate utilization by algae. 14. Distribution of adenosine-3'-phosphate-5'-phosphosulfate (PAPS) and adenosine-5'-phosphosulfate (APS) sulfotransferase in assimilatory sulfate reducers. *Plant Sci. Lett.* 4: 301-307.
- Tsang, M.L.-S., Schiff, J.A. (1976a) Studies of sulfate utilization by algae. 17. Reactions of the adenosine 5'-phosphosulfate (APS) sulfotransferase from *Chlorella* and studies of model reactions which explain the diversity of side products with thiols. *Plant Cell. Physiol.* 17: 1209-1220.
- Tsang, M.L.-S., Schiff, J.A. (1976b) Sulfate-reducing pathway in *Escherichia coli* involving bound intermediates. *J. Bacteriol.* 125: 923-933.
- Tsang, M.L.-S., Schiff, J.A. (1978) Studies of sulfate utilization by algae. 18. Identification of glutathione as a physiological carrier in assimilatory sulfate reduction by *Chlorella*. *Plant Sci. Lett.* 11: 177-183.
- Umbarger, H.E. (1978) Amino acid biosynthesis and its regulation. *Ann. Rev. Biochem.* 47: 533-606.
- Urlaub, H., Jankowski, G. (1982) Sulfate reduction in *Catharanthus roseus* (L.). Identification and subcellular localization of a particulate adenosine 5'-phosphosulfate-reducing activity in cells from cell suspension cultures. *Planta* 155: 154-161.
- Vaara, M. (1992) Eight bacterial proteins, including UDP-N-acetylglucosamine acyltransferase (LpxA) and three other transferases of *Escherichia coli*, consist of a six-residue periodicity theme. *FEMS Microbiol. Lett.* 97: 249-254.
- Van Camp, W., Bowler, C., Villarroel, R., Tsang, E.W.T., Van Montagu, M., Inzé, D. (1990) Characterization of iron superoxide dismutase cDNAs from plants obtained by genetic complementation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 87: 9903-9907.
- Vandenbol, M., Durand, P., Bolle, P.-A., Dion, C., Portetelle, D., Hilger, F. (1994) Sequence analysis of a 40.2 kb DNA fragment located near the left telomere of yeast chromosome X. *Yeast* 10:

1657-1662.

- Villand, P., Olsen, O.-A., Kleczkowski, L.A. (1993) Molecular characterization of multiple cDNA clones for ADP-glucose pyrophosphorylase from *Arabidopsis thaliana*. *Plant Mol. Biol.* 23: 1279-1284.
- von Arb, C., Brunold, C. (1983) Measurement of ferredoxin-dependent sulfite reductase activity in crude extracts from leaves using O-acetyl-L-serine sulphydrylase in a coupled assay system to measure the sulfide formed. *Anal. Biochem.* 131: 198-204.
- von Arb, C., Brunold, C. (1985) Ferredoxin-sulfite reductase and ferredoxin-nitrite reductase activities in leaves of *Pisum sativum*: Changes during ontogeny and in vitro regulation by sulfide. *Physiol. Plant.* 64: 290-294.
- von Heijne, G. (1992) Cleavage-site motifs in protein targeting sequences. In: Genetic engineering, Setlow, J.K. (ed.), Plenum Press, New York, pp 1-11.
- von Heijne, G., Nishikawa, K. (1991) Chloroplast transit peptides. The perfect random coil? *FEBS Lett.* 278: 1-3.
- von Heijne, G., Steppuhn, J., Herrmann, R.G. (1989) Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* 180: 535-545.
- Vuori, K., Myllylä, R., Pihlajaniemi, T., Kivirikko, K.I. (1992) Expression and site-directed mutagenesis of human protein disulfide isomerase in *Escherichia coli*. *J. Biol. Chem.* 267: 7211-7214.
- Vuorio, R., Härkönen, T., Tolvanen, M., Vaara, M. (1994) The novel hexapeptide motif found in the acyltransferases LpxA and LpxD of lipid A biosynthesis is conserved in various bacteria. *FEBS Lett.* 337: 289-292.
- Walker, J.E., Gay, N.J., Saraste, M., Eberle, A.N. (1984) DNA sequence around the *Escherichia coli* unc operon. *Biochem. J.* 224: 799-815.
- Warrilow, A.G.S., Clarkson, D.T., Hawkesford, M.J. (1995) Regulation of the sulphur-flux through the sulphur-assimilatory pathway. In: Book of abstracts, BBSRC Biochemistry of metabolic regulation in plants meeting, 11-13 June, University of Manchester, UK.

- Wenko, L.K., Treick, R.W., Wilson, K.G. (1985) Isolation and characterization of a gene encoding meso-Diaminopimelate Dehydrogenase from *Glycine max*. *Plant Mol. Biol.* 4: 197-204.
- Wheldrake, J.F., Pasternak, C.A. (1965) The control of sulphate activation in bacteria. *Biochem. J.* 96: 276-280.
- Wieles, B., van Soolingen, D., Holmgren, A., Offringa, R., Ottenhoff, T., Thole, J. (1995) Unique gene organization of thioredoxin and thioredoxin reductase in *Mycobacterium leprae*. *Mol. Microbiol.* 16: 921-929.
- Wilson, L.G., Asahi, T., Bandurski, R.S. (1961) Yeast sulfate-reducing system. I. Reduction of sulfate to sulfite. *J. Biol. Chem.* 236: 1822-1829.
- Wilson, Z.A., Dawson, J., Russell, J., Mulligan, B.J. (1991) *Arabidopsis thaliana*. *Biologist* 38: 163-169.
- Woodin, T.S., Wang, J.L. (1989) Sulfate permease of *Penicillium duponti*. *Exp. Mycol.* 13: 380-391.
- Wu, L., Ueda, T., Messing, J. (1995) The formation of mRNA 3'-ends in plants. *Plant J.* 8: 323-329.
- Wu, R. (1994) Development of the primer-extension approach: a key role in DNA sequencing. *Trends Biochem. Sci.* 19: 429-433.
- Yamaguchi, T., Masada, M. (1995) Comparative studies on cysteine synthase isozymes from spinach leaves. *Biochim. Biophys. Acta* 1251: 91-98.
- Yildiz, F.H., Davies, J.P., Grossman, A.R. (1994) Characterization of sulfate transport in *Chlamydomonas reinhardtii* during sulfur-limited and sulfur-sufficient growth. *Plant Physiol.* 104: 981-987.
- Young, L.J., Siegel, L.M. (1990) Alkaline low spin form of sulfite reductase heme protein subunit. *Biochem. Biophys. Res. Comm.* 168: 39-45.
- Youssefian, S., Nakamura, M., Sano, H. (1993) Tobacco plants transformed with the O-acetylserine (thiol) lyase gene of wheat are resistant to toxic levels of hydrogen sulphide gas. *Plant J.* 4: 759-769.
- Zink, M.W. (1984) Regulation of ATP-sulfurylase by various nitrogen and sulfur sources in cultured *Ipomoea* sp. *Can J. Bot.* 62: 2107-2113.