

ISOLATION AND CHARACTERISATION OF  
MOLYBDENUM COFACTOR BIOSYNTHESIS GENES  
FOR 'ASPERGILLUS NIDULANS'

Lindsey Jane Millar

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



1998

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/10023/14261>

This item is protected by original copyright

L

**ISOLATION AND CHARACTERISATION  
OF MOLYBDENUM COFACTOR BIOSYNTHESIS GENES  
FROM *Aspergillus nidulans***

**LINDSEY JANE MILLAR B. Sc. (Hons)**

**Sir Harold Mitchell Building  
School of Environmental and Evolutionary Biology  
University of St Andrews  
Greenside Place  
St Andrews  
Fife  
KY16 9TH  
SCOTLAND  
UK**



ProQuest Number: 10171184

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 10171184

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

Th 194

## DECLARATION

I, Lindsey Jane Millar, hereby certify that this thesis, which is approximately 66,000 words in length, has been written by me, that it is a record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Signature of candidate

Date

18<sup>th</sup> May 1998

I was admitted as a research student in October, 1994 and as a candidate for the degree of Doctor of Philosophy in October, 1994; the higher study for which this is a record was carried out in the University of St Andrews between 1994 and 1997.

Signature of candidate

Date

18<sup>th</sup> May 1998.

## COPYRIGHT

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.

Signature of candidate

Date

18<sup>th</sup> May, 1998

## CERTIFICATE

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.

Signature of supervisor

Date

12<sup>th</sup> May 1998

## ACKNOWLEDGEMENTS

I would firstly like to thank my supervisor Jim Kinghorn for his help and guidance throughout my studies. I am grateful for the helpful advice I received from Sheila Unkles and for the work which she carried out to sequence the *cnxE* mutant strains. I must also thank Ghassan Kan'an, Jacqueline Smith, Eddie Campbell, Virginia Appleyard, Takashi Kawaguchi and Robin Williams for help and friendship. Thanks also to Jörg Nieder (Braunschweig Technical University, Germany) for his valuable contribution to the isolation of the *cnxG* gene and for collaboration on sequence analysis of this locus.

Thanks also to David Boxer (Medical Sciences Institute, Dundee University) for allowing me to visit and work in his laboratory. Thank you to all the members of his laboratory for making me feel so welcome during my visits. Special thanks to Manuel Heck for his patient help in the preparation and analysis of HPLC samples and for many useful discussions.

I acknowledge the BBSRC (Biotechnology and Biological Sciences Research Council) for funding my Ph. D. research and for allowing me to attend the Fourth International Conference on Nitrate Assimilation: Molecular and Genetic Aspects, at the University of South Florida, Tampa, USA (3<sup>rd</sup>-9<sup>th</sup> May, 1997).

I would like to thank my parents Tom and June, my sister Hazel and my Grandmothers Helen and Jean for all their support and encouragement over the years, which has helped me to achieve this goal.

Finally, I dedicate this thesis to Kenny Gray, my fiancé who has been a great source of love and strength throughout my studies.

## CONTENTS

	<u>Page</u>
Declaration	i
Certificate	ii
Copyright	iii
Acknowledgements	iv
Contents	v
List of Figures and Tables	xiii
Abbreviations	xvii
Abstract	1
<b><u>Chapter 1 Introduction</u></b>	<b>3</b>
<b>1.1</b> <i>Aspergillus nidulans</i> as an Experimental Organism for Molecular Genetic Studies	3
<b>1.2</b> The Genetics of Molybdenum Cofactor Biosynthesis in <i>A. nidulans</i>	9
<b>1.2.1</b> Discovery of the <i>cnx</i> Genes and a 'Universal' Molybdenum Cofactor	9
<b>1.2.2</b> Individual <i>cnx</i> Loci	10
<b>1.3</b> The Molybdenum Cofactor	14
<b>1.4</b> Enzymes which Require the Molybdenum Cofactor for Activity	19
<b>1.4.1</b> Molybdoenzymes Present in <i>A. nidulans</i>	19
<b>1.4.1.1</b> Nitrate Reductase	19
<b>1.4.1.2</b> Xanthine Dehydrogenase (Purine Hydroxylase I) and Purine Hydroxylase II	25
<b>1.4.2</b> Other Molybdoenzymes	29

1.5	The Genetics of Molybdenum Cofactor Biosynthesis in Other Organisms	31
1.5.1	Other Molybdenum Cofactor Mutants Including Human Molybdenum Cofactor Deficiency	31
1.5.2	The Genetics of Molybdenum Cofactor Biosynthesis in <i>Escherichia coli</i>	33
1.5.3	Molybdenum Cofactor Biosynthesis Genes Isolated from Other Prokaryotes	38
1.5.4	Eukaryotic Molybdenum Cofactor Biosynthesis Genes	40
1.6	Research Objectives	41
<b><u>Chapter 2 Materials and Methods</u></b>		<b>42</b>
2.1	Strains	42
2.11	Bacterial Strains	42
2.12	Fungal Strains	43
2.2	Growth and Storage of Cultures	44
2.2.1	Bacterial Cultures	44
2.2.2	<i>A. nidulans</i> Cultures	44
2.3	Solutions	45
2.4	Media	45
2.4.1	<i>E. coli</i> Culture Media	45
2.4.2	<i>A. nidulans</i> Culture Media	47
2.5	Preparation of Plasmid/Cosmid DNA from <i>E. coli</i>	51
2.5.1	Small Scale ('mini') Preparation of Plasmid/Cosmid DNA	51
2.5.2	Medium Scale ('midi') and Large Scale ('maxi') Preparation of Plasmid/Cosmid DNA	51
2.6	Preparation of <i>A. nidulans</i> Chromosomal DNA	51

2.6.1	Small Scale Preparation	51
2.6.2	Large Scale Preparation	52
2.7	Spectrophotometric Quantitation of DNA	52
2.8	Restriction Endonuclease Digestion of DNA	53
2.9	Agarose Gel Electrophoresis of DNA	53
2.10	Isolation of DNA Fragments from Agarose Gels	54
2.11	Ethanol Precipitation of DNA	55
2.12	Protein Removal by Phenol Extraction of DNA	55
2.13	Drop Dialysis of DNA	55
2.14	<sup>32</sup> P Labelling of DNA	56
2.15	Southern Transfer and Hybridisation	57
2.16	Manipulation of the <i>A. nidulans</i> Cosmid Bank	60
2.16.1	Cosmid Propagation on LB Agar Plates	60
2.16.2	Hybridisation Reactions Using Cosmid DNA	60
2.17	Manipulation of the <i>A. nidulans</i> cDNA Library	61
2.17.1	Screening the cDNA Library	61
2.17.2	Excision of Purified cDNA Clones from the Uni-ZAP XR Vector	64
2.17.3	Mass Excision of the cDNA Library	64
2.18	Preparation of Competent <i>E. coli</i> Cells	64
2.19	Transformation of Competent <i>E. coli</i> Cells	65
2.20	Electro-transformation of <i>E. coli</i> Cells	66
2.20.1	Preparation of Cells	66
2.20.2	Electro-transformation	66
2.21	Preparation and Transformation of <i>A. nidulans</i> Protoplasts	67
2.21.1	Preparation of Protoplasts	67
2.21.2	Transformation Procedure	69
2.22	Polymerase Chain Reaction (PCR)	70

2.22.1	PCR Analysis of <i>A. nidulans</i> Transformant Colonies	71
2.22.2	PCR Amplification of Specific <i>A. nidulans</i> Genomic Gene Sequences	73
2.22.3	PCR Amplification of a DNA Fragment Suitably Engineered for Ligation into a Protein Expression Vector	76
2.22.4	PCR Amplification of DNA Fragments for use as Hybridisation Probes	77
2.23	Purification of PCR Products	78
2.24	Construction and Detection of Hybrid Plasmids	78
2.24.1	Ligation of DNA Fragments into Vectors	78
2.24.2	Detection of Recombinant Plasmids by $\alpha$ -complementation	79
2.24.3	Detection of Recombinant Plasmids by <i>in situ</i> Hybridisation	81
2.25	DNA Sequence Analysis	82
2.25.1	Sequence Analysis of Plasmid DNA	82
2.25.2	Sequence Analysis of PCR Products	87
2.26	Isolation of <i>A. nidulans</i> RNA	87
2.26.1	Total RNA Isolation	87
2.26.2	mRNA Isolation	90
2.27	Electrophoresis of RNA, Northern Transfer and Hybridisation	90
2.28	Spectrophotometric Quantitation of RNA	93
2.29	Nitrite Overlay Procedure for <i>E. coli</i>	94
2.30	Protein Over-Expression in <i>E. coli</i>	94
2.30.1	Time Course of Total Cell Protein Over-Expression	94
2.30.2	Lysozyme/Sonication Treatment of Cells to Determine Protein Solubility	95

2.31	SDS-Polyacrylamide Gel Electrophoresis of Proteins (SDS-PAGE)	96
2.32	HPLC Analysis of Molybdopterin Form A Dephospho and Compound Z Levels in <i>A. nidulans</i> Oxidised and Dephosphorylated Cell Free Extracts	98
2.33	Source of Plasmids	100
2.34	Chemicals	101
2.35	Containment and Safety	101

**Chapter 3 Isolation and Molecular Characterisation of the *cnxG* Gene** **102**

3.1	Brief Introduction	102
3.2	Initial Isolation of <i>cnxG</i> from the <i>A. nidulans</i> Genome	102
3.3	Molecular Analysis of pK	104
3.4	Molecular Analysis of the 2.2 kb <i>Bam</i> HI Fragment of pK	104
3.5	Isolation of <i>cnxG</i> Cosmid Clones	105
3.6	PCR Analysis of <i>A. nidulans</i> Colonies Transformed by pK and Cosmid Clone DNA	107
3.7	Isolation of <i>cnxG</i> cDNA clones	108
3.8	Nucleotide Sequence Analysis of the <i>cnxG</i> cDNA Clones	109
3.9	Construction of a <i>cnxG</i> Genomic Subclone	110
3.10	Structure of the <i>cnxG</i> Gene	112
3.11	CnxG Protein Sequence Similarities	114
3.12	Sequence Analysis of the <i>cnxG4</i> Mutant Strain	117
3.13	Northern Analysis of the <i>cnxG</i> Gene	118
3.14	Further Investigation of the Second <i>cnxG</i> Transcript	119
3.15	Over-expression of the CnxG Protein	120

3.15.1	Construction of a CnxG Protein Over-Expression Plasmid	120
3.15.2	Time Course and Yield of CnxG Protein Over-Expression	121
3.15.3	Determination of CnxG Protein Solubility	122

**Chapter 4 Isolation and Molecular Characterisation of the *cnxE* Gene** **123**

4.1	Brief Introduction	123
4.2	Initial Cloning Strategy: Transformation of <i>A. nidulans cnxE</i> Mutant Strains with Wild-type Genomic <i>A. nidulans</i> DNA	124
4.3	Alternative Cloning Strategy: Complementation of an <i>E. coli mogA</i> Mutant Strain with an Excised <i>A. nidulans</i> cDNA Library	125
4.3.1	Transformation of <i>E. coli</i> Strain JL3705 Competent Cells	125
4.3.2	Electro-transformation of <i>E. coli</i> Strain JL3705	127
4.4	Retransformation of the <i>E. coli mogA</i> Mutant (Strain JL3705) with pE1	128
4.5	Nitrate Reductase Activity in the pE1-Transformed <i>E. coli mogA</i> Mutant Strain	129
4.6	Sequence Analysis of pE1	130
4.7	Isolation of Genomic Cosmid Clones Equivalent to Region 1/Region 2 of pE1	131
4.8	Further Investigation of the pE1 Region 2 Sequence	132

<b>4.9</b>	Construction of a Subclone Containing the Genomic Counterpart of the ORF Contained Within Region 1 of pE1	133
<b>4.10</b>	Structure of the <i>cnxE</i> Gene	134
<b>4.11</b>	CnxE Protein Sequence Similarities	136
<b>4.12</b>	Sequence Analysis of the <i>cnxE</i> Gene from Mutant Strains	140
<b>4.13</b>	Northern Analysis of the <i>cnxE</i> Gene	141
<b>4.14</b>	Testing for pE1-Complementation of the <i>moeA</i> Mutant Phenotype	142
 <b><u>Chapter 5 HPLC Analysis of <i>cnx</i> Mutants</u></b>		<b>145</b>
<b>5.1</b>	Brief Introduction	145
<b>5.2</b>	HPLC Analysis of the <i>A. nidulans</i> Wild-type Strain Grown With Different Sole Nitrogen Sources	146
<b>5.2.1</b>	Precursor Z Levels	147
<b>5.2.2</b>	Molybdopterin Levels	147
<b>5.3</b>	HPLC Analysis of <i>A. nidulans cnxA</i> , <i>cnxB</i> and <i>cnxC</i> Mutant Strains	147
<b>5.3.1</b>	Precursor Z Levels	148
<b>5.3.2</b>	Molybdopterin Levels	148
<b>5.4</b>	HPLC Analysis of <i>A. nidulans cnxG</i> Mutant Strains	148
<b>5.4.1</b>	Precursor Z Levels	148
<b>5.4.2</b>	Molybdopterin Levels	149
<b>5.5</b>	HPLC Analysis of an <i>A. nidulans cnxE</i> Mutant Strain	149
<b>5.5.1</b>	Precursor Z Level	149
<b>5.5.2</b>	Molybdopterin Level	149

<b><u>Chapter 6 Attempts to Isolate an <i>E. coli moaD</i> Gene</u></b>	<b>150</b>
<b><u>Equivalent in <i>A. nidulans</i></u></b>	<b>150</b>
6.1 Brief Introduction	150
6.2 Cloning Strategy: Complementation of an <i>E. coli moaD</i> Mutant Strain with an Excised <i>A. nidulans</i> cDNA Library	150
6.3 Testing the Putatively Complemented <i>moaD</i> Mutant Colonies for Nitrate Reductase Activity	152
6.4 Subsequent Attempts at Electro-transformation of Strain MJ431 with an Excised <i>A. nidulans</i> cDNA Library	153
<b><u>Chapter 7 Discussion and Conclusions</u></b>	<b>155</b>
7.1 The <i>A. nidulans cnxG</i> Locus	155
7.2 The <i>A. nidulans cnxE</i> Locus	164
7.3 The <i>A. nidulans cnxABC</i> Locus	177
7.4 Regulation of MoCo Biosynthesis in <i>A. nidulans</i>	178
7.5 An Overview of Molybdenum Cofactor Biosynthesis in <i>A. nidulans</i>	181
7.5.1 The Proposed Function of Other <i>cnx</i> Genes	181
7.5.2 Does a <i>moaD</i> Gene Equivalent Exist in <i>A. nidulans</i> ?	184
7.6 Concluding Remarks	186
Bibliography	188
Appendix	236

## LIST OF FIGURES AND TABLES

### LIST OF FIGURES

#### Chapter 1

- Figure 1: Linkage map of *A. nidulans*
- Figure 2a: Structure of the molybdenum cofactor -molybdopterin form
- Figure 2b: Structure of the molybdenum cofactor -molybdopterin guanine dinucleotide form
- Figure 3: Schematic diagram of the nitrate reductase homodimer
- Figure 4: The purine degradation pathway in *A. nidulans*
- Figure 5: Putative consensus sequence for the molybdenum cofactor binding domain
- Figure 6: Representative model of MoCo biosynthesis in *E. coli*

#### Chapter 3

- Figure 7: Complementation of strain *cnxG4* with pK
- Figure 8: Restriction digest pattern of the *cnxG* cosmid clones
- Figure 9: PCR analysis of plasmid K, cosmid clone 4 and cosmid clone 5 *A. nidulans* transformant colonies
- Figure 10: Identification of a *cnxG* cosmid fragment suitable for subcloning
- Figure 11: Complementation of strain *cnxG4* by pLJM18
- Figure 12: Nucleotide sequence and inferred amino acid sequence of the *cnxG* gene
- Figure 13: Alignment of the deduced amino acid sequence of *A. nidulans* CnxG with the proteins Yhr2 from

*S. cerevisiae*, MetB from *E. coli*, MetB from *H. pylori*,  
MegL from *P. putida* and Cys3 from *S. cerevisiae*.

Figure 14: Northern analysis of the *cnxG* gene

Figure 15: Southern blot analysis of the second *cnxG* transcript

Figure 16: Time course and yield of CnxG protein over-expression

Figure 17: Determination of CnxG solubility

#### **Chapter 4**

Figure 18: Complementation of *E. coli* strain JL3705 with the recombinant vector pE1

Figure 19: Detection of nitrate reductase activity in pE1-transformed *E. coli* strain JL3705

Figure 20: Schematic map of pE1

Figure 21: Restriction digest pattern of the cosmid clones which hybridised to a probe taken from within the ORF identified in region 1 of the cDNA clone pE1

Figure 22: Identification of a cosmid fragment containing the genomic counterpart of the pE1 region1 ORF suitable for subcloning

Figure 23: Nucleotide sequence and inferred amino acid sequence of the *cnxE* gene

Figure 24: The two similar amino acid regions found between the *A. nidulans* CnxE and *A. thaliana* Cnx1 sequences

Figure 25: Alignment of the deduced amino acid sequence of *A. nidulans* *cnxE* with Gephyrin from *R. norvegicus*, Cinnamon from *D. melanogaster* and MoeA, MogA and MoaB from *E. coli*.

Figure 26: Northern analysis of the *cnxE* gene

Figure 27: Testing for complementation of *E. coli* strain NS9 with pE1

## **Chapter 5**

- Figure 28: Conversion of precursor Z to compound Z and molybdopterin to Form A dephospho
- Figure 29: Levels of precursor Z detected in the *A. nidulans* wild-type strain grown with various different sole nitrogen sources
- Figure 30: Levels of molybdopterin detected in the *A. nidulans* wild-type strain grown with various different sole nitrogen sources
- Figure 31: Levels of precursor Z detected in the *A. nidulans* *cnxA9*, *cnxB11* and *cnxC3* mutant strains
- Figure 32: Levels of molybdopterin detected in the *A. nidulans* *cnxA9*, *cnxB11* and *cnxC3* mutant strains
- Figure 33: Levels of precursor Z detected in the *A. nidulans* *cnxG4*, *cnxG1222* and *cnxG2* mutant strains
- Figure 34: Levels of molybdopterin detected in the *A. nidulans* *cnxG4*, *cnxG1222* and *cnxG2* mutant strains
- Figure 35: Levels of precursor Z detected in the *A. nidulans* *cnxE14* mutant strain
- Figure 36: Levels of molybdopterin detected in the *A. nidulans* *cnxE14* mutant strain

## **Chapter 7**

- Figure 37: Sulphur amino acid metabolism in *A. nidulans*
- Figure 38: Domain organisation in the *A. nidulans* CnxE protein
- Figure 39: Proposed overview of MoCo biosynthesis in *A. nidulans*

## **LIST OF TABLES**

### **Chapter 3**

- Table 1: Isolation of *cnxG* by transformation of mutant strain *cnxG4* with *A. nidulans* genomic DNA
- Table 2: Protein similarities to *cnxG*
- Table 3: Location, nucleotide sequence and amino acid sequence of the *cnxG4* mutation

### **Chapter 4**

- Table 4: Attempts to isolate *cnxE* by transformation of *cnxE* mutant strains with *A. nidulans* genomic DNA
- Table 5: Protein similarities to *cnxE*
- Table 6: Location, nucleotide sequence and amino acid sequence of *cnxE* mutations

### **Chapter 7**

- Table 7: Signature pattern for proteins with similarity to *E. coli* MogA and MoaB

## ABBREVIATIONS

A	adenine
APS	ammonium persulphate
BCA	bichinchoninic acid
bp	base pair(s)
BPB	bromophenol blue
BSA	bovine serum albumin
camMPT	dicarboxamidomethylmolybdopterin
cm	centimetre(s)
C	cytosine
cDNA	complementary deoxyribonucleic acid
CHEF	contour-clamped homogeneous electric field
dATP	2'-deoxy-adenosine-5'-triphosphate
dCTP	2'-deoxy-cytidine-5'-triphosphate
dGTP	2'-deoxy-guanosine-5'-triphosphate
dNTP	2'-deoxy-nucleoside-5'-triphosphate
DEPC	diethyl pyrocarbonate
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
EC	Enzyme Commission
EDTA	ethylenediaminetetraacetic acid
<i>e.g.</i>	<i>exempli gratia</i> (for example)
<i>et al.</i>	<i>et alia</i> (and others)
FeMoCo	iron molybdenum cofactor
FAD	flavin adenine dinucleotide (oxidised form)
g	gram(s)
G	guanine
GMP	guanosine 5'-monophosphate

GuSCN	guanidine isothiocyanate
HPLC	High Performance Liquid Chromatography
h	hour(s)
<i>i.e.</i>	<i>id est</i> (that is)
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
kb	kilobase pair(s)
kDa	kiloDalton(s)
kV	kiloVolt(s)
l	litre(s)
lb	pounds
MAD	molybdopterin adenine dinucleotide
Mb	megabase(s)
MCD	molybdopterin cytosine dinucleotide
$\mu$ Ci	microCuries
$\mu$ g	microgram(s)
MGD	molybdopterin guanine dinucleotide
MHD	molybdopterin hypoxanthine dinucleotide
$\mu$ l	microlitre(s)
$\mu$ M	microMolar
M	molar
mg	milligram(s)
min	minute(s)
ml	millilitre(s)
mm	millimetre(s)
mM	millimolar
MoCo	molybdenum cofactor
MOPS	morpholinopropanesulfonic acid
MPT	molybdopterin
mRNA	messenger ribonucleic acid

MVH	reduced methyl viologen
NAD	nicotinamide adenine dinucleotide (oxidised form)
NADH	nicotinamide adenine dinucleotide (reduced form)
NADP	nicotinamide adenine dinucleotide phosphate (oxidised form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
ng	nanogram(s)
N-1-NED	N-1-naphthylethylenediamine
nm	nanometre(s)
No.	number
NR	nitrate reductase
°C	degrees Centigrade
OD	optical density
ORF	open reading frame
PABA	para-amino benzoic acid
PAGE	polyacrylamide-gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units
PHI	purine hydroxylase I
PHII	purine hydroxylase II
PVP	polyvinyl pyrrolidone
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
SAP	shrimp alkaline phosphatase

SDS	sodium dodecyl sulphate
sec	second(s)
S	Svedberg unit
T	thymine
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	tris (hydroxymethyl) methylamine
UV	ultra violet
V	Volt(s)
vol	volume(s)
W	Watt(s)
x	times
X-gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside
=	equal to
$\equiv$	equivalent to
$\div$	divided by
-	minus
+	plus
%	percent

## ABSTRACT

The genes *cnxG* and *cnxE* from *Aspergillus nidulans* have been cloned and characterised. *cnxG* was isolated by functional complementation of the *A. nidulans* *cnxG4* mutant strain. Within the *cnxG4* mutant a single amino acid change from valine 250 to phenylalanine has been identified suggesting that this residue is critical for CnxG protein functionality. *cnxG* exhibits similarity at the amino acid level to several enzymes involved in catalysing transsulphuration reactions during the biosynthesis of cysteine and methionine. HPLC analysis of several *cnxG* mutant strains has demonstrated that precursor Z is accumulated to 17.2 -25.3 times that observed in the wild-type and that molybdopterin is not detectable. The evidence suggests that CnxG may be involved in the initial donation of sulphur to the MoCo biosynthesis pathway. *cnxE* was isolated by functional complementation of an *Escherichia coli* *mogA* mutant strain. *cnxE* exhibits similarity at the amino acid level to the *E. coli* MoeA, MogA and MoaB proteins and to the eukaryotic proteins Gephyrin from rat, Cnx1 from *Arabidopsis thaliana* and Cinnamon from *Drosophila melanogaster*. Analysis of the deduced amino acid sequences of several *cnxE* mutants has allowed the identification of residues which are critical for CnxE to function. HPLC analysis of the *cnxE14* mutant indicated a modest accumulation of precursor Z and molybdopterin levels in comparison to the wild-type however the increase was too slight to conclude a significant effect. It would appear that CnxE is involved in the processing and/or incorporation of molybdenum into molybdopterin. Northern analysis has indicated that the transcription of both *cnxG* and *cnxE* is not subject to regulation by nitrate or ammonium. Mutations within *cnxABC* were shown to result in a significant reduction of precursor Z levels and

abolition of molybdopterin levels in comparison to the wild-type, suggesting a function for CnxABC in the synthesis of precursor Z. In the wild-type, precursor Z levels were found to be 3.5 times higher in cells grown on nitrate than in those grown on ammonium as sole nitrogen source, perhaps indicating a degree of nitrate induction exerted on the early part of the pathway.

# CHAPTER 1

## INTRODUCTION

### 1.1      *Aspergillus nidulans* as an Experimental Organism for Molecular Genetic Studies

In the Orient species of the filamentous fungus *Aspergillus* such as *A. oryzae* have been used by man for centuries during the 'koji' stage of several food fermentations, including soy sauce production (Hesseltine, 1965). Today, many aspergilli are used in the biotechnology industry where their metabolic capabilities to produce several organic acids (*e.g.* citric acid) and many useful enzymes (*e.g.* lipase) have been exploited by man (Bodie *et al.*, 1994). Although *A. nidulans* is not of intrinsic metabolic commercial significance, this species has a number of properties which make it ideal for use as an experimental organism in which to study the molecular genetics of the filamentous fungi.

Before the introduction of recombinant DNA procedures *A. nidulans*, together with *Neurospora crassa*, was the main organism of choice for physiological, biochemical and classical genetic studies in the filamentous fungi. Indeed, the only other lower eukaryotic system equally well studied in this respect was the unicellular budding yeast *Saccharomyces cerevisiae*. The advantages of *A. nidulans* for these early studies (as well as for studies today) included its ease of handling, simple nutrient requirements, metabolic versatility and rapid life cycle. Features of the sexual, asexual and parasexual life cycles of *A. nidulans* were exploited for many aspects of classical genetic study including the mapping of gene mutations to specific chromosomal locations (reviewed by Martinelli, 1994 and references therein). On the basis of this work a

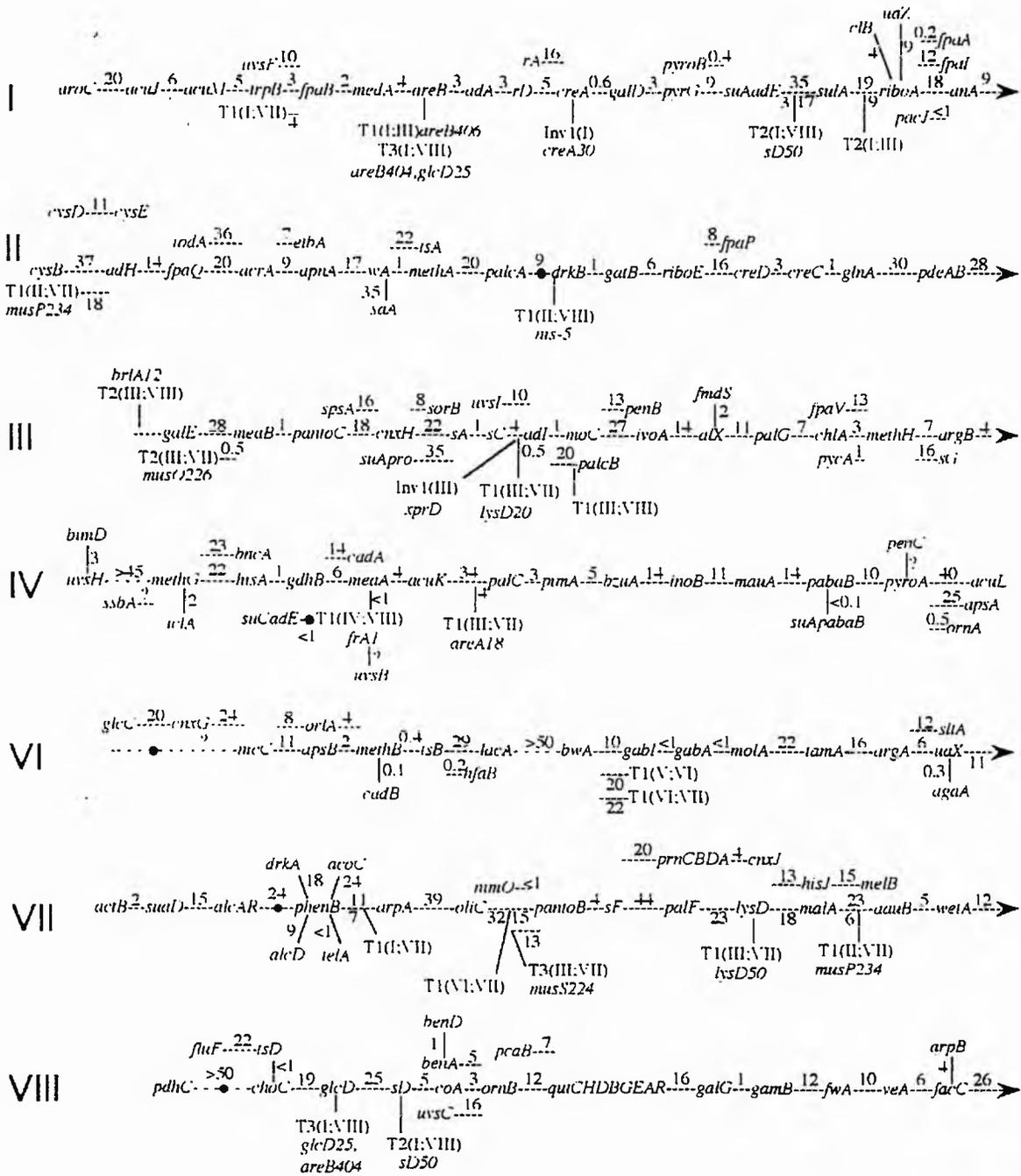
detailed genetic linkage map of *A. nidulans* has been established (Clutterbuck, 1994, refer to Figure 1; Clutterbuck, 1997).

In the early 1970's, with the advent of recombinant DNA technology in bacteria, the main thrust for the development of molecular genetic systems in filamentous fungi was given to those organisms which had previously been well-characterised by classical methods *i.e.* *A. nidulans* and *N. crassa*.

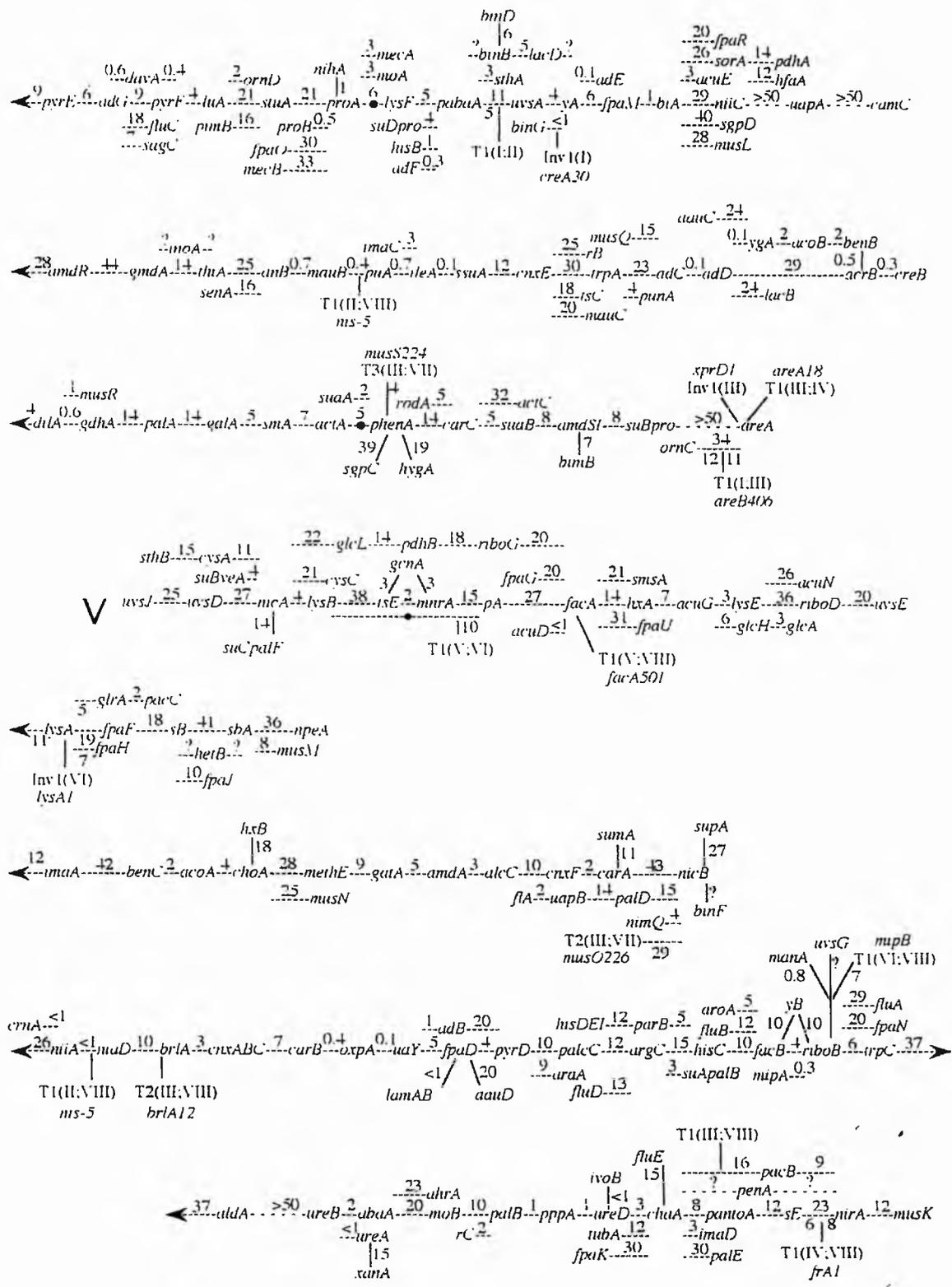
The advancement of molecular genetic research in *A. nidulans* was boosted with the development of an efficient transformation (gene transfer) system in the early 1980's. A transformation system permits the introduction of exogenous DNA into the fungal cell and allows the selection of those cells expressing incorporated DNA. The earliest, and by far the most commonly used transformation system, involved the uptake of DNA by protoplasts in the presence of calcium ions ( $\text{CaCl}_2$ ) and polyethylene glycol (PEG) followed by their regeneration to produce transformed fungal colonies (Balance *et al.*, 1983; Tilburn *et al.*, 1983; John and Peberdy, 1984; Yelton *et al.*, 1984). Although rarely used, successful transformation of filamentous fungi has also been achieved by electroporation of protoplasts (Richey *et al.*, 1989), using lithium acetate treatment of intact fungal spores (Dhawale *et al.*, 1984; Bej and Perlin, 1989) and 'biolistics' (biological-ballistics) whereby tungsten microprojectiles coated with DNA are accelerated at high velocity directly into fungal spores or hyphae (Lorito *et al.*, 1993). For a recent, comprehensive discussion of transformation in filamentous fungi refer to Riach and Kinghorn (1996).

In *A. nidulans* successful transformation is most commonly detected by the use of nutritional selective markers or dominant selection markers. Nutritional selective markers are previously cloned genes which have the ability to complement corresponding auxotrophic mutations *e.g.*

Figure 1: Linkage map of *A. nidulans* (taken from Clutterbuck, 1994)



Linkages are given as uncorrected percentages. Where the distance but not the orientation of a linkage is known, it is shown as a vertical or diagonal line. Centromeres are shown as closed circles.



the *argB* gene, which allows transformed *argB* mutants to utilise arginine as sole nitrogen source (John and Peberdy, 1984). Dominant selection markers are mainly antibiotic resistance genes which permit both wild-type and mutant strains to grow in the presence of the appropriate antibiotic *e.g.* the *hph* gene derived from *E. coli*, which confers resistance to hygromycin (Punt *et al.*, 1987).

Selection markers are particularly useful when cotransformed with a gene for which there is no direct selection (Wernars *et al.*, 1987). Cotransformation is based on the finding that when recipient cells are simultaneously exposed to two different kinds of transforming DNA, there is a high probability that a cell which takes up one molecule will also take up the other. One example of the usefulness of cotransformation is in determining whether vectors for the expression of mammalian proteins have been successfully incorporated into fungal cells.

The efficient transformation system available in *A. nidulans* provides an attractive method for gene isolation and is an alternative to other, sometimes more limited, cloning approaches which include: differential hybridisation; DNA homology to heterologous probes; chromosomal position; antibody recognition and functional complementation of *E. coli* or *S. cerevisiae* mutants (reviewed by Kinghorn and Unkles, 1994 and references therein). So called self-cloning of genes by transformation involves the use of a genomic library to complement an appropriate mutant strain (from the same species), followed by isolation of the DNA responsible. In *A. nidulans*, many loci are represented by well-defined mutations making gene isolation by this method especially convenient.

Most plasmids transformed into *A. nidulans* are maintained by becoming integrated into the chromosomal DNA by recombination (reviewed by Fincham *et al.*, 1989). Much effort has been directed

towards the search for sequences that enable autonomous replication of plasmids within the cell. Recently, a plasmid designated Arp1 (*Aspergillus* replicating plasmid) was isolated. This plasmid consists of sequences derived from the gene bank vector pILJ16 which carries the *argB* gene marker and a 6.1 Kb insert sequence, AMA1, responsible for autonomous replication in *A. nidulans* (Gems *et al.*, 1991; Aleksenko *et al.*, 1996; Aleksenko and Clutterbuck, 1997). The AMA1 sequence within Arp1 results in a 250-fold enhancement of transformation frequency when compared with frequencies of integrative vectors (Gems *et al.*, 1991). With Arp1, cotransformation frequencies are also increased *e.g.* the transformation of a *trpC*<sup>-</sup> strain with the normally integrative *trpC*<sup>+</sup> plasmid pTA11 was increased approximately 100-fold with the addition of Arp1 to the transformation mixture (Gems *et al.*, 1991).

The vector Arp1 and its derivatives pDHG25 and pHELP1 are useful genetic tools which enhance the ability to clone genes by a mutant complementation strategy. An 'instant gene bank' can be synthesised by cotransforming an autonomously replicating vector (Arp1, pDHG25 or pHELP1) with fragmented wild-type *A. nidulans* chromosomal DNA (Gems *et al.*, 1993). Alternatively, a conventional genomic library constructed in a typical integrative vector can be cotransformed with the autonomously replicating vector of choice.

In such gene cloning experiments, a replicating cointegrate plasmid derived from the two transforming DNA species is formed. In *A. nidulans* cointegrate plasmids are formed by homologous and non-homologous recombination as well as by end-to-end ligation of linear fragments (Aleksenko, 1994). The vector pDHG25 and the markerless pHELP1 are particularly suitable for this method of gene isolation as they both contain a unique *Bam*HI site which may be used to direct the insertion of cotransformed (genomic) DNA sequences at this position. For

the molecular structures of Arp1, pDHG25 and pHELP1 refer to the Appendix. Recombinant plasmids existing within transformant colonies can then be recovered with relative ease and without recourse to excision from the chromosomal DNA. One disadvantage of using this technique to isolate genes is that rearrangement of sequences during cointegrate formation is observed. Fragments of a gene may still however be isolated which can be used as hybridisation probes to an appropriate gene library (Kinghorn and Unkles, 1994).

Contour-clamped homogeneous electric field (CHEF) electrophoresis has been used to separate the chromosomes of *A. nidulans* (Brody and Carbon, 1989). Eight chromosomes were identified and equivalence with the linkage groups, previously identified by classical genetic analysis over many decades, was established. In order of decreasing size, the *A. nidulans* chromosomes have been estimated to be 5.0 Megabases (Mb) (VIII), 4.5 Mb (VII), 4.2 Mb (II), 3.8 Mb (I and V), 3.5 Mb (III and VI) and 2.9 Mb (IV) with a total genome size of 31 Mb. The ability to separate the chromosomes has permitted the generation of a chromosome-specific cosmid DNA library for *A. nidulans* which is proving useful in molecular genetic studies (Brody *et al.*, 1991). The chromosome-specific library was constructed using two complete cosmid libraries which consist of 35-50 kb genomic wild-type *A. nidulans* DNA fragments held within LORIST 2 (Gibson *et al.*, 1987) and pWE15 (Wahl *et al.*, 1987) vectors (for molecular structures refer to the Appendix).

The development of a physical or contig map for *A. nidulans* is currently underway using the chromosome-specific subcollections identified by Brody and colleagues (1991). Overlapping cosmid clones can be organised into large contiguous chromosomal regions designated contigs which will eventually build into a complete physical map where all of the appropriate cosmid clones are positioned relative to one another

along each of the chromosomes. The development of a such a contig map should greatly facilitate position-based cloning by chromosome walking from a previously cloned gene, with reference to the classical genetic map.

The first *bona fide* gene cloned from *A. nidulans* was *aromA* (involved in the shikimate pathway) which was isolated on the basis of complementation of a corresponding *E. coli* mutant (Kinghorn and Hawkins, 1982). By 1994 at least 82 genes had been isolated from *A. nidulans* and their sequences determined (reviewed by Martinelli, 1994), and since then many more genes have been cloned and characterised. The isolation and characterisation of loci in *A. nidulans* provides important information on gene structure, function and regulation which may then be applied to the biotechnological manipulation of other fungal genes. *A. nidulans*, along with *N. crassa*, is often termed the model filamentous fungus because molecular genetic systems and procedures developed in this organism can then be applied to less tractable, but more industrially important, species such as *Penicillium chrysogenum* (Penicillin production), *Acremonium chrysogenum* (Cephalosporin production) and *Aspergillus niger* (citric and gluconic acid production) (Timberlake and Marshall, 1989). Although *A. nidulans* is not of intrinsic commercial importance in the production of useful metabolites it is now being exploited by biotechnology companies as a host organism for the heterologous expression and secretion of industrial proteins (van den Hondel *et al.*, 1991). Particular interest has centred around the use of *A. nidulans* as a host for the production of novel mammalian proteins. Successful expression of many such proteins including human interferon  $\alpha$ -2 (Gwynne *et al.*, 1987), cattle tick cell-surface glycoprotein (Bm86) (Turnbull *et al.*, 1990) and human interleukin-6 (Contreras *et al.*, 1991) has already been achieved.

## 1.2 The Genetics of Molybdenum Cofactor Biosynthesis in *A. nidulans*

### 1.2.1 Discovery of the *cnx* Genes and a 'Universal' Molybdenum Cofactor

Over three decades ago, Pateman and colleagues generated a number of *A. nidulans* mutants, termed *cnx*, which lacked both nitrate reductase and xanthine dehydrogenase (since renamed purine hydroxylase I and purine hydroxylase II) activities (Pateman *et al.*, 1964). They proposed that these enzymes share a common molybdenum cofactor which is synthesised by the products of five unlinked genetic loci, designated *cnxABC*, *cnxE*, *cnxF*, *cnxG* and *cnxH*. This was the first suggestion that a molybdenum cofactor common to several molybdenum-containing enzymes might exist.

The *A. nidulans* *cnx* mutants are distinguished by their inability to grow on nitrate or hypoxanthine as sole nitrogen source (Pateman *et al.*, 1964). A mutant strain of *N. crassa* termed *nit-1*, which displays the same growth characteristics as the *cnx* mutants was investigated by Nason and coworkers (Nason *et al.*, 1970, 1971; Ketchum *et al.*, 1970). Their studies showed that the inactive apoprotein of nitrate reductase in the *nit-1* strain could be reconstituted *in vitro* by an appropriate source of molybdenum cofactor (*i.e.* cofactor provided by wild-type *N. crassa* or *nit* mutants other than *nit-1*). This provided evidence for the transferability of the cofactor and the *nit-1* system became the first biological assay for molybdenum cofactor activity. It was also shown that acidified preparations of aldehyde oxidase from rabbit liver, xanthine dehydrogenase from chicken liver, nitrate reductase from *E. coli* and foxtail or bovine xanthine oxidase and sulphite oxidase all provided a substance capable of reconstituting nitrate reductase activity in extracts of the *nit-1* strain (Ketchum *et al.*, 1970; Nason *et al.*, 1971). These findings

suggested the existence of a universal molybdenum-containing cofactor which is present in a similar if not indistinguishable form in a wide range of molybdoenzymes from bacterial, plant and animal sources.

Wild-type or *niaD* mutant (defective in the nitrate reductase structural protein) strains of *A. nidulans* have since tested positively as sources of molybdenum cofactor in the *nit-1* reconstitution assay (Ketchum and Downey, 1975; Garrett and Cove, 1976). Garrett and Cove (1976) also reported reconstitution of *A. nidulans* *cnxA6*, *cnxE29*, *cnxF12*, *cnxG4* and *cnxH3* mutant strains using the *A. nidulans* deletion mutant *niaD26* as cofactor source and concluded that all *A. nidulans* *cnx* mutants tested were similar to the *N. crassa nit-1* strain in their ability to produce NADPH-nitrate reductase in the presence of molybdenum cofactor.

To date, all evidence indicates that the molybdenum cofactor is basically an identical structural and functional component of all molybdoenzymes, with the exception of nitrogenase (Pienkos *et al.*, 1977). Nitrogenase possesses a distinct cofactor containing iron as well as molybdenum *i.e.* FeMoCo (Shah and Brill, 1977; Hoover *et al.*, 1989; Kim and Rees, 1992). For a recent review of FeMoCo biosynthesis refer to Allen *et al.*, (1994).

### **1.2.2 Individual *cnx* Loci**

Five *cnx* loci (*cnxABC*, *cnxE*, *cnxF*, *cnxG* and *cnxH*) were originally described in *A. nidulans* (Pateman *et al.*, 1964). Later a sixth locus termed *cnxJ* was identified, however mutations within this gene only seem to impair rather than abolish MoCo biosynthesis (Arst *et al.*, 1982).

*cnxABC* (located on chromosome VIII):

The *cnxABC* locus comprises three tightly linked mutations, *cnxA*, B and C which exhibit a complex complementation pattern (Cove and Pateman, 1963; Pateman *et al.*, 1964; Hartley, 1970). *cnxA* and *cnxC* mutants were found to complement each other, but not *cnxB* mutants, when combined as heterokaryons and grown on nitrate (Cove and Pateman, 1963). This complementation pattern is similar to that found in fungal genes which code for multifunctional proteins *e.g.* the first two steps of pyrimidine biosynthesis in *N. crassa* which are specified by the single gene *pyr-3* (Williams *et al.*, 1970). It was unclear whether *cnxA* and *cnxC* represented two separate genes with *cnxB* mutants lacking both activities or if a single gene was present with *cnxA* and *cnxC* mutants exhibiting intragenic complementation.

*cnxE* (located on chromosome II):

Strains possessing mutations within the *cnxE* locus are interesting in that they can be repaired, for growth on nitrate and hypoxanthine, by the addition of 33 mM sodium molybdate to the medium (Arst *et al.*, 1970). When assayed for nitrate reductase and xanthine dehydrogenase it was determined that growth of *cnxE* mutants on molybdate resulted in the production of active enzymes. The enzyme activities were however only repaired if molybdate was added to the growth medium and not if added directly to the enzyme assay suggesting that the repair process requires metabolic transformation of molybdate which was not possible in the *in vitro* assay system (Arst *et al.*, 1970). Temperature-sensitive mutations of *cnxE* have been identified which produce functional nitrate reductase (NR) only at a permissive temperature (MacDonald and Cove, 1974). Such temperature-sensitive *cnxE* mutants produce a NR with the same half-life as NR from wild-type

strains suggesting that *cnxE* is probably involved in the enzymatic synthesis of MoCo. It has been proposed that the *cnxE* gene product may be an enzyme which acts directly on molybdate, perhaps to incorporate it into the MoCo molecule as a final step in the biosynthesis process (Arst *et al.*, 1970). Under high (33 mM) molybdate levels alternative processes, perhaps non-enzymatic, may be sufficient for the incorporation of molybdenum into the cofactor therefore explaining the repairable phenotype exhibited by *cnxE* mutant strains.

*cnxF* (located on chromosome VII):

Temperature-sensitive mutations within *cnxF* have been identified which, like the *cnxE* temperature-sensitive strains, produce nitrate reductase (only at the permissive temperature) with a half-life which is comparable to that of wild-type nitrate reductase (MacDonald and Cove, 1974). This suggests that the *cnxF* gene product may play an enzymatic role in MoCo biosynthesis.

*cnxG* (located on chromosome VI):

No temperature-sensitive alleles of *cnxG* have been isolated (MacDonald and Cove, 1974) and the possible role of the *cnxG* gene product in MoCo biosynthesis has never attracted speculation.

*cnxH* (located on chromosome III):

This locus is interesting in that the nitrate reductase molecules produced by temperature-sensitive mutants (only at the permissive temperature) have shorter half-lives than wild-type nitrate reductase (MacDonald and Cove, 1974). This may imply that the *cnxH* gene codes for a protein which forms a structural component of the nitrate reductase molecule, perhaps a polypeptide chain to which the cofactor is attached.

The three mutant strains (*biA1 cnxH11*, *biA1 cnxH17* and *biA1 puA2 pyroA4 cnxH21*) studied by MacDonald and Cove, while being temperature-sensitive for NR, were found to behave as mutants at all temperatures for both purine hydroxylases (Scazzocchio *et al.*, 1973). This finding was in contrast to temperature-sensitive *cnxE* and *cnxF* mutant strains which can produce all three MoCo-containing enzymes at the permissive temperature and lack all three enzyme activities at the non-permissive temperature (Scazzocchio *et al.*, 1973). This finding could be explained if the requirement for the protein encoded by the *cnxH* gene was more stringent for the purine hydroxylases than for nitrate reductase (Scazzocchio, 1980). It was also suggested that the *cnxH* gene could not specify an enzymatic step of MoCo synthesis because any mutation in such a gene should have a uniform effect on the three cofactor-containing enzymes (Scazzocchio, 1974).

*cnxJ* (located on chromosome VII):

The existence of a sixth *cnx* locus designated *cnxJ* was postulated by Arst *et al.*, (1982). Studies showed that strains carrying *cnxJ1* or *cnxJ2* mutations, when grown in the presence (but not in the absence) of tungstate or methylammonium, had reduced molybdoenzyme levels as judged from both growth properties and enzyme determinations. *cnxJ* mutations differ from other types of *cnx* mutation in that they effectively lower but do not abolish any or all of the three molybdoenzyme activities found in *A. nidulans*. If the *cnxJ* gene product is dispensible for cofactor synthesis it does not seem likely that it forms a structural component of the cofactor or functions as an enzyme catalysing a step in the cofactor biosynthesis pathway. It has been suggested that *cnxJ* might be a positive-acting regulatory gene whose product is necessary for a high, but not a moderate level of cofactor synthesis (Arst *et al.*, 1982).

### **1.3      The Molybdenum Cofactor**

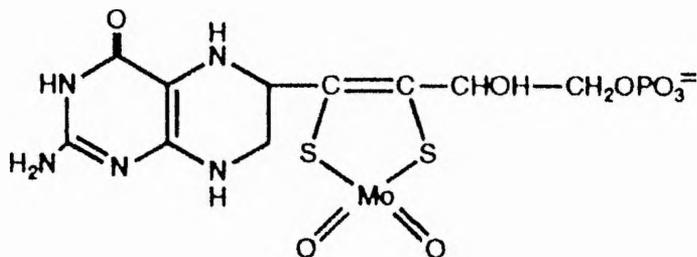
The evidence to date suggests that, with the exception of nitrogenase, all molybdoenzymes contain a basically 'universal' cofactor in which molybdenum is found as part of an organo-metallic structure, termed the molybdenum cofactor.

The extreme lability of the free molybdenum cofactor when released from its protective enzyme environment has denied direct structural characterisation of the molecule. The first insight into the chemical nature of MoCo was derived from the observation that upon aerobic denaturation of molybdoproteins, fluorescence developed which resembled that of oxidised pterins. This led to the proposal that the organic moiety of the cofactor was a novel pterin which was called molybdopterin (Johnson *et al.*, 1980a). Structural studies on four stable degradation products of the active cofactor have inferred the structural characterisation of MoCo. Two of these degradation products, termed Form A and Form B, are stable fluorescent compounds formed *in vitro* by the oxidation of molybdopterin. Form A is produced in the presence of KI and I<sub>2</sub> (by I<sub>2</sub> oxidation) while Form B is produced by air oxidation in the absence of KI and I<sub>2</sub> (Johnson and Rajagopalan, 1982; Johnson *et al.*, 1984). The third degradation product, dicarboxamidomethylmolybdopterin (camMPT), is a stable alkylated derivative of molybdopterin which is produced by a procedure involving treatment with iodoacetamide under mild denaturing conditions (Kramer *et al.*, 1987). Form A, Form B and camMPT were originally isolated from chicken liver sulphite oxidase. The fourth compound used to deduce the structure of MoCo is a metabolic degradation product called urothione which is excreted in human urine (Johnson and Rajagopalan, 1982), and was first isolated more than 50 years ago (Koschara, 1940).

A chemical structure for the active molybdenum cofactor was proposed (Figure 2a) (Kramer *et al.*, 1987), based on the structural elements observed in the different MoCo derivatives. The dioxo form of MoCo which is shown in Figure 2a is characteristic of that found in the majority of molybdoenzymes including sulphite oxidase and nitrate reductase. The molybdenum hydroxylases such as xanthine oxidase however, require a terminal sulphide ligand in place of one of the oxo groups for catalytic activity (Cramer *et al.*, 1981). In addition, the oxidation state of the pterin ring may vary between different molybdoenzymes (Kramer *et al.*, 1987). The cofactor therefore consists of a molybdenum atom attached to a dithiol group on a 4-carbon side chain attached to C-6 of the pterin ring. Other studies using the MoCo degradation products identified in chicken liver sulphite oxidase have confirmed that the same basic MoCo structure (as proposed by Kramer *et al.*, 1987) is found in a wide variety of molybdoenzymes from many different species (Rajagopalan and Johnson, 1992).

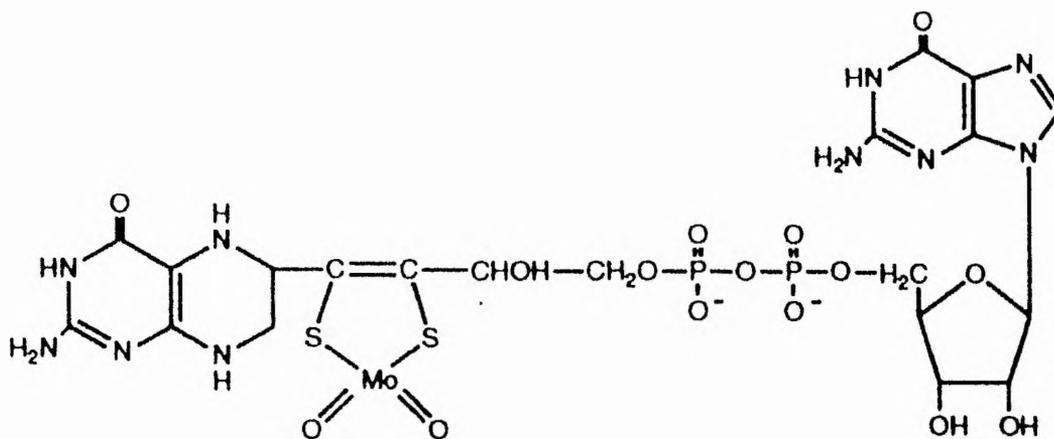
It however became apparent that in several prokaryotes a variant form of the molybdenum cofactor existed. This was initially proposed by Krüger and Meyer (1986), (1987), who reported that carbon monoxide dehydrogenase from *Pseudomonas carboxydoflava* yielded on purification, an inactive pterin, larger than the corresponding derivative from milk xanthine oxidase. Attempts at obtaining camMPT from *Rhodobacter sphaeroides* dimethyl sulphoxide reductase failed to yield that derivative but did however yield a new pterin derivative which was characterised as the alkylated derivative of molybdenum guanine dinucleotide in which molybdopterin is linked to 5'-GMP by a pyrophosphate bond (Johnson *et al.*, 1990a). Figure 2b details the structure of the active molybdopterin guanine dinucleotide form of the molybdenum cofactor. Soon afterwards the pterin in the molybdenum

Figure 2a: Structure of the molybdenum cofactor -molybdopterin form (taken from Rajagopalan and Johnson, 1992 -after Kramer *et al.*, 1987)



**MOLYBDENUM COFACTOR  
(Molybdopterin Form)**

Figure 2b: Structure of the molybdenum cofactor -molybdopterin guanine dinucleotide form (taken from Rajagopalan and Johnson, 1992)



**MOLYBDENUM COFACTOR  
( Molybdopterin Guanine Dinucleotide Form)**

cofactor of carbon monoxide dehydrogenase from *P. carboxydoflava* was found to be molybdopterin cytosine dinucleotide (MCD) (Johnson *et al.*, 1990b). Since then it has been shown that a number of prokaryotic species, but as yet no eukaryotic species, utilise the MGD or MCD form of the molybdenum cofactor (Rajagopalan and Johnson, 1992). Molybdopterin guanine dinucleotide (MGD), molybdopterin adenine dinucleotide (MAD) and molybdopterin hypoxanthine dinucleotide (MHD) forms of the molybdenum cofactor have been identified in formylmethanofuran dehydrogenase from *Methanobacterium thermoautotrophicum* (Börner *et al.*, 1991). It appears that the dinucleotide residue is added on to the 'basic' MPT-only cofactor structure as a final step of MoCo biosynthesis in some prokaryotic species (Johnson *et al.*, 1991a). The specific function of the dinucleotide attachment has not yet been determined. It should be noted that some molybdoenzymes from prokaryotic species utilise the 'unaltered' molybdopterin form of MoCo *e.g.* xanthine dehydrogenase from *Pseudomonas aeruginosa* (Johnson *et al.*, 1991b).

All molybdoenzymes examined to date contain a molybdenum cofactor which can be converted to Form A and Form B derivatives indicating that molybdopterin is the minimum common organic constituent of the universal MoCo. Some thermophilic bacteria appear to use an active form of cofactor with tungsten substituted in place of molybdenum. Formate dehydrogenase from *Clostridium thermoaceticum* is an example of such a species (Yamamoto *et al.*, 1983) however, the purified enzyme still yields fluorescent species indistinguishable from Form A and Form B indicating that the basic molybdopterin structure is still present. More recently, studies on aldehyde ferredoxin oxidoreductase from the hyperthermophilic *Pyrococcus furiosus* and the yet unclassified ES-4, and formaldehyde ferredoxin oxidoreductase from

*P. furiosus* and *Thermococcus litoralis* have revealed the same substitution of tungsten for molybdenum in the active enzymes which can still produce Form A and camMPT (Johnson *et al.*, 1993). The use of these 'tungsto' enzymes may well be limited to thermophilic species and may have evolved to allow survival in such an extreme and unusual environment. Strong chemical similarities exist between tungsten and molybdenum and it appears that the catalytic functions of tungsten may be analogous to that of molybdenum in non-thermophilic organisms.

It appears that in non-thermophilic organisms tungsten can become incorporated into molybdopterin in place of molybdenum to form an inactive cofactor. Rats have been shown to synthesise inactive, demolybdosulphite oxidase when tungsten was administered in conjunction with a low molybdenum diet (Johnson *et al.*, 1974a; 1974b). Approximately 30% of the sulphite oxidase molecules had tungsten incorporated at the molybdenum site with the remaining sulphite oxidase molecules (70%) apparently devoid of cofactor. While inactive, tungsten-containing cofactor has been observed in a number of systems (Johnson, 1980), in only one report has vanadium been implicated as an alternative metal component in an inactive form of the cofactor (Lee *et al.*, 1974). Evidence indicates that under conditions of molybdenum (or other metal) deficiency, organisms will continue to synthesise molybdopterin *i.e.* the organic portion of the molecule without metal (Johnson, 1980).

There is some evidence to suggest that the molybdenum cofactor does not exist free within the cell but is found either incorporated into a molybdoenzyme or in association with a carrier protein (Johnson, 1980). So far, a carrier protein which associates with active MoCo has been identified in several organisms including the prokaryote *Escherichia coli* (Amy and Rajagopalan, 1979) and the

eukaryotes, *Chlamydomonas reinhardtii* (Aguilar *et al.*, 1991) and *Vicia faba* (Kalakoutskii and Fernández, 1996). The carrier protein in *E. coli* is reported to be around 40 kDa and appears to protect the MoCo from inactivation by heat and oxygen (Amy and Rajagopalan, 1979; Hinton and Dean, 1990). In contrast, the 40-50 kDa carrier protein found in *C. reinhardtii* does not seem to protect the MoCo from air inactivation (Aguilar *et al.*, 1992). The 70 kDa carrier protein found in seeds of *V. faba* and the *C. reinhardtii* carrier protein have both been implicated in the direct transfer of MoCo to aponitrate reductase. Free MoCo (unassociated with a carrier protein and not incorporated into a molybdoenzyme) has also been observed in these organisms however this does not necessarily suggest that free MoCo exists *in vivo* because *in vitro* analysis may result in MoCo becoming released from the carrier proteins (Amy and Rajagopalan, 1979; Aguilar *et al.*, 1992). It remains to be seen whether such a MoCo carrier protein is present in *A. nidulans*, or indeed other species, and if so what function it serves.

As well as its catalytic role in various enzyme reactions, the molybdenum cofactor seems to function in maintaining the structural stability of molybdoenzymes. In *A. nidulans* it was shown that wild-type nitrate reductase (which has NADPH: nitrate reductase activity) sedimented at 7.6 S, whereas certain *cnx* mutant nitrate reductase enzymes (which retain only cytochrome *c* reductase activity) sedimented at 4.5 S and other *cnx* mutants contained both the 7.6 S and the 4.5 S species of NR enzymes (MacDonald *et al.*, 1974). *cnxE* mutants were unique in that they produced only the 7.6 S form of mutant NR. From these observations it was postulated that the presence of MoCo allowed the subunits of the native enzyme to become aggregated. It was assumed that only 7.6 S cytochrome *c* reductase was detected in the *cnxE* mutants due to the production of molybdenum-free cofactor precursor ('empty'

cofactor) in these strains which could still function to aggregate the NR monomers. This hypothesis did not explain why both 4.5 S and 7.6 S forms of cytochrome *c* reductase should be found in the other *cnx* strains. A study on xanthine dehydrogenase from *A. nidulans* (Lewis and Scazzocchio, 1977) suggested that the molybdenum cofactor serves to stabilise the dimeric enzyme structure but is not solely responsible for subunit aggregation therefore some 7.6 S dimerised NR could still be formed even in the absence of MoCo (or a Mo-free precursor) in the other *cnx* mutants. It has been suggested that cofactor binding results in slight conformational changes within the NR subunits which leads to strengthened subunit interactions. The role of the molybdenum cofactor in stabilising native quaternary molybdoenzyme structure appears to be universal (Johnson, 1980).

#### **1.4** Enzymes which Require the Molybdenum Cofactor for Activity

##### **1.4.1** Molybdoenzymes Present in *A. nidulans*

Three molybdoenzymes can be found within *A. nidulans*, nitrate reductase, xanthine dehydrogenase (purine hydroxylase I) and purine hydroxylase II. These enzymes all require active MoCo to function.

##### **1.4.1.1** Nitrate Reductase

Nitrate reductase is probably the most widely studied of all molybdoenzymes. The two step enzymatic reduction of nitrate to ammonium, via nitrite, is known as the pathway of nitrate assimilation. In *A. nidulans* the enzyme nitrate reductase (NADPH: nitrate oxidoreductase, EC 1.6.6.3) catalyses the two-electron reduction of nitrate to nitrite. The reduction of nitrite to ammonium is catalysed by the enzyme nitrite reductase (NAD(P)H: nitrite oxidoreductase, EC 1.6.6.4).

The ammonium thus formed is assimilated into glutamate and glutamine catalysed by NADP-linked glutamate dehydrogenase and glutamine synthetase activities. Reduced nitrogen is required by all organisms to synthesise proteins and nucleic acids.

Many bacteria, fungi, algae and plants (animals cannot assimilate nitrate) gain the majority of their reduced nitrogen from the nitrate assimilation pathway however, atmospheric nitrogen can be 'fixed' to produce ammonium by a limited number of bacteria which often form symbiotic relationships with plants. It has been calculated that nitrate assimilation is the major contributor to reduced nitrogen in the biosphere exceeding the total contribution of nitrogen fixation by 100-fold (Guerrero *et al.*, 1981). Nitrates, added as agricultural fertilizers in an aim to increase crop yield are often leached out of the soil and accumulate in surface and ground waters (Campbell and Kinghorn, 1990). This is one of the major causes of eutrophication in water bodies worldwide. Pollution of drinking water with nitrates imposes a health risk as nitrate is reduced in the human gut to toxic compounds. These and other problems associated with the overuse of nitrate fertilizers make nitrate reductase an important enzyme for scientific study. Understanding the structural and regulatory genetics of nitrate reductase and its molybdenum cofactor may help in the engineering of major crop plants which require less fertilizers for optimal yield. The nitrate reductase enzyme may also prove useful in the construction of nitrate determination and elimination devices for dealing with drinking water contaminated with nitrate (Crawford, 1995). An example is a recently developed experimental bioreactor with immobilised nitrate reductase which reduces nitrate to dinitrogen gas using electrical currents (Mellor *et al.*, 1992).

In *A. nidulans* the genes *niaD* and *niiA* encode the apoproteins of the enzymes nitrate reductase and nitrite reductase respectively (Cove

and Pateman, 1963; Pateman *et al.*, 1967). A third gene, *crnA* encodes a permease involved in the uptake of nitrate into the cell (Tomsett and Cove, 1979; Brownlee and Arst, 1983). The three genes are found in a tightly linked cluster on chromosome VIII. The gene order is *crnA-niiA-niaD* as detailed in the fine structure genetic map constructed by Tomsett and Cove (1979). All three genes have been characterised at the nucleotide level (Greaves, 1989; Johnstone *et al.*, 1990; Kinghorn *et al.*, 1990; Unkles *et al.*, 1991).

For reviews which discuss the control of the *A. nidulans crnA*, *niiA* and *niaD* structural genes for nitrate assimilation see Cove, (1979); Scazzocchio and Arst, (1989); Tomsett, (1989) and citations therein. In brief, the *niiA* and *niaD* genes are under strong genetic regulation mediated by two unlinked control genes, *nirA* (located on chromosome VIII) and *areA* (located on chromosome III). *nirA* is involved in nitrate induction of the NiiA and NiaD enzymes and is pathway specific. *areA* is a 'global' control gene which regulates the expression of the nitrate assimilation pathway, and a number of other pathways which utilise alternative nitrogen sources, by ammonium or nitrogen metabolite repression. There is also evidence to suggest that nitrate reductase plays a role in the regulation of its own synthesis as well as that of nitrite reductase (Cove, 1979; Hawker *et al.*, 1992). This autoregulation is thought to operate by a mechanism in which in the absence of nitrate, the functional nitrate reductase molecule is able to bind and inactivate the *nirA* gene product. The phenomenon of NR autoregulation was indicated by the presence of *niaD* and *cnx* mutants with altered regulation. Mutation in the *cnx* or *niaD* genes usually results in the constitutive synthesis of both nitrite reductase and of a mutant nitrate reductase (which can be detected because it still retains cytochrome *c* reductase activity) however, some mutations in the *cnx* genes or the *niaD* gene result in 'normal'

nitrate inducible synthesis of both nitrite reductase and the mutant nitrate reductase (Cove, 1970; Cove and Pateman, 1969; Pateman *et al.*, 1964; Pateman *et al.*, 1967).

A mutation in at least 16 different *A. nidulans* genes can prevent or reduce utilisation of nitrate as a nitrogen source, however only mutations in one of these, the *niaD* gene, can abolish NR activity alone (Cove, 1979). *niaD* mutants were first isolated on the basis of their inability to use nitrate as nitrogen source, while still remaining able to use nitrite and ammonium (Cove and Pateman, 1963; Pateman *et al.*, 1964; Pateman *et al.*, 1967). Another method for the selection of *niaD* mutations (and other mutations which affect NR functionality *e.g.* *cnx* mutations ) is on the basis of chlorate resistance. This is a positive selection method for nitrate non-utilising mutants, based on the finding that cells which lack nitrate reductase are resistant to the toxic analogue of nitrate, chlorate (Åberg, 1947). Chlorate resistance has also been successfully employed to select mutant strains which are defective in nitrate reductase activity (including MoCo mutants) from many other organisms including the bacterium *E. coli* (Miller and Amy, 1983), the algae *Nostoc muscorum* (Singh and Sonie, 1977) and the model plant *Arabidopsis thaliana* (Braaksma and Feenstra, 1973; Wilkinson and Crawford, 1993), however the exact mechanism of chlorate toxicity remains unknown. The simplest explanation is that chlorate is converted to chlorite by nitrate reductase activity and that chlorite is toxic to wild-type cells (Åberg, 1947). This hypothesis is however insufficient to explain the wealth of data generated by Cove for *A. nidulans* (refer to Cove, 1976a, 1976b, 1979 for an extensive discussion) and it would appear that the mechanism of chlorate toxicity is far more complex than that hypothesised by Åberg.

Early studies indicated that the subunit structure of nitrate reductase in *A. nidulans* was difficult to define: Downey and Focht (1974)

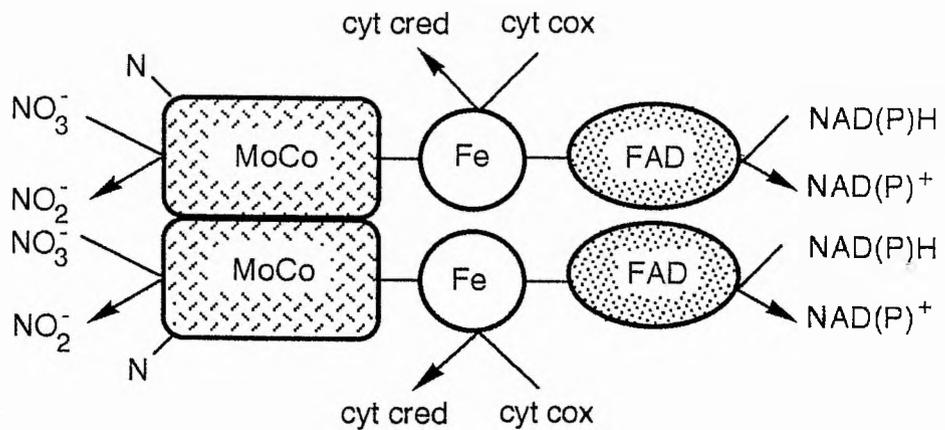
isolated a protomer of 49 kDa and later, Downey and Steiner (1979) identified proteins which had subunit molecular weights of 49, 50 and 75 kDa. Steiner and Downey (1982) identified a single polypeptide with a molecular weight of 54 kDa, suggesting a homo-tetrameric nitrate reductase structure and Minagawa and Yoshimoto (1982) described a 180 kDa nitrate reductase molecule consisting of two 59 kDa subunits and two 38 kDa subunits. Such studies involved extensive purification procedures during which the enzyme was likely to be exposed to proteolysis, thus accounting for the variation in the subunit sizes identified. A later study by Cooley and Tomsett (1985) employed rapid purification in the presence of higher concentrations of protease inhibitors and reported that the native nitrate reductase of *A. nidulans* was a homodimer composed of two 91 kDa subunits with a combined molecular weight of approximately 180 kDa. Molecular cloning and DNA sequencing of the *niaD* gene from *A. nidulans* has revealed that the inferred nitrate reductase apoprotein contains 873 amino acids with a subunit size of 95 kDa (Johnstone *et al.*, 1990), broadly in agreement with the estimation of Cooley and Tomsett (1985). The structure and subunit composition of nitrate reductase has been studied in a wide variety of organisms, both prokaryotic and eukaryotic. The molecular weights of the NR molecules and their subunit sizes and number seem to vary somewhat between different species *e.g.* the NADPH-nitrate reductase enzyme from *Chlorella vulgaris* is believed to be a homotetramer with a native molecular weight of 360-380 kDa consisting of four 90-100 kDa subunits (Solomonson *et al.*, 1986).

Each polypeptide subunit of NR in *A. nidulans* contains three functional domains which bind three cofactors namely FAD (Flavin Adenine Dinucleotide), haem (cytochrome *b557*) and molybdenum pterin (MoCo). Sequence comparisons of the NR subunit polypeptides from a variety of species with functionally-related proteins such as human

erythrocyte NADH-cytochrome *b*<sub>5</sub> reductase (FAD containing protein), human liver cytochrome *b*<sub>5</sub> (haem binding protein) and rat sulphite oxidase (MoCo containing protein) suggested an N to C termini arrangement of the MoCo, haem and FAD domains (Crawford *et al.*, 1988; Kinghorn and Campbell, 1989; Campbell and Kinghorn, 1990). Two 'hinge' interdomain regions were also identified which contain proteolytic cleavage sites (Campbell and Kinghorn, 1990). During the reaction, electrons flow from the physiological electron donor, NADPH, to nitrate through, successively, FAD, haem and MoCo (Meyer *et al.*, 1995). Figure 3 shows a schematic diagram of the NR homodimer, representative of that found in *A. nidulans*. Each functional domain of NR has activity independent of the other domains, for example reduction of the non-physiological electron acceptor cytochrome *c* only requires the haem and FAD containing parts of the enzyme (this activity is shown in Figure 3). NO<sub>3</sub><sup>-</sup> reduction can still be achieved by only the haem and MoCo domains, using reduced methyl viologen (MVH) as artificial electron donor or by the MoCo domain alone, using bromophenol blue (BPB) to artificially donate electrons.

Many eukaryotic NR amino acid sequences have been determined and are available from diverse organism such as plants, algae and fungi (Rouzé and Caboche, 1992; Solomonson and Barber, 1990). Several studies on NR defective mutants have pinpointed single amino acid residues within each of the MoCo, haem and FAD domains of the NR protein which are thought to have a critical role in the functionality of the enzyme (Okamoto and Marzluf, 1993; Garde *et al.*, 1995; Meyer *et al.*, 1995 and references therein). The amino acid residues which appear to be critical for the MoCo domain to function are considered here. In *Arabidopsis thaliana*, two point mutations in the MoCo domain of the NR gene have been identified by sequencing and indicate alanine 192 and

Figure 3: Schematic diagram of the nitrate reductase homodimer  
(adapted from Crawford, 1995).



FAD, flavin domain; Fe, haem domain; MoCo, molybdenum cofactor domain; N, N terminus of nitrate reductase; cyt c, cytochrome c, which can be used as an alternative electron acceptor *in vitro*, cyt cred and cyt cox, the reduced and oxidised forms of cytochrome c respectively.

glycine 308 as critical residues (Wilkinson and Crawford, 1993; Braaksma and Feenstra, 1982; LaBrie and Crawford, 1994). In 1995, Meyer and colleagues published results of the sequence analysis of four nitrate reductase mutants from *Nicotina plumbaginifolia* which were presumed to be defective in the NR MoCo domain. In each of the four mutants, a single base change was revealed when the MoCo domain sequences were compared with the wild-type sequence, suggesting that these four residues are critical for NR MoCo domain functionality in *N. plumbaginifolia*. The four critical residues were identified as glycine 293, arginine 298, tryptophan 459 and a second glycine, glycine 463. Each of these mutated residues is conserved among all known plant NR sequences as well as algal and fungal sequences (with the exception of glycine 463 which is replaced by a serine residue in the fungus *Ustilago maydis*) and the same residues are also conserved in the MoCo domain of sulphite oxidases. It was suggested that these mutations may affect the quaternary structure of the NR molecule (Meyer *et al.*, 1995). In *A. nidulans*, site-directed mutagenesis was used to substitute alanine for cysteine 150, which is the only totally conserved cysteine residue found in the MoCo domains of both NR and sulphite oxidase (Garde *et al.*, 1995). This mutation was found to abolish NR activity as well as MoCo domain partial activities (measured using the non-physiological electron donors MVH and BPB). It was suggested that cysteine 150 may be involved in the attachment of MoCo to the NR molecule or may have a role in NR subunit dimerisation.

#### 1.4.1.2 Xanthine Dehydrogenase (Purine Hydroxylase I) and Purine Hydroxylase II

Xanthine dehydrogenase/purine hydroxylase I (EC 1.2.1.37, Xanthine: NAD<sup>+</sup> oxidoreductase, purine hydroxylase I) and purine

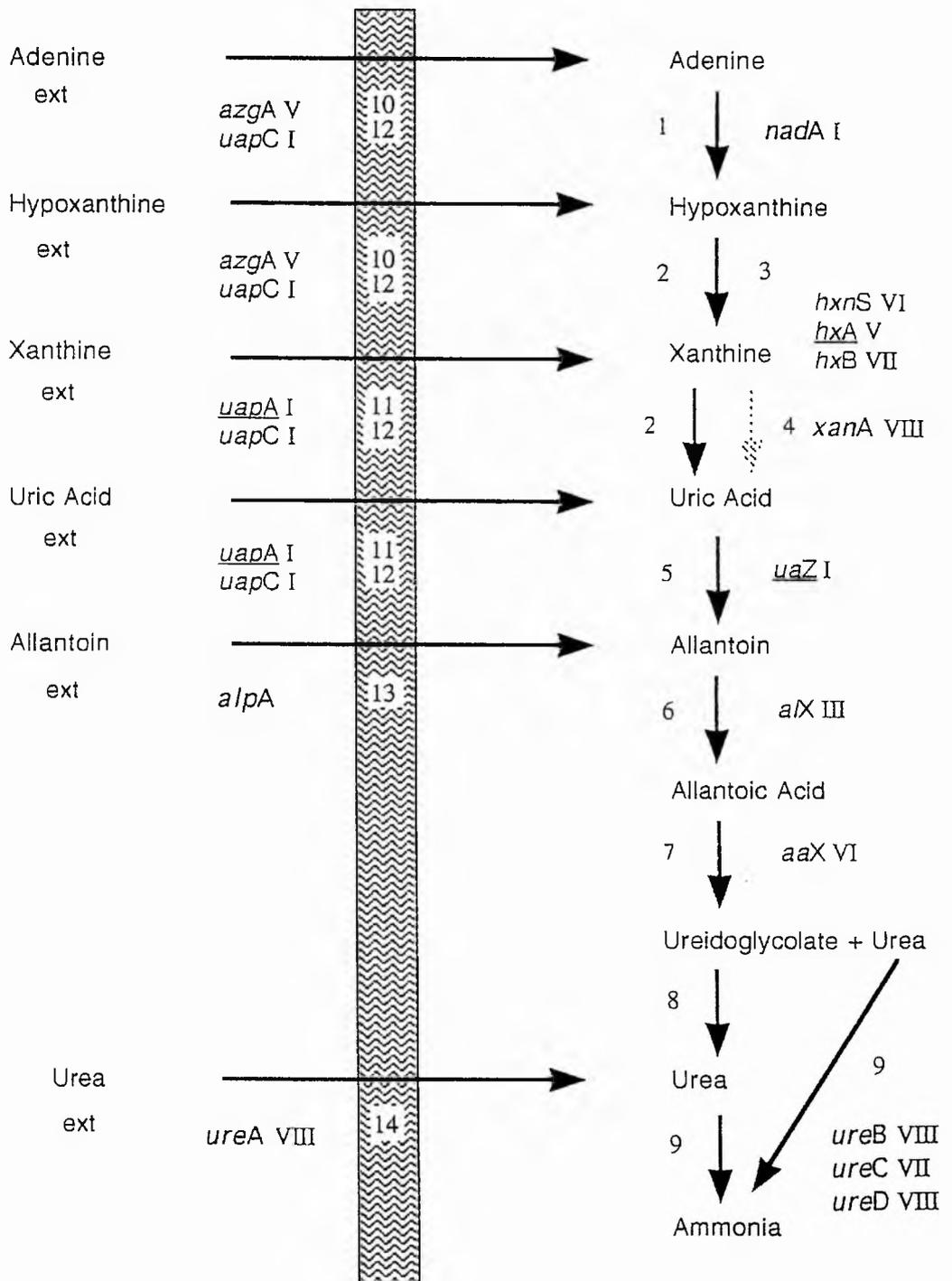
hydroxylase II are enzymes involved in the *A. nidulans* purine degradation pathway. *Via* this pathway the purines adenine, guanine and hypoxanthine, xanthine, uric acid and their metabolites, allantoin, allantoic acid and urea, can be broken down and used as sole nitrogen sources (however not as sole carbon sources). The nitrogen atoms of the purine ring become reduced to ammonium while the carbon gives rise to four molecules of CO<sub>2</sub> and one molecule of glyoxylate (Vogels and van der Drift, 1976). The purine degradation pathway is shown in Figure 4. The genes involved in purine degradation are scattered throughout the *A. nidulans* genome, the only exception being *ureA* and *ureB*, which are closely linked on chromosome VIII.

The xanthine dehydrogenases belong to a group of related enzymes (sometimes referred to as the molybdenum hydroxylases) including aldehyde oxidases and xanthine oxidases, which exhibit the same overall general structure but have different substrate specificities. Scazzocchio and colleagues (1973) reported that two xanthine dehydrogenases (I and II) were present in *A. nidulans*. The enzymes have since been renamed purine hydroxylase I (also still referred to as xanthine dehydrogenase) and purine hydroxylase II. The two enzymes show partial identity in double diffusion plates and immunoelectrophoresis however differ in their induction, repression, substrate and inhibitor specificities (Scazzocchio *et al.*, 1973; Scazzocchio, 1974). Xanthine dehydrogenase (PHI) catalyses the conversion of hypoxanthine and xanthine to uric acid. Purine hydroxylase II (PHII) is physiologically a nicotinate hydroxylase (Scazzocchio *et al.*, 1973) and accepts nicotinate as substrate in the nicotinate degradation pathway catalysing the conversion of nicotinate to 6-hydroxynicotinate. PHII can however also accept hypoxanthine, but not xanthine, as a substrate during purine degradation (Scazzocchio, 1974).

Figure 4: The purine degradation pathway in *A. nidulans* (adapted from Scazzocchio, 1994)

1, adenine deaminase; 2, xanthine (and hypoxanthine) dehydrogenase (purine hydroxylase I); 3, purine hydroxylase II; 4, xanthine alternative pathway; 5, urate oxidase; 6, allantoinase; 7, allantoicase; 8, ureidoglycolase; 9, urease; 10, narrow specificity purine permease (adenine, hypoxanthine and also guanine); 11, narrow specificity purine permease (xanthine and uric acid); 12, wide specificity purine permease; 13, allantoin permease; 14, urea permease.

The genes corresponding to each step are marked along with the chromosomes where they map (the *alpA* gene has not been mapped to a chromosome). The genes which are underlined have been shown formally to code for the cognate enzyme or permease and have been cloned and sequenced. The dashed arrow indicates that the xanthine alternative pathway is only known from mutational evidence. The presence of guanine or guanosine deaminase has not been investigated and is not indicated on the figure.



The gene *hxA*, which encodes the xanthine dehydrogenase (PHI) apoprotein has been cloned and sequenced (Glatigny and Scazzocchio, 1995). Mutations in *hxA* were originally identified by their inability to utilise hypoxanthine as sole nitrogen source (Darlington *et al.*, 1965; Alderson and Scazzocchio, 1967; Darlington and Scazzocchio, 1967). The *hxA* gene (located on chromosome V) specifies a polypeptide with a predicted molecular weight of 149.41 kD, which compares well with the earlier estimate of 304 kD for the xanthine dehydrogenase dimer (Lewis *et al.*, 1978). Xanthine dehydrogenases (and related enzymes such as aldehyde oxidases) from a variety of species including *A. nidulans* appear to be dimers, with each monomer of approximately 150 kDa comprising four domains. Starting at the N-terminus the xanthine dehydrogenase monomer contains two small domains which each bind a single 2Fe:2S cluster followed by a flavin binding domain and completed (at the C-terminal) by a MoCo binding domain (Kisker *et al.*, 1997). The *A. nidulans* xanthine dehydrogenase, in common with related enzymes, uses NAD as the terminal electron acceptor and *in vitro* is capable of transferring electrons from NADH to artificial electron acceptors such as a tetrazolium salt (*e.g.* 2-p-iodophenyl-3-(p-nitrophenyl-5-phenyl-tetrazolium chloride). This NADH-tetrazolium reductase: NADH dehydrogenase ancillary activity is maintained in *cnx* and *hxB* mutants (see below) (Scazzocchio, 1980). The electron flow is from substrate to NAD through, successively, MoCo, FAD and the two 2Fe:2S domains.

Another gene, *hxB* (located on chromosome VII) codes for a protein needed for the substrate-specific hydroxylation activity of both purine hydroxylases (I and II), but not for the ancillary NADH dehydrogenase activity common to these enzymes (Scazzocchio *et al.*, 1973; Scazzocchio, 1980). Mutations in *hxB* do not affect nitrate reductase activity. The *hxB* gene seems to code for a protein involved in

the post-synthesis modification of the molybdenum cofactor which is necessary for both purine hydroxylase, but not nitrate reductase, activities. It has been suggested that *hxB* may encode an enzyme involved in the generation of the molybdenum bound -S<sup>-</sup> group (which is typical of the molybdenum hydroxylases) in a manner analogous to the *ma-l* gene of *Drosophila melanogaster* (Wahl *et al.*, 1982).

The gene *hxnS* (located on chromosome VI) codes for the purine hydroxylase II apoenzyme (Scazzocchio, 1980). Mutations in *hxnS* can prevent growth on hypoxanthine as nitrogen source in the presence of 1.8  $\mu$ M allopurinol (a complete inhibitor of xanthine dehydrogenase) and 1 mM nicotinate (which at this concentration cannot be used as a nitrogen source however serves to induce PHII) or on 10 mM nicotinate as sole nitrogen source (Scazzocchio, 1973). Purine hydroxylase II has been purified and has an estimated molecular weight of 300 kDa, consisting of two 153 kDa subunits which each contain an FAD, MoCo and two 2Fe:2S domains (Mehra and Couglan, 1984), as in xanthine dehydrogenase (PHI). PHII also uses NAD as terminal electron acceptor and exhibits transfer of electrons from NADH to artificial electron acceptors *in vitro*. It has been suggested that when xanthine dehydrogenase activity is inhibited and purine hydroxylase II activity is induced (by nicotinate), the product of the *xanA* gene can provide the so-called 'xanthine alternative pathway' for the conversion of xanthine to uric acid, however only mutational evidence for the existence of this pathway is so far available (Darlington and Scazzocchio, 1968; Scazzocchio, 1994).

The positive acting regulatory gene *uaY* (located on chromosome VIII) mediates uric acid induction of several enzymes involved in purine catabolism including xanthine dehydrogenase (Darlington *et al.*, 1965; Darlington and Scazzocchio, 1967, 1968; Scazzocchio *et al.*, 1973). The *hxnS* gene (encoding PHII) is controlled, in

response to nicotinate/6-hydroxynicotinate availability, by *hxnR* (a positive acting regulatory gene) and *aplA* (a negative acting regulatory gene) and these three genes (*hxnS*, *hxnR* and *aplA*) form a cluster on chromosome VI (Scazzocchio, 1994). Indirect data indicate that the xanthine alternative pathway is also inducible by uric acid (Sealy-Lewis *et al.*, 1978) and that the expression of the *hxB* gene can be induced by uric acid (mediated by the *uaY* gene product) but also by nicotinate or 6-hydroxynicotinate (mediated by the *hxnR-aplA* gene products) (Scazzocchio *et al.*, 1973; Scazzocchio *et al.*, 1982). All the genes involved in purine degradation are under the control of the 'global' nitrogen regulatory gene *areA*, which mediates ammonium and glutamate repression of the pathway.

#### 1.4.2 Other Molybdoenzymes

The molybdenum cofactor is required by a diverse range of molybdoenzymes, which catalyse basic reactions in the metabolism of nitrogen, sulphur and carbon. MoCo-containing enzymes all catalyse the transfer of an oxygen atom, ultimately derived from or incorporated into water, to or from a substrate in a two-electron redox reaction. Enzymes which contain MoCo (including the recently identified enzymes which utilise a tungsten form of the MoCo) have been classified into four recognised families: (1) the DMSO (dimethylsulphoxide) reductase family which includes DMSO reductase (*Rhodobacter sphaeroides*), biotin sulphoxide reductase (*Escherichia coli*) and formate dehydrogenase (*Clostridium thermoaceticum*); (2) the xanthine oxidase family which includes xanthine oxidase (*Bos taurus*), xanthine dehydrogenase (*Homo sapiens*) and aldehyde oxidoreductase (*Desulphovibrio gigas*); (3) the sulphite oxidase family which includes sulphite oxidase (*Rattus norvegicus*) and assimilatory nitrate reductase (*Aspergillus nidulans*); (4)

the aldehyde ferredoxin oxidoreductase family which includes aldehyde ferredoxin oxidoreductase (*Pyrococcus furiosus*) and hydrocarboxylate viologen oxidoreductase (*Proteus vulgaris*). For a detailed description of the four currently recognised classes of MoCo-containing enzymes refer to Kisker *et al.*, (1997).

A comparison of several types of molybdoenzyme from different species has allowed a 'consensus' sequence for the molybdenum cofactor binding domain to be suggested (Glatigny and Scazzocchio, 1995). The putative consensus sequence is shown in Figure 5. Ten nitrate reductase sequences (from *A. nidulans*, *Fusarium oxisporum*, *Neurospora crassa*, *Ustilago maydis*, *Volvox carteri*, *Arabidopsis thaliana*, barley, *Nicotinum tobacum*, rice and spinach), eight xanthine dehydrogenase sequences (from *A. nidulans*, *Drosophila melanogaster*, *Drosophila pseudoobsura*, *Calliphora vicina*, rat liver, mouse liver, human liver H1 and human liver H2), two sulphite oxidase sequences (rat liver and chicken liver), the aldehyde oxidoreductase of *Desulphovibrio gigas*, the aldehyde dehydrogenase of *Acetobacter polyoxogenes* and the nicotinate dehydrogenase of *Arthrobacter nicotinovorans* were aligned and compared. It appears that in the different molybdoenzymes several amino acids are conserved indicating an important role for these residues in MoCo binding. Only one residue which is conserved within this model, namely cysteine 150, has been investigated in *A. nidulans* with site-directed mutagenesis of this amino acid resulting in abolition of NR activity as well as MoCo domain partial activities (Garde *et al.*, 1995). Sequence analysis of NR mutants has revealed several single amino acid residues within the MoCo domain which are thought to have a critical role in enzyme functionality (refer to section 1.4.1.1 and references therein), however none of these implicated residues feature in the putative consensus sequence displayed in Figure 5.

Figure 5: Putative consensus sequence for the molybdenum cofactor binding domain (taken from Glatigny and Scazzocchio, 1995)

S.O., the consensus sequence for the sulphite oxidases (corresponding to positions 137-214 of the rat enzyme); N.R., the consensus sequence for the nitrate reductases (positions 72-157 of *A. nidulans*); XHD, the consensus sequence for the xanthine dehydrogenases (positions 773-862 of *A. nidulans*); Nic.D., the nicotine dehydrogenase sequence (positions 203-288); Ald.D., the aldehyde dehydrogenase sequence (positions 368-452); Ald.O., the aldehyde oxidoreductase sequence (positions 369-446); cons., the "consensus of consensus" for all eukaryotic sequences;  $\varphi$ , hydrophobic amino acids;  $\gamma$ , asparagine or glutamine (an exception is found in the *U. maydis* sequence);  $\beta$ , histidine or asparagine or glutamine;  $\alpha$ , aspartic or glutamic acid.

The double boxes represent the universally conserved amino acids; the single boxes represent the amino acids where there are conservative substitutions; and the numbers represent the the residues between conserved amino acids. For 9 out of the 10 nitrate reductases, the sixth spacing between conserved residues (between  $\beta$  and G) is of 16 or 19 amino acids; the exception is the nitrate reductase from *N. crassa*, which shows a spacing of 30 amino acids. Residues marked with an asterisk are those where the human liver H2 enzyme differs from other, conventional xanthine dehydrogenases.

S.O.	S	3	F	2	E	5	L	4	I	9	N	H	20	G	21	L	1	C	3	R	2	E
									L						or							
															22							
N.R.	G	3	F	2	E	5	L	5	I	9	N	H	16	G	24	L	1	C	3	R	2	E
			L						L		S	Q	or		or	F						
													19		27							
XDH	G	3	F	2	E	4	I	10	I	4	Q	H	28	G	25	V	1	C	3	R	1	E
							L	or	L		N			A*	I						or	
							V	11	V		F*										2	D
cons.	G		F		E		$\phi$		$\phi$		$\gamma$	$\beta$		G		$\phi$		C		R		$\alpha$
Nic.D.					E	4	L	11	L	4	Q		29	G	24	V	1	W	3	R	1	E
Ald.D.					E	4	L	9	I	4	Q		29	G	24	V	1	L	3	R	2	D
Ald.O.					E	4	F	10	I	4	I		28	G	24	V	1	L				

The wide distribution of molybdoenzymes in organisms ranging from bacteria to humans and the diverse array of reactions they catalyse highlights the importance of the MoCo which in all systems studied maintains the same basic structure.

## **1.5 The Genetics of Molybdenum Cofactor Biosynthesis in Other Organisms**

Although mutants lacking functional MoCo activity have been identified in a large range of species, the genes responsible for MoCo biosynthesis have so far only been isolated and cloned in a few organisms. As yet, no fungal MoCo biosynthesis genes have been sequenced. The cloning of MoCo biosynthesis genes in lower eukaryotes (*e.g.* fungi) may reveal important information about the evolution of the pathway when compared with prokaryotic and higher eukaryotic systems.

### **1.5.1 Other Molybdenum Cofactor Mutants Including Human Molybdenum Cofactor Deficiency**

Molybdenum cofactor-deficient mutants, equivalent to the *cnx* mutants of *A. nidulans*, have been identified in many other organisms including: *Escherichia coli* (Haddock and Jones, 1977; Bachmann, 1983; Miller and Amy, 1983); *Salmonella typhimurium* (Stouthamer, 1969; Sanderson and Hartman, 1978); *Nostoc muscorum* (Bagchi and Singh, 1984); *Nicotiana tabacum* (Buchanan and Wray, 1982) and *Hordeum vulgare* (Bright *et al.*, 1983); *Drosophila melanogaster* (Warner and Finnerty, 1981) and *Homo sapiens* (Johnson *et al.*, 1980b; Wadman *et al.*, 1983; Slot *et al.*, 1993).

Molybdenum cofactor deficiency in humans was first identified in 1978 (Duran *et al.*, 1978). By 1989, more than 20 patients suffering from this devastating condition had been noted (Johnson *et al.*, 1989) and

since then many more cases have been described. Individuals affected by MoCo deficiency display the symptoms of a combined loss of sulphite oxidase and xanthine dehydrogenase activity due to a genetic defect leading to incomplete biosynthesis of the molybdenum cofactor shared by these molybdoenzymes (Arnold, 1993; Bonioli, 1996). A third human molybdoenzyme, aldehyde oxidase, is also affected by MoCo deficiency however no symptoms specifically attributable to a loss of aldehyde oxidase activity have been identified.

The condition is inherited as an autosomal recessive trait (Johnson, 1989) and antenatal diagnosis can be performed by an assay of sulphite oxidase activity in a chorionic villus biopsy sample (Desjacques *et al.*, 1985) or in a culture of amniotic cells (Ogier *et al.*, 1983). Clinical symptoms include: feeding difficulties; mental retardation; abnormal physiognomy; severe neurological abnormalities including seizures; myclonia; an abnormal muscle tone and dislocation of the ocular lens (Wadman, 1983; Mize, 1995; Lueder and Steiner, 1995). The biochemical hallmarks of MoCo deficiency include a marked decrease in the excretion of uric acid and sulphate; abnormally high excretion of hypoxanthine, xanthine, sulphite, thiosulphate and S-sulphocysteine in the urine and the absence of urothione in urine samples (Bamforth *et al.*, 1990; van Gennip, 1994; Yurdakök and Coskun, 1997).

Although molybdenum cofactor deficiency is rare, most afflicted patients fail to survive beyond the first few years of life. So far attempts at treatment have resulted in only mild alleviation of symptoms (Boles *et al.*, 1993; Kurlemann *et al.*, 1996). A better understanding of the genes involved in MoCo biosynthesis may allow a more enlightened therapeutic approach to the problems of individuals affected by this genetic disease. If the genes specifying MoCo biosynthesis are isolated from several different species it may become possible to use PCR (with

degenerate primers deduced from conserved areas of the predicted protein sequences) to clone the human counterparts.

### 1.5.2 The Genetics of Molybdenum Cofactor Biosynthesis in *Escherichia coli*

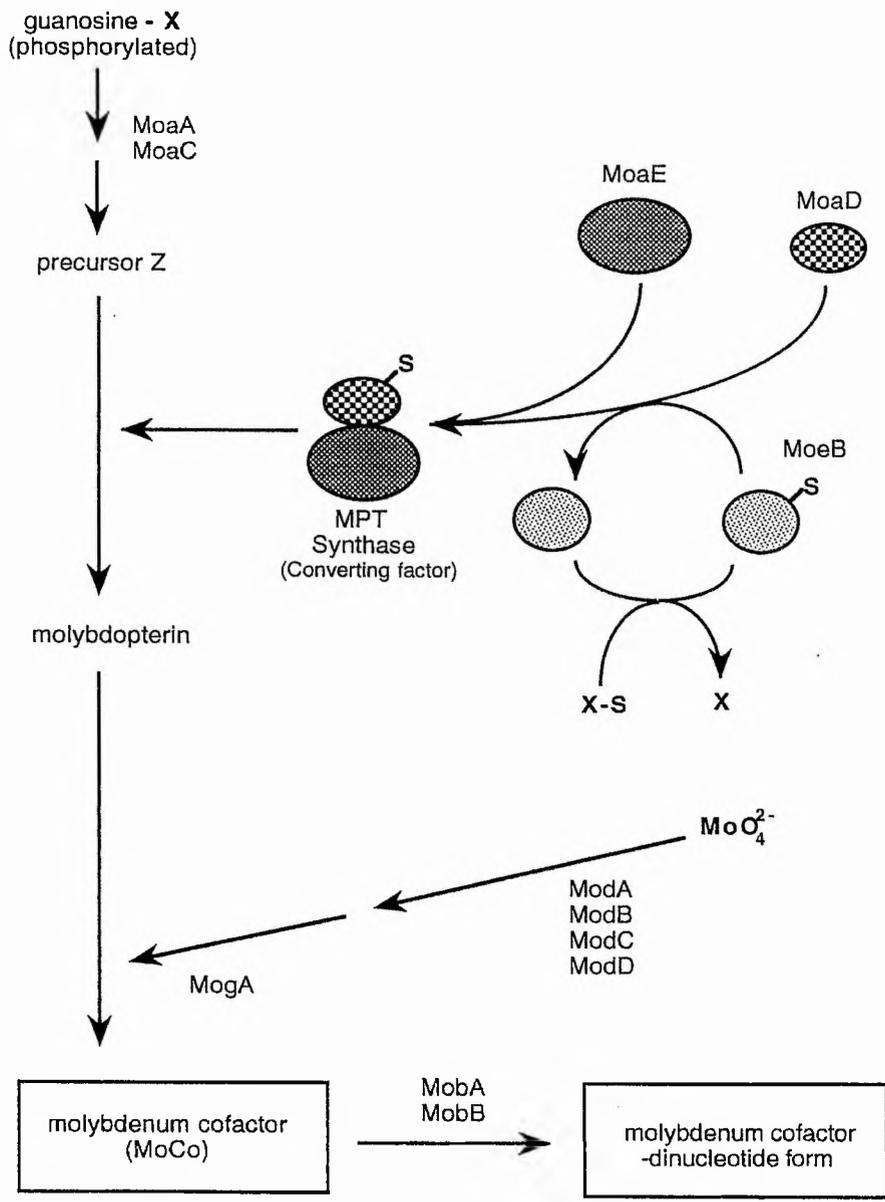
The molecular genetics of molybdenum cofactor biosynthesis has been most extensively studied in *Escherichia coli*. The work on *E. coli* has been facilitated by its relatively small genome size and the grouping together of genes with similar functions into operons. In *E. coli*, mutants which are pleiotropically defective in molybdoenzyme activities have been selected on the basis of chlorate resistance (Miller and Amy, 1983; Baker and Boxer, 1991). These mutants were originally designated *chl* mutants but have recently been redesignated *mo* (Shanmugam *et al.*, 1992). Figure 6 depicts the proposed model for MoCo biosynthesis in *E. coli* (adapted from Stallmeyer *et al.*, 1995). This first model was based on results of the biochemical analysis of the different *mo* mutants.

The products of at least five separate loci are thought to be required for the active expression of molybdoenzymes within the *E. coli* cell. These five loci are: *moa*; *mob*; *mod*; *moe* and *mog* (formerly termed *chlA*, B, D, E and G respectively). Each of the corresponding mutants are unable to synthesise functional molybdenum cofactor with the exception of *mob* (Miller and Amy, 1983).

*mob* mutants produce molybdenum cofactor which is functional when tested in the *nit-1* complementation assay, but lack the ability to carry out a maturation step necessary to form active MoCo in *E. coli*. The *mob* locus is involved in the addition of a GMP moiety to molybdopterin to produce molybdopterin guanine dinucleotide (MGD) which is the functional form of MoCo in *E. coli* (Johnson *et al.*, 1991a). This late step

Figure 6: Representative model of MoCo biosynthesis in *E. coli*  
(adapted from Stallmeyer *et al.*, 1995)

The gene products of *moaA* and *moaC* are involved in the conversion of a phosphorylated guanosine residue into a molybdopterin precursor, namely precursor Z. Precursor Z is transformed into molybdopterin by the enzyme MPT synthase (also known as the converting factor), which is composed of the products of the *moaD* and *moaE* loci. MPT synthase is activated by a sulphotransferase (encoded by *moeB*). Molybdenum, which is transported into the *E. coli* cell by proteins encoded by the *mod* operon (*modABCD*), is then incorporated into molybdopterin possibly by the *mogA* gene product to form the molybdenum cofactor molecule. The *mob* protein products then function in the attachment of a dinucleotide residue to the molybdeum cofactor.



appears to be confined to prokaryotes (Krüger and Meyer, 1987; Rajagopalan and Johnson, 1992)

A region of the *E. coli* chromosome including the *mob* locus was sequenced by Plunkett *et al.*, 1993, as part of the *E. coli* genome sequencing project. As a result two genes were identified at this locus which have been termed *mobA* and *mobB*. The discovery of a second *mob* gene, *mobB*, was unexpected as all available *mob* mutants can be complemented by plasmids carrying only the *mobA* gene (Iobbi-Nivol *et al.*, 1995). The *mobA* gene encodes protein FA (protein factor d'association) which is able to restore molybdoenzyme activities to cell extracts of *mob* mutants. Protein FA has been purified to homogeneity (Palmer *et al.*, 1994). It was found that activation of purified, inactive *mob* nitrate reductase could not be achieved in the presence of protein FA and GTP alone. The addition of a further protein fraction, factor X, was required in order to activate the mutant nitrate reductase (Santini *et al.*, 1992). It has been suggested that the product of the *mobB* gene may contribute to the activity of factor X (Iobbi-Nivol *et al.*, 1995). The existence of a putative nucleotide-binding motif in the predicted amino acid sequence of *mobB* suggested a role in the GTP-dependent synthesis of MGD from molybdopterin (Iobbi-Nivol *et al.*, 1995). Recent purification of the MobB protein has confirmed the function of this protein as a nucleotide binding protein with high affinity for GTP and has shown that MobB enhances the *in vitro* activation of nitrate reductase in a concentration-dependent manner (Eaves *et al.*, 1997). Expression of the *mob* locus is very low and appears to be constitutive (Iobbi-Nivol *et al.*, 1995).

The *moa* locus has been sequenced and found to contain an operon with five cistrons (*moaA-moaE*) each encoding a protein (MoaA-MoaE) (Rivers *et al.*, 1993). Mutant strains carrying lesions

corresponding to all of the individual *moa* genes have been generated except in *moaB* (Rivers *et al.*, 1993). The *moaA* and *moaC* protein products are thought to be involved in the conversion of a poorly defined phosphorylated guanosine derivative into a molybdopterin intermediate called precursor Z (Rivers *et al.*, 1993; Wuebbens and Rajagopalan, 1993, 1995). The precursor Z molecule is in turn converted to molybdopterin by a so called 'converting factor' designated MPT synthase which is composed of two protein subunits, the products of the *moaD* and *moaE* genes, with molecular masses of 10 kDa and 16 kDa respectively (Pitterle and Rajagopalan, 1993). Expression studies using gene fusions have indicated that the *moa* locus is subject to repression by the molybdenum cofactor (Baker and Boxer, 1991). The overall control of MoCo biosynthesis in *E. coli* is possibly mediated through feedback regulation of the *moa* operon by the end product of the pathway (active MoCo).

Sequencing of the *moe* locus has revealed two open reading frames namely *moeA* and *moeB* transcribed from a common promoter (Nohno *et al.*, 1988). The *moeB* gene encodes an enzyme involved in the transfer of sulphur to the 10 kDa subunit of MPT synthase (Pitterle and Rajagopalan, 1993; Pitterle *et al.*, 1993). The small subunit of MPT synthase (encoded by *moaD*) requires activation by the addition of a single reactive sulphur before the enzyme can function (Pitterle *et al.*, 1993). Two sulphur atoms are transferred from the small subunit of the converting factor to precursor Z to generate the dithiolene group unique to molybdopterin. The stoichiometry of this sulphur transfer has not yet been established. It is the dithiolene moiety within molybdopterin that serves to bind molybdenum. The reactions of sulphur transfer during molybdopterin biosynthesis are thought to compose a novel cycle (Pitterle *et al.*, 1993). The initial donor of sulphur to the *moeB* protein is unclear. As yet, a function for the *moeA* gene product has not been determined

however it has recently been suggested that MoeA is responsible for generating a form of 'activated' molybdenum produced by the combination of sulphide with molybdate (Hasona *et al.*, 1998).

The *mod* locus was found to consist of four separate genes in an operon (Johann and Hinton, 1987; Walkenhorst *et al.*, 1995; Maupin-Furlow *et al.*, 1995). These genes (*modA*, *B*, *C* and *D*) collectively encode a high-affinity molybdate uptake and transport system (Johann and Hinton, 1987; Miller *et al.*, 1987). It has been suggested that the *modA* gene product, with a signal peptide at the N-terminus, is probably the periplasmic binding protein for molybdate. It is also probable that *ModB* is the inner membrane protein and the product of *modC* is the ATP-binding protein of the molybdate transport system. The possible function of the *modD* gene has not yet been predicted.

A further operon, containing two genes, has been identified immediately upstream of the *modABCD* operon (Walkenhorst *et al.*, 1995; Grunden *et al.*, 1996). The two genes in this operon namely *modE* and *modF* encode proteins with predicted molecular weights of 28 kDa and 54 kDa respectively. Expression of the *modABCD* operon is repressed by *ModE* in a molybdate-dependent fashion (Rech *et al.*, 1995; Walkenhorst *et al.*, 1995; Grunden *et al.*, 1996). *ModE* has been shown to bind to the operator/promoter region of the *modABCD* operon, when bound to molybdate (Grunden *et al.*, 1996; McNicholas *et al.*, 1997). It has also been shown that *ModE* functions as a modest activator of *moaABCDE* expression by binding to the promoter region of this operon in a manner independent of molybdate availability (McNicholas *et al.*, 1997). The role of *ModF* has not yet been determined.

The *mog* locus has been cloned (Reiss *et al.*, 1987) and sequenced (Yura *et al.*, 1992; James *et al.*, 1993) and found to contain a single open reading frame encoding a relatively small protein of 21 kDa.

This gene is alternatively termed *mogA* even although no other ORF has been identified at this locus. The amino acid sequence of the MogA protein suggests that it is not a membrane protein but gives no other clues as to its function (Hinton and Dean, 1990). It has been suggested that the *mogA* gene product may be involved in intracellular molybdate processing or the insertion of molybdate into the cofactor because the *mogA* mutant phenotype seems to be repairable by the addition of molybdate to the growth medium (Stewart and MacGregor, 1982; Miller and Amy, 1983).

Both *mod* and *mog* mutants have been reported to respond to elevated levels of molybdate in the culture medium, however the two types of mutant differ in the extent of phenotypic repair. While molybdoenzymes in *mod* mutants are restored to wild-type levels by the addition of 0.1 mM molybdate to the growth medium (Glaser and DeMoss, 1971), *mog* mutants display much lower increases in nitrate reductase (Stewart and MacGregor, 1982) and biotin sulfoxide reductase (del Campillo-Campbell and Campbell, 1982) levels, even at nearly toxic concentrations of molybdate (*i.e.* 10 mM).

Somewhat conflicting evidence is available from the studies of Miller and Amy (1983) who demonstrated that wild-type levels of molybdopterin (demolybdo-cofactor) and molybdenum-containing cofactor were present in *mog* and *mod* mutants when grown in 1 mM molybdate. They also showed that *mog* mutant strains grown in 1  $\mu$ M molybdate produced near wild-type levels of molybdopterin (demolybdo cofactor) but were unable to synthesise active molybdenum-containing cofactor, while *mod* strains had only 15-20% of the wild-type levels of both cofactor forms when grown with 1  $\mu$ M added molybdate. This led to the suggestion that the protein encoded by the *mog* gene was involved in the insertion of molybdenum into molybdopterin during MoCo

biosynthesis. It was presumed that under high molybdate levels alternative processes, perhaps non-enzymatic were sufficient for the supply and incorporation of molybdenum into the cofactor. A more recent study showed that addition of 1 mM molybdate increased the production of MoCo to wild-type levels in a *mod* mutant strain but did not significantly alter the level of MoCo found in a *mog* mutant strain (Joshi *et al.*, 1996), which is perhaps more in accord with the observed levels of phenotypic repair. On the basis of this study it was suggested that the MogA protein may function as a molybdochelate responsible for specific intracellular molybdenum binding and processing for MoCo assembly.

### 1.5.3 Molybdenum Cofactor Biosynthesis Genes Isolated from Other Prokaryotes

The molecular genetics of MoCo biosynthesis is far less well characterised in other prokaryotes. The function of the genes which have been isolated has been predicted by their similarities to previously isolated *E. coli* protein sequences. Some examples of MoCo biosynthesis genes which have been cloned from other prokaryotes are given here, however this is by no means a complete overview.

A gene which is similar to the MoaA protein from *E. coli* and which was able to complement *E. coli moaA* mutant strains has been isolated in *Arthrobacter nicotivorans* (Menendez *et al.*, 1995). Purification of MoaA from *A. nicotivorans* has revealed an Fe-S protein which contains two conserved cysteine clusters each of which is essential for MoaA to function (Menendez *et al.*, 1996). The *mob* locus from *Rhodobacter sphaeroides* has recently been characterised (Palmer *et al.*, 1998). Two open reading frames (*mobA* and *mobB*) were identified, each exhibiting amino acid sequence similarity to their *E. coli* counterparts. The *mobA* ORF from *R. sphaeroides* was sufficient to complement an

*E. coli mob* mutant strain. One difference between the two organisms is that the *R. sphaeroides* MobB protein appears to be some 277 amino acids longer than the *E. coli* equivalent and it has been suggested that this extra domain may be uniquely required for MoCo biosynthesis in *R. sphaeroides*. Operons encoding proteins with similarity to the *E. coli mod* (molybdate transport) proteins have been identified in a number of other prokaryotes including *Azotobacter vinelandii* and *Rhodobacter capsulatus* (Luque *et al.*, 1993; Wang *et al.*, 1993). It appears that these *mod* operons are also regulated in a similar manner to that found in *E. coli* (Wang *et al.*, 1993; Mouncey *et al.*, 1993; Grunden and Shanmugan, 1997).

A gene has been identified in the cyanobacterium *Anabaena* (strain PCC 7120) whose protein sequence displays similarity to MoeA from *E. coli* (Ramaswamy *et al.*, 1996). It was concluded that mutation of the *moeA* gene in this *Anabaena* strain was responsible for the resulting heterocyst formation exhibited on nitrate-containing media. In another cyanobacterial species *Synechococcus* (strain PCC 7942) five open reading frames forming the *narA* gene cluster have been isolated which complement nitrate reductase mutants. The ORF's appear to have similarity to several *moa* genes and the *moeA* gene from *E. coli* (Rubio *et al.*, 1998).

In the nitrogen-fixing soybean microsymbiont bacterium *Bradyrhizobium japonicum*, sequencing of a DNA region containing the tryptophan biosynthetic genes *trpDC* revealed the presence of a third ORF with high similarity to the *E. coli* MoeC protein, located downstream of *trpDC* (Kuykendall and Hunter, 1997). In *E. coli*, it has been observed that the predicted amino acid sequence of *moaB* has significant homology to that of *trpC* (Rivers *et al.*, 1993). Similarity was also observed between the gene products of *moaA* and *moaB* from *E. coli* and those of *trpDC*

from *B. japonicum*. It is unclear if there is any functional or evolutionary significance to these observations.

Genome sequencing projects carried out on a number of prokaryotes (*e.g. H. influenzae*) have also revealed loci which exhibit similarity to previously cloned *E. coli* genes (Fleischmann *et al.*, 1995) and are presumably involved in MoCo biosynthesis within these organisms.

#### **1.5.4 Eukaryotic Molybdenum Cofactor Biosynthesis Genes**

Thus far only a few eukaryotic MoCo biosynthesis genes have been cloned and characterised. Kamdar and colleagues reported the cloning of *cinnamon* from the fruit fly *Drosophila melanogaster* (Kamdar *et al.*, 1994) representing the first eukaryotic molybdenum cofactor biosynthesis gene to be isolated. The Cinnamon protein was found to be similar to three of the *E. coli* MoCo proteins, namely MoaB, MoeA and MogA, as well as to Gephyrin, a protein found in the rat central nervous system. More recently, a cDNA clone of the gene *cnx1* from the plant *Arabidopsis thaliana* has been cloned and sequenced (Stallmeyer *et al.*, 1995). Cnx1 was found to be similar to Cinnamon, Gephyrin, MoaB, MoeA and MogA. The similarity of both Cinnamon and Cnx1 to several different *E. coli* MoCo biosynthesis proteins suggests that in higher eukaryotes a multifunctional protein has evolved from the separate prokaryotic forms. Two other MoCo biosynthesis genes from *A. thaliana*, *cnx2* and *cnx3*, have also been cloned and sequenced (Hoff *et al.*, 1995). The inferred Cnx2 and Cnx3 proteins were found to be similar to the *E. coli* proteins MoaA and MoaC respectively. Cnx2 and Cnx3 therefore appear to encode monofunctional proteins which each carry out a separate enzymatic step in the early stages of plant MoCo biosynthesis. A recent communication has revealed that the genes *cnx4*, *cnx5*, and *cnx6* in

*A. thaliana* have been isolated (Mendel, 1997). The gene *cnx4* appears to code for the protein  $\gamma$ -adaptin which is involved in the vesicle transport system of eukaryotic cells. The function of Cnx4 is unclear, but it does not appear to be directly involved in MPT synthesis or Mo insertion. It would appear that Cnx5 is similar to *E. coli* MoeB and that Cnx6 is similar to the *E. coli* MoaE protein. Again, each of the genes (*cnx4-6*) seem to encode monofunctional proteins.

## 1.6      Research Objectives

The major objectives of this research were to isolate and characterise two of the molybdenum cofactor biosynthesis genes (i) *cnxG* and (ii) *cnxE* from *A. nidulans* *vis à vis* to sequence both these genes and to determine if any nitrate or ammonium-mediated gene regulation was apparent at the transcriptional level. In addition, it was hoped that sequencing strains carrying *cnxG* or *cnxE* mutations would allow identification of critical amino acid residues essential for protein functionality. It was also expected that HPLC determination of precursor Z and molybdopterin levels in appropriate mutant strains would shed light on the possible metabolic functions of the proteins encoded by these two *cnx* genes within the molybdenum cofactor biosynthesis pathway.

During my studies the *A. nidulans* *cnxABC* locus was successfully cloned and a number of mutants sequenced ( S.E. Unkles, personal communication). In this regard a facet of my research became the HPLC analysis of *cnxA*, *cnxB* and *cnxC* mutants, to provide further information on the levels of MoCo biosynthesis intermediates found within such strains. The effect of different sole nitrogen sources on wild-type strain levels of compound Z and molybdopterin was also investigated by HPLC to establish whether any nitrogen metabolite regulation of MoCo biosynthesis was apparent.

## CHAPTER 2

### MATERIALS AND METHODS

#### 2.1 Strains

##### 2.1.1 Bacterial Strains

Plasmids were propagated in *Escherichia coli* strain DH5 $\alpha$  (F',  $\emptyset$ 80d *lacZ* M15, *endA1*, *recA1*, *hsdR17*, *r<sub>k</sub>-m<sub>k</sub><sup>+</sup>*, *supE44*, *thi-1*, *l<sup>-</sup>*, *gyrA*, *relA1*,  $\Delta$ (*lacZYA-argF*) U169).

The *E. coli mogA* mutant strain JL3705 (F-*araD139*,  $\Delta$ (*argF-lac*) U169, *deoC1*, *flbB5201*, *ptsF25*, *relA1*, *rpsL150*, *rbsR*,  $\Delta$ *mog-1Km<sup>r</sup>*), the *E. coli moeA* mutant strain NS9 (F-*araD139*,  $\Delta$ (*argF-lac*) U169, *deoC1*, *flbB5201*, *ptsF25*, *relA1*, *rpsL150*, *rbsR*, *moeA-009*), the *E. coli moaD* mutant strain MJ431 (F- *rpsL* dGal<sup>r</sup> *chlM*) and the wild-type *E. coli* strain MC4100 (F-*ara D139*,  $\Delta$ (*argF-lac*) U169, *deoC1*, *flbB5201*, *ptsF25*, *relA1*, *rpsL150*, *rbsR*) were obtained from Professor D.H. Boxer, University of Dundee.

The *Aspergillus nidulans* cosmid library was held within *E. coli* strain HB101 (*supE44 ara14 gal K2 lacY1*  $\Delta$ (*gpt-proA*) 62 *rpsL20* (*Str<sup>r</sup>*) *xyl-5 mtl-1 recA13*  $\Delta$ (*mcrC-mrr*) *HsdS<sup>-</sup>* (*r<sup>-</sup> m<sup>-</sup>*)). The cosmid library was obtained commercially from the University of Kansas Fungal Genetics Stock Centre.

The *E. coli* strains XL1-Blue MRF' ( $\Delta$ (*mcrA*) 183  $\Delta$ (*mcrCB-hsdSMR-mrr*) 173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI $\phi$  Z*  $\Delta$ M15 Tn10 (*Tet<sup>r</sup>*)]<sup>c</sup> and SOLR (*el4<sup>-</sup>* (*McrA<sup>-</sup>*)  $\Delta$ (*mcrCB-hsdSMR-mrr*) 171 *sbcC recB recJ uvrC umuC* :: TnS (*Kan<sup>r</sup>*) *lac gyrA96 relA1 thi-1 endA1*  $\lambda^R$  [F' *proAB lacI $\phi$  Z*  $\Delta$ M15]<sup>c</sup> Su<sup>-</sup>) were used during manipulation of the *A. nidulans* cDNA library. These strains were obtained as part of the Stratagene ZAP-cDNA synthesis kit.

Protein over-expression was carried out in *E. coli* strain BL21 (DE3) (*hsdS gal  $\lambda$ clts857 ind1 Sam7 nin5 lacUV5-T7* gene 1). This strain was obtained from Dr. S.E. Unkles, Monash University.

The *Pseudomonas aeruginosa* strain which was used as an indicator organism during anaerobic growth, was provided by Dr. D. Thirkell, University of St Andrews. This strain was a recent clinical isolate obtained from the Fife Public Health Laboratories at Victoria Hospital, Kirkcaldy, and as such has no definitive genotype and strain number.

## 2.12 Fungal Strains

The *Aspergillus nidulans* strain *biA1* (a biotin auxotroph) was used as wild-type. The mutant *A. nidulans* strains used in experiments are listed below.

<u>Strain</u>	<u>Genotype</u>
$\alpha$ 8	<i>biA1 cnxA9</i>
G055	<i>biA1 cnxB11</i>
G832	<i>yA2 pyroA4 cnxC3</i>
$\alpha$ 34	<i>yA2 wA3 cnxE10</i>
G059	<i>biA1 cnx E14</i>
$\alpha$ 36	<i>biA1 cnxE12</i>
GK849	<i>biA1 cnxE849</i>
$\alpha$ 27	<i>yA2 wA3 cnxE3</i>
G058	<i>biA1 wA2 cnxE16</i>
$\alpha$ 37	<i>biA1 cnxE13</i>
G614	<i>yA2 pyroA4 cnxG4</i>
$\alpha$ 23	<i>yA2 wA3 cnxG2</i>
GK1222	<i>biA1 cnxG1222</i>

Strains  $\alpha 8$ ,  $\alpha 34$ ,  $\alpha 36$ ,  $\alpha 27$   $\alpha 37$  and  $\alpha 23$  were supplied by Mr. A. Simpson, University of Cambridge. Strains GK849 and GK1222 were donated by Mr. G. Kan'an, University of St Andrews. All the other mutant strains and the wild-type strain were obtained from Dr. A.J. Clutterbuck, University of Glasgow. All of the strains held within this collection were derived from a single wild isolate of *A. nidulans*, NRRL 194 (Pontecorvo *et al.*, 1953).

## 2.2 Growth and Storage of Cultures

### 2.2.1 Bacterial Cultures

Short-term storage of bacterial strains was on LB agar plates kept at 4°C. For long-term storage cultures were held at -70°C in LB medium containing a final concentration of 20% glycerol. All cultures were grown aerobically at 37 °C unless otherwise stated. When required, anaerobic growth was achieved using an anaerobic gas jar (Oxoid) containing an activated catalyst (palladium-coated pellets of alumina) in conjunction with Anaerocult A gas packs (Merck). An indicator strip was included within the jar to establish that anaerobiosis had been achieved. When stated, a plate culture of a *Psuedomonas aeruginosa* strain was also placed in the jar to check that the conditions were anaerobic. Anaerobic growth was carried out at room temperature.

### 2.2.2 *A. nidulans* Cultures

Stock cultures of strains were maintained on complete medium slants at 4 °C (or at room temperature). Long-term storage of *A. nidulans* cultures was on silica gel (Roberts, 1969). All cultures were grown aerobically at 37 °C unless otherwise specified.

### **2.3**      **Solutions**

Unless otherwise stated, solutions were prepared according to Sambrook *et al.*, (1989) using sterile distilled H<sub>2</sub>O. Where appropriate, solutions were sterilised by autoclaving at 15 lb/inch<sup>2</sup> for 15 min. When required, solutions were filter-sterilised through a 0.2 µm filter unit (MILLEX-GS).

### **2.4**      **Media**

Solid medium was prepared by the addition of 1.2% agar to the appropriate liquid media unless otherwise stated. Medium was sterilised by autoclaving at 15 lb/inch<sup>2</sup> for 15 min.

#### **2.4.1**    ***E. coli* Culture Media**

##### **Luria-Bertani (LB) Broth**

25 g Luria broth base (Difco) was dissolved in 1 litre distilled H<sub>2</sub>O.

##### **Selective LB Media**

Where appropriate, selective LB medium was prepared by the addition of the antibiotics ampicillin (sodium salt), kanamycin or tetracycline at the concentrations indicated below.

<b><u>Antibiotic</u></b>	<b><u>Stock Solution</u></b>	<b><u>Final Concentration</u></b>
Ampicillin	10 mg/ml in H <sub>2</sub> O	50 µg/ml
Kanamycin	10 mg/ml in H <sub>2</sub> O	25 µg/ml
Tetracycline	5 mg/ml in ethanol	10 µg/ml

The stock solutions of ampicillin and kanamycin were filter-sterilised prior to use. Antibiotics were added to medium which had

previously been autoclaved and pre-cooled to 55 °C. Antibiotic stock solutions were stored at 4 °C for no more than 30 days.

### SOC Medium

<u>Component</u>	<u>Amount Added</u>
Bactotryptone	2 g
Yeast extract	0.5 g
1M NaCl	1 ml
1M KCl	0.25 ml
H <sub>2</sub> O	97 ml

This solution was autoclaved and cooled to room temperature before the addition of 1 ml 2M Mg<sup>++</sup> (1 M MgCl<sub>2</sub>.6H<sub>2</sub>O + 1 M MgSO<sub>4</sub>.7H<sub>2</sub>O) filter-sterilised stock solution and 1 ml 2M glucose filter-sterilised stock solution.

### NZY Agar

21.1 g of NZYDT (Gibco BRL) broth base was dissolved in 1 litre distilled H<sub>2</sub>O.

### Top Agarose

0.7% agarose was added to 100 ml NZY broth.

### Lactate Nitrate (LN) Medium

Lactate nitrate medium was prepared according to Venables and Guest (1968), with peptone from casein substituted for vitamin-free casamino acids.

<u>Component</u>	<u>Amount per Litre</u>
NaH <sub>2</sub> PO <sub>4</sub> .2H <sub>2</sub> O	6.08 g
K <sub>2</sub> HPO <sub>4</sub>	10.6 g
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	2 g
KNO <sub>3</sub>	4 g
Potassium lactate (50 %)	7 ml
Peptone from casein	0.4 g
Salts solution	5 ml

The volume was made up to 1 litre with distilled H<sub>2</sub>O.

#### Salts Solution

<u>Component</u>	<u>Amount per Litre</u>
MgSO <sub>4</sub> .7H <sub>2</sub> O	10 g
MnCl <sub>2</sub> .4H <sub>2</sub> O	1 g
FeSO <sub>4</sub> .7H <sub>2</sub> O	0.05 g
CaCl <sub>2</sub> .2H <sub>2</sub> O	0.1 g

The volume was made up to 1 litre with distilled H<sub>2</sub>O. A trace of concentrated HCl was added to clarify. The solution was stored at 4 °C.

#### 2.4.2 A. nidulans Culture Media

The recipes for minimal and complete medium were based on that described by Pontecorvo *et al.* (1953) and Cove (1966).

### Minimal Medium (MM)

<u>Component</u>	<u>Amount per Litre</u>
Salts solution	50 ml
Trace elements	1 ml
Vitamin solution (optional)	1 ml
Glucose	10 g

The volume was made up to 1 litre with distilled H<sub>2</sub>O and the pH brought to 6.5 with 5 M NaOH.

### Complete Medium (CM)

<u>Component</u>	<u>Amount per Litre</u>
Salts solution	50 ml
Trace elements	1 ml
Vitamin solution	1 ml
Glucose	10 g
Peptone from casein	2 g
Yeast extract	1 g
Casein hydrolysate	1 g

The volume was made up to 1 litre with distilled H<sub>2</sub>O and the pH brought to 6.5 with 5 M NaOH.

### Salts Solution

<u>Component</u>	<u>Amount per Litre</u>
KCl	26 g
MgSO <sub>4</sub> .7H <sub>2</sub> O	26 g
KH <sub>2</sub> PO <sub>4</sub>	76 g

The volume was made up to 1 litre with distilled H<sub>2</sub>O. The solution was stored at 4 °C.

### Trace Elements Solution

<u>Component</u>	<u>Amount per Litre</u>
Mo <sub>7</sub> O <sub>24</sub> .4H <sub>2</sub> O	1.1 g
H <sub>3</sub> BO <sub>4</sub>	11.1 g
CoCl <sub>2</sub> .6H <sub>2</sub> O	1.6 g
CuSO <sub>4</sub> .5H <sub>2</sub> O	1.6 g
EDTA	50 g
FeSO <sub>4</sub> .7H <sub>2</sub> O	5 g
MnCl <sub>2</sub> .4H <sub>2</sub> O	5 g
ZnSO <sub>4</sub> .7H <sub>2</sub> O	22 g

The volume was made up to 1 litre with distilled H<sub>2</sub>O and boiled with stirring. After cooling to 60 °C the solution was brought to pH 6.5 with 5M KOH. Storage was in the dark at 4 °C.

## Vitamin Solution

<u>Component</u>	<u>Amount per Litre</u>
Biotin	2.5 g
Nicotinic acid	2.5 g
PABA	0.8 g
Pyrimidine hydrochloride	1 g
Pantothenate	2 g
Riboflavin	2.5 g
Aneurine	1.5 g
Choline hydrochloride	20 g

The volume was made up to 1 litre with distilled H<sub>2</sub>O. The prepared vitamin solution was stored in the dark at 4 °C.

## Growth Supplements

Where required supplements were added to the growth medium as detailed below. All supplements were filter-sterilised prior to use and added to autoclaved, precooled (55 °C) medium. Solutions containing the supplements were stored at 4 °C.

<u>Supplement</u>	<u>Stock Solution</u>	<u>Amount Added to 100 ml</u>
Biotin	0.1 g/100 ml	1 ml
Pyridoxine HCl	0.5 g/100 ml	1 ml

## Nitrogen Sources

The different nitrogen sources were sterilised by autoclaving. Sterile stock solutions were stored at 4 °C and added, when required, to nitrogen-free minimal medium at the concentrations indicated overleaf.

<u>Nitrogen Source</u>	<u>Stock Solution</u>	<u>Final Concentration</u>
Ammonium tartrate	1 M	10 mM
Sodium nitrate	1 M	10 mM
L-proline	1 M	10 mM
Adenine	1 M	10 mM

## **2.5** Preparation of Plasmid/Cosmid DNA from *E. coli*

### **2.5.1** Small Scale ('mini') Preparation of Plasmid/Cosmid DNA

Small scale preparation of plasmid/cosmid DNA was carried out using the QIAprep Spin Plasmid Kit (QIAGEN) according to the manufacturer's instructions. The QIAprep plasmid kit procedure is based on the modified alkaline lysis method of Birnboim and Doly (1979).

### **2.5.2** Medium Scale ('midi') and Large Scale ('maxi') Preparation of Plasmid/Cosmid DNA

Medium and large scale preparation of plasmid/cosmid DNA was carried out using the QIAGEN Plasmid Midi Kit (QIAGEN tip100) for a medium preparation or the QIAGEN Plasmid Maxi Kit (QIAGEN tip 500) for a large preparation, according to the manufacturer's instructions. The QIAGEN plasmid purification procedure is based on the optimised alkaline lysis method of Birnboim and Doly (1979).

## **2.6** Preparation of *A. nidulans* Chromosomal DNA

### **2.6.1** Small Scale Preparation

Small scale preparation of fungal chromosomal DNA was carried out as described by Leach *et al.* (1986).

A 10 ml culture (minimal medium plus the appropriate strain supplements) was inoculated with a loopful of conidia and incubated overnight at 37 °C, 260 rpm (New Brunswick Scientific orbital shaker).

The resulting mycelium was harvested over sterile Miracloth and washed with 10 ml H<sub>2</sub>O. The mycelium was then freeze-dried and mechanically diced before adding 0.5 ml LETS buffer and vortexing to mix. 1 ml phenol:chloroform:isoamyl alcohol (25:24:1) was added and the mixture vortexed (20 sec) before centrifugation at 3000 rpm (Eppendorf Bench Centrifuge 5415 C). 500 µl of the aqueous phase was combined with 1 ml 100 % ethanol and placed at -70 °C for 15 min. The sample was then centrifuged for 15 min (14000 rpm, bench centrifuge) and the resulting pellet air dried and resuspended in 40 µl TE buffer (10:1) (pH 8.0).

#### LETS Buffer

0.1 M LiCl  
10 mM EDTA  
10 mM Tris.HCl (pH 8.0)  
0.5 % SDS

#### TE Buffer (10:1) (pH 8.0)

10 mM Tris.HCl (pH 8.0)  
1 mM EDTA (pH 8.0)

#### 2.6.2 Large Scale Preparation

Large scale preparation of fungal chromosomal DNA was carried out using the Nucleon II DNA Extraction Kit (Scotlab), according to the manufacturer's instructions.

#### 2.7 Spectrophotometric Quantitation of DNA

The concentration of DNA preparations was estimated spectrophotometrically (SP6-550 UV/VIS PYE Unicam spectrophotometer) by measuring the optical density of an appropriate

sample dilution at 260 nm, and then substituting this value into the following equation.

$$\text{OD}_{260 \text{ nm}} \times \text{Sample Dilution} \times 50 = \text{DNA Concentration } (\mu\text{g/ml})$$

To determine the quality of a DNA preparation the OD of a suitable sample dilution was measured at 260 nm and 280 nm. The ratio of the  $\text{OD}_{260 \text{ nm}}/\text{OD}_{280 \text{ nm}}$  measurements allowed the purity to be estimated, with a ratio of approximately 1.8 indicating a pure DNA preparation.

## **2.8**      **Restriction Endonuclease Digestion of DNA**

Type II restriction endonucleases were used to digest DNA samples at the temperature and time recommended by the manufacturer (GibcoBRL or Pharmacia) in the appropriate commercially prepared reaction buffer.

## **2.9**      **Agarose Gel Electrophoresis of DNA**

Agarose gel electrophoresis of DNA samples was carried out essentially as described by Sambrook *et al.*, (1989). The concentration and size of the horizontal gel slabs were chosen with regard to the number and nature of the samples analysed. Agarose gels were used routinely at a concentration of 0.8%, however 2% agarose gels were prepared for resolution of the ampicillin gene PCR products because of their small size (approximately 300 bp). For both 0.8% and 2% gel concentrations the agarose was solubilised in 1x TAE electrophoresis buffer by heating in a microwave oven. 5  $\mu\text{l}$ /100 ml of a 10 mg/ml ethidium bromide stock solution (Sharp *et al.*, 1973) was added before the gel was poured into the appropriate gel former using a comb to provide the sample wells. Once

set, the comb was removed and the gel submerged in an electrophoresis tank containing 1 x TAE buffer.

Before loading into the wells DNA samples were mixed with a 0.1 volume of gel loading buffer (30% sucrose, 100 mM EDTA, several crystals of bromophenol blue). Electrophoresis was performed at 15-50 volts until the required separation of DNA had been achieved (visualised by the migration of the bromophenol blue dye). The intercalation of ethidium bromide molecules within the DNA allowed visualisation using a short wavelength UV transilluminator. Gels were photographed under UV light using a Polaroid MP-4 land camera and Polaroid type 667 film.

*Hind*III digested lamda phage DNA or 100 bp ladder DNA (Pharmacia) were used as molecular weight markers.

#### 50 x TAE Electrophoresis Buffer

242 g Tris.HCl base

57.1 ml Glacial acetic acid

100 ml 0.5 M EDTA (pH 8.0)

#### **2.10** Isolation of DNA Fragments from Agarose Gels

The restricted DNA sample (known to contain the fragment of interest) was loaded onto a 0.8% agarose gel and electrophoresis carried out to ensure sufficient separation of the desired DNA band. The required fragment (visualised under UV light) was excised from the gel using a surgical blade and transferred to an Eppendorf tube. If appropriate the tube was stored at -70°C until further analysis.

DNA was extracted from the surrounding agarose gel using the Pharmacia Gene Clean Kit. The procedure followed by this kit is based on that originally described by Vogelstein and Gillespie (1979).

### **2.11**      **Ethanol Precipitation of DNA**

DNA was concentrated by ethanol precipitation. This involved the addition of 0.1 volume of 3 M sodium acetate ( $\text{CH}_3\text{COONa}\cdot 3\text{H}_2\text{O}$ ) (pH 5.2) and 2 volumes of chilled ( $-20^\circ\text{C}$ ) 96% ethanol to the sample to be concentrated. These components were mixed and incubated at  $-20^\circ\text{C}$  for at least 1 h. Precipitates were pelleted by centrifugation at 14000 rpm (bench centrifuge) for 15 min and the pellet washed in 70% ethanol before air drying at room temperature and resuspending in the appropriate volume of sterile TE (10:1) (pH 8.0).

### **2.12**      **Protein Removal by Phenol Extraction of DNA**

Phenol was equilibrated with 100 mM Tris.HCl (pH 8.0) and then 10 mM Tris.HCl (pH 8.0) until the pH of the aqueous layer was greater than 7.6 and contained 0.1% 8-hydroxyquinoline, and then mixed with an equal volume of chloroform:isoamylalcohol (24:1). An equal volume of the resulting phenol:chloroform:isoamylalcohol (25:24:1) solution was added to the aqueous DNA solution, mixed to form an emulsion and then centrifuged at 4000 rpm for 5 min (bench centrifuge) at room temperature to separate the aqueous and organic phases. The upper, aqueous layer was then removed to a fresh tube taking care not to disturb the precipitated protein at the interface. The extraction was repeated until the interphase was observed to be clear. Any traces of phenol in the aqueous sample were then removed by extraction with chloroform:isoamylalcohol (24:1) alone using the same method.

### **2.13**      **Drop Dialysis of DNA**

DNA samples were dialysed to remove salt residues and other impurities. The DNA sample to be dialysed was placed on to the surface of a Millipore VS filter (0.025  $\mu\text{M}$  pore size) floating on sterile  $\text{H}_2\text{O}$ , and

left for 1-2 h. After this period of time the DNA sample was removed from the filter surface.

#### **2.14**     **<sup>32</sup>P Labelling of DNA**

DNA probes were radiolabelled with  $\alpha$ -<sup>32</sup>P-dCTP (ICN Pharmaceuticals) as in the method described by Feinberg and Vogelstein (1983). DNA was initially denatured by boiling for 2 min and then placed on ice (quenched) to prevent the re-annealing of strands. The denatured double-stranded linear DNA molecules were then added to the other components in the labelling reaction mix, as detailed below.

<u>Component</u>	<u>Amount Added</u>
Denatured DNA	20-50 ng
$\alpha$ - <sup>32</sup> P-dCTP	30 $\mu$ Ci
Klenow fragment*(5U/ $\mu$ l)	1 $\mu$ l
BSA (10 mg/ml)	2 $\mu$ l
dNTP-C (500 $\mu$ M)	2 $\mu$ l
5 x Labelling buffer	10 $\mu$ l

\*DNA Polymerase I Klenow fragment (Pharmacia)

The reaction volume was made up to 50  $\mu$ l with sterile distilled H<sub>2</sub>O.

The labelling reaction was incubated for 30 min at 37 °C and then 10  $\mu$ l of a concentrated solution of blue dextran dye was added. Labelled probe was separated from unincorporated nucleotides by passing the sample through a Sephadex G50 Nick column (Pharmacia) using TE (10:1) (pH 8.0) as the eluting agent. Labelled nucleotides were boiled for 2 min, quenched on ice and used in hybridisation reactions with the appropriate filter(s).

## **2.15 Southern Transfer and Hybridisation**

### **Southern Transfer**

The capillary method of DNA transfer is based on the technique described by Southern (1975). The DNA fragments to be analysed were separated according to size by electrophoresis through an agarose gel. Concentrations of approximately 1 µg plasmid/cosmid DNA and 6 µg *A. nidulans* genomic DNA were used routinely. Sequential depurination, denaturation and neutralisation of the DNA was then applied by shaking the gel gently (50-100 rpm, Luckham Model R100/TW rotatest plate shaker) in the appropriate solution for the time indicated.

#### 1) Depurination Solution

0.25 M HCl, shaken for 30 min.

#### 2) Denaturation Solution

1.5 M NaCl

0.5 M NaOH, shaken for 40 min.

#### 3) Neutralising Solution

1.5 M NaCl

1 M Tris.HCl, pH 7.4, shaken for 50 min.

The DNA fragments were then transferred from the gel to an Electran nylon membrane (BDH) using Southern's method of capillary transfer (an updated version of which is detailed in Sambrook *et al.*, 1989) with 20 x SSC as buffer. After overnight blotting the filter was briefly washed in 2 x SSC and then air dried on a piece of Whatman 3MM paper at room temperature. DNA was fixed to the filter using an XL-1500 UV

crosslinker on 'optimal crosslink'. The fixed blot could then be stored at room temperature between 2 sheets of 3MM paper until required.

### 20 x SSC

3 M NaCl

0.3 M Sodium Citrate, pH 7.0

### DNA Hybridisation

The filter was incubated at 65 °C for 3 h to overnight in an appropriate volume of hybridisation solution, shaking at 70 reciprocal strokes per min in a waterbath (Grant).

### Hybridisation Solution (100 ml)

<u>Solution</u>	<u>Volume Added</u>
20 x SSPE	25 ml
30% PEG 6000	20 ml
10% SDS	10 ml
10% Skimmed milk (Marvel)	5 ml
5% Na <sub>4</sub> P <sub>2</sub> O <sub>7</sub>	2 ml
Herring sperm DNA (5 mg/ml)	5 ml
Distilled H <sub>2</sub> O	33 ml

The Herring sperm DNA solution was prepared according to Sambrook *et al.*, (1989). All of the other solutions were autoclaved prior to use except for the skimmed milk which was steamed for 15 min.

## 20 x SSPE

3 M NaCl

0.2 M NaH<sub>2</sub>PO<sub>4</sub>.H<sub>2</sub>O

25 mM EDTA

The pH was brought to 7.4 with 5M NaOH.

The appropriate <sup>32</sup>P-labelled DNA probe (see section 2.14) was added to the hybridisation solution and incubation continued overnight.

## Filter Washing

Unhybridised probe was removed by washing the filter first with 5 x SSC, 0.1% SDS, 0.1% Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> and then 2 x SSC, 0.1% SDS, 0.1% Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> for 20 min periods until background radiation on the membrane was insignificant. Washes were performed at 65 °C and 70 strokes per min.

## Autoradiography

After washing, the filter was sealed in thick plastic sheeting and placed in an autoradiography cassette (fitted with white intensifying screens) against Kodak X-OMAT S film or X-OMAT AR fast film and placed at -70°C. The length of exposure varied according to the strength of the signal expected. Films were developed in a Fuji RG II X-Ray film processor.

## Removal of Radioactive Probe from Filters

Filters were stripped of radioactive probe, if required, by boiling the membrane in a solution of 0.1 % SDS and then cooling to room temperature. This process was completed until the radiation present

on the membrane was found to be negligible. The filter could then be rehybridised with a different probe.

## **2.16**     **Manipulation of the *A. nidulans* Cosmid Bank**

The chromosome-specific *A. nidulans* cosmid bank was purchased from the University of Kansas Fungal Genetics Stock Centre, USA. The individual cosmids (each approximately 35-50 Kb in length) were ligated into either LORIST2 or pWE15 vectors (see Appendix for maps) and held within *E. coli* strain HB101. The cosmid bank was stored at -70 °C, mixed with glycerol.

### **2.16.1**     **Cosmid Propagation on LB Agar Plates**

Cells containing the cosmid of interest were revived by transferring a small sample from the surface of the frozen stock to a LB agar plate containing the appropriate antibiotic (ampicillin or kanamycin) and streaking across the plate surface before overnight incubation at 37 °C. Cosmid DNA was prepared from single bacterial colonies as described in section 2.5.

### **2.16.2**     **Hybridisation Reactions Using Cosmid DNA**

An Electran nylon filter was placed on a dry LB agar plate containing the appropriate antibiotic. A small sample from each of the frozen cosmid stocks of interest was touched onto the surface of the filter using a sterile plastic pipette tip. The plate was then incubated at 37 °C overnight, after which time 'spots' of bacterial growth were visible on the filter where the samples had been applied. The filter was removed from the surface of the agar and sequentially transferred onto pieces of Whatman 3 MM chromatography paper saturated with the appropriate solution for the times shown overleaf.

<u>Solution</u>	<u>Contact Time with Filter</u>
10% SDS	3 min
Denaturing solution	10 min
1 M Tris.HCl, pH 8.0	5 min
1 M Tris.HCl, pH 8.0	5 min
Neutralising solution	5 min

After drying in air on a piece of Whatman 3 MM paper the cosmid DNA was fixed to the filter using an XL-1500 UV crosslinker on 'optimal crosslink'. The filter could then be stored between 2 sheets of 3 MM paper at room temperature until proceeding with the hybridisation. Prehybridisation, hybridisation with the radiolabelled probe of choice, filter washing and autoradiography were carried out in an identical manner to that detailed in section 2.15.

## **2.17 Manipulation of the *A. nidulans* cDNA Library**

The *A. nidulans* cDNA library was created by Dr. R. Williams (St Andrews University) using the Stratagene ZAP-cDNA synthesis kit according to the manufacturer's instructions. cDNA clones were held within the Uni-ZAP XR lambda-phage vector (see appendix for map). The amplified library was held as a phage stock at 4 °C. The *E. coli* strains XL1-MRF' and SOLR and the ExAssist helper phage used during manipulation of the library were provided as part of the kit.

### **2.17.1 Screening the cDNA Library**

#### **Primary Screen**

The *E. coli* strain XL1-MRF' was grown in 10 ml LB broth (containing 0.2 % maltose and 10 mM MgSO<sub>4</sub>) to an OD<sub>600</sub> of 0.4 and then centrifuged for 5 min at 2000 rpm (using an angle rotor in a Mistral

100 centrifuge). The resulting pellet was resuspended in 10 mM MgSO<sub>4</sub> at an OD<sub>600</sub> of 0.4. An appropriate volume of the cDNA library was added to 150 µl of the cell suspension and incubated for 15 min at 37 °C (100 rpm). 4 ml top agarose was subsequently added and the sample quickly poured over the surface of an NZY agar plate. Plates were incubated at 37 °C for 7 h.

An Electran nylon filter was placed on to the surface of the agar plate and left in contact for 1 min. To determine orientation, holes were pushed through the filter into the agar with a hypodermic needle. The filter was then turned over and sequentially transferred onto pieces of Whatmann 3 MM paper saturated with the appropriate solution for the times shown below.

<u>Solution</u>	<u>Contact Time with Filter</u>
Denaturing solution	30 sec
Neutralising solution	30 sec
2 x SSC	5 min

The process was then repeated with a second filter, which was left in contact with the agar surface for 3 min. After drying in air on a piece of Whatmann 3 MM paper the DNA was fixed to the filters using an XL-1500 UV crosslinker on 'optimal crosslink'. Filters could then be hybridised with the radiolabelled probe of choice. Prehybridisation, hybridisation, filter washing and autoradiography were carried out in an identical manner to that detailed in section 2.15.

### Subsequent Screens

Using the orientation marks, the position of any positive plaque(s) was identified. An area of top agarose from around the positive signal was removed and placed in 100  $\mu$ l SM buffer. 5  $\mu$ l  $\text{CHCl}_3$  was added to the sample and vortexed for 20 sec, then spun for 2 min at 14000 rpm (bench centrifuge). The supernatant was diluted 1/50 and 1/500 and 10  $\mu$ l of each dilution was used to infect 100  $\mu$ l *E. coli* XL1-MRF' cells (resuspended in 10 mM  $\text{MgSO}_4$  to an  $\text{OD}_{600}$  of 0.4 as in the primary screen). Samples were incubated for 15 min at 37 °C shaking at 100 rpm, then each combined with 4 ml top agarose and quickly poured over the surface of an NZY agar plate. Plates were incubated at 37 °C for 7 h. Filters were applied to the plates for 1 min and 3 min periods and then treated as in the primary screen. The screening procedure was repeated until all of the plaques present on the plate were positive. An isolated plaque could then be removed from the plate and excised from the Uni-ZAP XR lambda phage vector as detailed overleaf (section 2.17.2).

### SM Buffer (1 litre)

<u>Component</u>	<u>Amount Added</u>
NaCl	5.8 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	2 g
1 M Tris.HCl (pH 7.5)	50 ml
2 % Gelatin solution	5 ml

The volume was made up to 1 litre with distilled  $\text{H}_2\text{O}$ .

### **2.17.2** **Excision of Purified cDNA Clones from the Uni-ZAP XR Vector**

Excision of the pBluescript phagemid (for map refer to the Appendix) held within the Uni-ZAP XR vector allowed the cDNA inserts to be characterised in a plasmid system. Single-clone phagemid excision was carried out essentially as described in the Stratagene ZAP-cDNA synthesis kit instruction manual. The excised pBluescript phagemid was plated in *E. coli* strain SOLR (as detailed in the instruction manual), after which plasmid DNA could be prepared in the usual manner (as detailed in sections 2.5.1 and 2.5.2).

### **2.17.3** **Mass Excision of the cDNA Library**

Mass excision of the cDNA library was carried out as detailed in the ZAP-cDNA synthesis kit manual. This resulted in the production of an excised cDNA library represented as phagemids packaged as filamentous phage particles.

#### **Preparation of DNA Following Mass Excision**

An appropriate volume of *E. coli* strain SOLR was grown at 37 °C, shaking at 250 rpm to an OD<sub>600</sub> of 0.5-1.0. The appropriate volume of phage stock (containing the excised cDNA library) was added to the cells and incubated for 1 h at 37 °C, shaking at 100 rpm, followed by a further 3 h at 37 °C, shaking at 250 rpm. The cells were then used for plasmid DNA preparation as detailed in section 2.5.2.

### **2.18** **Preparation of Competent *E. coli* Cells**

The preparation of competent *E. coli* cells was based on the method described by Cohen *et al.*, (1972). 1 ml of an overnight culture of *E. coli* in LB broth was used to inoculate 200 ml fresh LB broth and

incubated at 37°C, shaking at 250 rpm in a New Brunswick Scientific orbital shaker until the culture reached an OD of approximately 0.25 at a wavelength of 600 nm. Cells were pelleted at 3000 rpm in a Sorvall GS 3 rotor (Sorvall RC-58 refrigerated superspeed centrifuge) for 10 min at 4 °C and gently resuspended in 100 ml ice-cold 100 mM MgCl<sub>2</sub>, 5 mM Tris.HCl, pH 7.4. A further centrifugation was performed at 3000 rpm (GS 3 rotor) for 5 min at 4 °C and cells gently resuspended in 100 ml ice-cold 100 mM CaCl<sub>2</sub>, 5 mM Tris.HCl, pH 7.4. The cells were incubated on ice for 60 min after which time they were competent for transformation. The competent cells were centrifuged (as previously) and the pellet resuspended in 1 ml 100 mM CaCl<sub>2</sub>, 5 mM Tris.HCl, pH 7.4, 14% glycerol. 100 µl aliquots of the suspension were dispensed into pre-cooled Eppendorf tubes and frozen immediately. The cells were stored at -70 °C until required.

### **2.19 Transformation of Competent *E. coli* Cells**

An appropriate amount of DNA (typically 50-500 ng) was added to 100 µl competent *E. coli* cells (thawed on ice from a frozen stock), mixed and incubated on ice for 30 min. Cells were heat-shocked by incubation at 42 °C for 90 sec and then returned to ice for 1-2 min. 900 µl SOC liquid medium was then added to the mixture and cells incubated at 37°C for 1 h at 125 rpm. Appropriate aliquots of the transformation mix were spread onto LB agar plates containing the required antibiotic for the selection of transformed cells. The selection plates were incubated at 37 °C, overnight. A 100 µl sample of competent *E. coli* cells which was not exposed to transforming DNA served as a negative control when plated onto selective medium.

Where required (*i.e.* when selecting for successfully-complemented *E. coli* MoCo biosynthesis mutants) the transformed

competent cells were plated onto LN agar plates with the appropriate selective antibiotic. Such plates were incubated under anaerobic conditions at room temperature for the appropriate length of time.

## **2.20**     **Electro-transformation of *E. coli* Cells**

The method used for electro-transformation of *E. coli* was based on that originally described by Dower *et al.*, (1988).

### **2.20.1**     **Preparation of Cells**

50 ml LB broth was inoculated with 500  $\mu$ l of an *E. coli* overnight culture and incubated at 37 °C, shaking at 250 rpm until an OD<sub>600</sub> of 0.6 was reached. The cells were chilled on ice for 30 min, then centrifuged at 3500 rpm in a Sorvall RC-58 refrigerated superspeed centrifuge with an HB-4 rotor for 15 min at 4 °C. The cell pellet was resuspended in 50 ml ice cold sterile distilled H<sub>2</sub>O and centrifuged as before. The pellet was then resuspended in 25 ml ice cold water and centrifuged (as previously) and the resulting pellet resuspended in 1 ml ice cold 10 % glycerol. After a final spin (as before) the pellet was resuspended in 300  $\mu$ l ice cold 10 % glycerol and stored in 50  $\mu$ l aliquots at -80 °C.

### **2.20.2**     **Electro-transformation**

50  $\mu$ l cells (prepared as above) and an electroporation cuvette were placed on ice for 30 min. 0.5  $\mu$ g DNA was added to the cells, mixed gently with the pipette tip, and incubated on ice for 30 sec before transferring to the cuvette. The cuvette was placed into the sample holder of an electroporation unit (EASYJECT) and pulsed with 2.5 kV for 5 msec. 1 ml SOC broth was immediately added to the cells and subsequently incubated for 1 h at 37 °C, shaking at 100 rpm. The cells

were then plated on the appropriate selective medium and incubated aerobically overnight at 37 °C or anaerobically at room temperature for the appropriate length of time.

## **2.21**     **Preparation and Transformation of *A. nidulans* Protoplasts**

The protoplast preparation and transformation procedures are based on the methods described by Balance *et al.* (1983); Tilburn *et al.* (1983); John and Peberdy (1984); Yelton *et al.* (1984).

### **2.21.1**     **Preparation of Protoplasts**

2 or 3 flasks each containing 400 ml appropriately supplemented minimal medium were inoculated with a conidial suspension in 20 ml saline Tween 80 (of approximately  $5 \times 10^6$  conidia per ml) and grown for 12 hours at 26°C, shaking at 250 rpm in a New Brunswick Scientific orbital shaker. Young mycelial cells were harvested through sterile Miracloth, washed with 25 ml cold 0.6 M MgSO<sub>4</sub> and suspended in 5 ml filter-sterilised osmotic medium. 1.5 ml osmotic medium containing 20 mg Novozyme 234 (NovoEnzyme Products Ltd) was added to the cell suspension and incubated for 5 min on ice. After this time 1.25 ml of a 12 mg/1.5 ml solution of BSA in osmotic medium was introduced and the whole mixture incubated at 30 °C for 2-2.5 h with gentle orbital shaking (100 rpm). The formation of protoplasts was checked by microscopic examination. Protoplasts were separated from the mycelial debris by centrifugation at 5000 rpm for 20 min at 4 °C in a Sorvall RC-58 refrigerated superspeed centrifuge using an HB-4 swing-out rotor, with an overlay of an equal volume of cold osmotic medium and 2 x volume cold trapping buffer. The protoplasts formed a thick 'bushy' band halfway up the tube and the cell debris pelleted at the bottom. Protoplasts were retrieved and resuspended in an equal volume of 1 x STC before

centrifugation at 7500 rpm for 5 min at 4 °C (using the same centrifuge and rotor). The protoplast pellet was then resuspended in 10 ml 1 x STC and spun down for 5 min at 7500 rpm, 4 °C (as before). The final protoplast pellet was resuspended in an appropriate volume of 1 x STC, depending on the number of samples to be transformed (50 µl per sample plus another 50 µl for a negative 'no DNA' control) and leaving 32 µl to use in the determination of protoplast number and viability.

#### Determination of Protoplast Number and Viability

2 µl of the resuspension was removed and placed on the central plateau of a haemocytometer (Improved Neubauer Haemocytometer, Weber). The number of protoplasts in 4 squares of the haemocytometer was counted under the microscope (with each square split into 16 smaller squares, therefore 64 smaller squares counted). This value was then applied to the following equation in order to determine the number of protoplasts per ml.

$$\frac{\text{No. protoplasts in 4 squares}}{64} \times 4 \times 10^6 = \text{No. protoplasts per ml}$$

The volume over each small square is 1/4000 mm<sup>3</sup> or for all practical purposes 1/4000000 ml, therefore we must multiply the average number of protoplasts present in each small square by a value of 4 x 10<sup>6</sup> to calculate the number of protoplasts per ml.

A concentration of approximately 2 x 10<sup>8</sup> protoplasts/ml was desirable in order to allow efficient transformation. This value however, did not distinguish between viable and non-viable protoplasts and the proportion of viable non-protoplast cells with undegraded cell walls. The

concentration of viable protoplasts was tested by plating a serial dilution ( $10^{-2}$  to  $10^{-5}$  in 1 x STC) of 10  $\mu$ l untransformed protoplasts on non-selective medium. 10  $\mu$ l untransformed protoplasts, diluted 1:100 in H<sub>2</sub>O, were plated onto non-selective medium to determine the proportion of cells with undegraded cell walls. The subtraction of the number of colonies resulting from cells with undegraded cell walls (able to withstand an H<sub>2</sub>O environment) from the number of viable protoplast colonies allowed the number of viable protoplasts without cell walls which were able to be transformed to be determined. 10  $\mu$ l untransformed protoplasts (diluted 1:100 in 1 x STC) were plated onto selective medium to test for mutant reversion or for non-sufficient selection conditions.

### 2.21.2 Transformation Procedure

An appropriate concentration of plasmid or cosmid DNA in a volume of no more than 25  $\mu$ l was used in each sample to be transformed. An equal volume of 2 x STC was added to the DNA and the total volume of each sample made up to 50  $\mu$ l by adding 1 x STC. 50  $\mu$ l of the protoplast suspension was added to each 50  $\mu$ l sample and mixed gently. 25  $\mu$ l 60% PEG 6000 was added to each sample and incubated on ice for 30 min. A further 1 ml 60% PEG 6000 was then added to each sample and incubated at room temperature for 30 min before centrifugation with 5 ml 1 x STC at 4500 rpm using an angle rotor in a Mistral 100 centrifuge. The resulting pellet was resuspended in 300  $\mu$ l 1 x STC and 2 x 150  $\mu$ l aliquots plated onto transformation selection medium to provide duplicate plates. Selective medium consisted of minimal medium at pH 6.5 containing 1.2 M sorbitol and the appropriate supplements except for that required for growth of untransformed cells. Plates were incubated at 37 °C for up to 5 days. A negative control was provided by treating protoplasts as

above but without the addition of DNA (*i.e.* using 50  $\mu$ l 1 x STC only in the transformation procedure).

### Osmotic Medium

1.2 M MgSO<sub>4</sub>

10 mM Sodium phosphate buffer, pH 7.0

The pH was adjusted to 5.8 with 0.2 M Na<sub>2</sub>HPO<sub>4</sub>. The solution was filter-sterilised before use and stored at 4 °C.

### 1 x STC

1.2 M Sorbitol

10 mM Tris.HCl, pH 7.5

10 mM CaCl<sub>2</sub>

### Trapping Buffer

0.6 M Sorbitol

100 mM Tris.HCl, pH 7.0

### Saline Tween 80

0.01 % Tween 80

0.9 % NaCl

## **2.22** Polymerase Chain Reaction (PCR)

The PCR method used was derived from Saiki *et al.*, (1988). PCR was used to amplify specific 'target' DNA sequences. In this study, PCR was used in the amplification of: (1) the *E. coli* ampicillin resistance gene sequence, if present, in putative *A. nidulans* transformant colonies; (2) the *cnxE* and *cnxG* gene sequences from wild-type and mutant

*A. nidulans* strains for use as templates in sequencing; (3) a fragment of DNA containing the *cnxG* gene coding region, with engineered enzyme restriction sites, suitable for ligating into a protein over-expression vector and (4) specific DNA fragments for use as hybridisation probes.

In each case reactions were carried out in a total volume of 100 µl in 0.5 ml PCR tubes. Clean gloves were worn throughout the procedure and all solutions and pipette tips were double autoclaved before use to prevent contamination with DNA from another source (Kwok and Higuchi, 1989). PCR was carried out in a HYBAID thermal reactor programmed with the appropriate procedures. Typically 100 ng template DNA and 100 µM of each primer were used in reactions. *Taq* polymerase (GibcoBRL) was used in PCR reactions when testing for the presence of the *E. coli* ampicillin resistance gene sequence in putative transformants, under optimised conditions. All other PCR reactions were carried out using Dynazyme (Flowgen) with a standard set of reaction conditions but with annealing temperatures altered according to the primer sequences to ensure specific annealing of the primers to the desired sequences in the template DNA. The annealing temperature was calculated using the formula  $2(A + T) + 4(G + C)$  for the bases present in each primer. Where possible, two primers with approximately the same annealing temperature were chosen for each PCR reaction.

### **2.22.1 PCR Analysis of *A. nidulans* Transformant Colonies**

PCR was used to identify the presence of the *E. coli* ampicillin resistance gene by allowing amplification of this sequence (if present) from genomic DNA extracted from putative *A. nidulans* transformant colonies. The genomic DNA used in reactions was prepared from transformant colonies by the method described in section 2.6.1. Wild-type *A. nidulans* (*biA1*) DNA was prepared in the same manner to act as a

negative control. The primers used (AMP 1 and AMP 2) anneal to sequences within the *E. coli* ampicillin resistance gene *bla* encoding  $\beta$ -lactamase which is carried on a number of vectors.

### Primers

AMP 1: 5'- CTGTGACTGGTGAGTAC -3'

AMP 2: 5'- CAACATTTCCGTGTCGC -3'

The PCR mixture was prepared as detailed below.

<u>Component</u>	<u>Amount Added</u>
DNA template	100ng
10 x buffer	10 $\mu$ l
20 $\mu$ M AMP 1 primer	5 $\mu$ l
20 $\mu$ M AMP 2 primer	5 $\mu$ l
2 mM dNTP	5 $\mu$ l
5U/ $\mu$ l <i>Taq</i> polymerase	0.5 $\mu$ l
25 mM MgCl <sub>2</sub>	6 $\mu$ l

The volume was made up to 100  $\mu$ l with sterile distilled water and the sample overlaid with mineral oil (Sigma) to prevent evaporation. A negative 'no template DNA' control was also performed to provide a test for any contamination. PCR was then carried out using the reaction conditions specified below/overleaf.

Cycle 1 -	Denaturing	94°C for 2 min
	Annealing	50°C for 20 s
	DNA Synthesis	72°C for 30 s
Cycles 2-26 -	Denaturing	94°C for 20 s

Annealing	50°C for 20 s
DNA Synthesis	72°C for 30 s

After the reaction was complete 10 µl of each PCR product was analysed by electrophoresis on a 2% agarose gel.

### 2.22.2 PCR Amplification of Specific *A. nidulans* Genomic Gene Sequences

PCR was used to amplify specific gene (*cnxE* and *cnxG*) sequences from chromosomal *A. nidulans* DNA to provide templates for sequencing. Genomic DNA was prepared from the wild-type *A. nidulans* strain and the appropriate *cnxG* and *cnxE* mutant strains as described in section 2.6.2. An appropriate amount of the DNA was digested with an enzyme which did not cut the gene of interest. The resulting digest was ethanol precipitated and resuspended in an appropriate volume of TE (10:0.1). Approximately 100 ng DNA from this suspension was used in the PCR reaction mix. Primers were selected which would amplify suitable fragments of genomic template DNA to allow sequencing of *cnxE* or *cnxG*.

### Primers used to Obtain Overlapping Genomic DNA Fragments Containing *cnxE*

#### Pair 1

E1A: 5'- GTGCCTGAGGTGTCAAT -3'

E1B: 5'- AGTGACTIONTGTGTCGGGT -3'

Annealing temperature = 52 °C

Pair 2

E2A: 5'- GTTGAAGGTCACGCCTT -3'

E2B: 5'- CAGCATATCCATCAACG -3'

Annealing temperature = 50 °C

Pair 3

E3A: 5'- CGAGGTTCCGGTGAATA -3'

E3B: 5'- CCCGTAGGCTTTGTTCC -3'

Annealing temperature = 52 °C

Pair 4

E4A: 5'- CTGAAATCATGGGGGAT -3'

E4B: 5'- CTTTGGGAGACGCTGTA -3'

Annealing temperature = 50 °C

Pair 5

E5A: 5'- CCCTTGTAGCAGTAAC -3'

E5B: 5'- GAATAAGTCTTTGGAGG -3'

Annealing temperature = 48 °C

Primers used to Obtain a Single Genomic DNA Fragment Containing *cnxG*

3R1: 5'- GGAGGTCTGATCGATTC -3'

FMUT1: 5'- TTCTACCAGCTACTCTG -3'

Annealing temperature = 50 °C

## PCR Mixture

<u>Component</u>	<u>Amount Added</u>
DNA Template	100 ng
10 x Buffer	10 $\mu$ l
10 mM dNTP	1 $\mu$ l
100 $\mu$ M Primer 1	1 $\mu$ l
100 $\mu$ M Primer 2	1 $\mu$ l
Dynazyme	0.5 $\mu$ l

The volume was made up to 100  $\mu$ l with sterile distilled H<sub>2</sub>O and the sample overlaid with mineral oil to prevent evaporation. A negative 'no template DNA' control was also performed. PCR was then carried out using the reaction conditions specified below. The annealing temperature (represented by an asterisk) was varied according to the specific set of primers used during each reaction.

Cycle 1 -	Denaturing	94 °C for 1 min
	Annealing	* °C for 15 s
	DNA Synthesis	72 °C for 50 s
Cycles 2-31 -	Denaturing	94 °C for 20 s
	Annealing	* °C for 15 s
	DNA Synthesis	72 °C for 50 s

10  $\mu$ l of each PCR product was analysed by electrophoresis on a 0.8 % agarose gel. The rest of each PCR product (90  $\mu$ l) could then be purified (discussed in section 2.23) and the appropriate volume used as a template for DNA sequencing.

### 2.22.3 PCR Amplification of a DNA Fragment Suitably Engineered for Ligation into a Protein Expression Vector

PCR was used to amplify the *cnxG* gene coding region and to introduce appropriate restriction sites into the PCR product so as to allow ligation of the resulting fragment into a protein expression vector. A *cnxG* cDNA clone was used as template. Primers were designed to anneal to the DNA sequences encoding the first few amino acids at the amino-terminal end and the last few amino acids (discounting the translational stop codon) at the carboxy-terminal end of CnxG. An appropriate restriction site was included in each primer sequence so as to generate a PCR fragment which could be ligated into an appropriate expression vector. Sequences which would allow efficient cleavage of the fragment during restriction were also included in the primers.

#### Primers

*cnxG* 5' Primer (*EcoRI* site):

5'- GCCGGAATTCATGCTGCCAGAAGGCCAG -3'

*cnxG* 3' Primer (*SalI* site):

5'- TGCGGTCGACAACCAGTGCTCTAAAAGCATT -3'

The reaction mix contained the same amounts of the various components as described in section 2.22.2. As previously, the reaction volume was made up to 100  $\mu$ l with sterile distilled H<sub>2</sub>O and the sample overlaid with mineral oil. Again, a 'no template DNA' test for contamination was performed. The PCR conditions were also as described in section 2.22.2, however an annealing temperature of 58 °C was used in each cycle.

10  $\mu$ l of each PCR product was analysed on a 0.8 % agarose gel. The rest of the product (90  $\mu$ l) could then be purified (see section 2.23) and the appropriate volume used in the ligation reaction.

#### 2.22.4 PCR Amplification of DNA Fragments for use as Hybridisation Probes

PCR was used to amplify specific DNA fragments, from within or around the *cnxE* or *cnxG* genes, for use as probes in hybridisation reactions. In each case an appropriate clone of the gene was used as the template. Primers were chosen which would enable the required region of DNA to be amplified.

##### Primers used to Create Probes for the Analysis of *cnxE*

###### Pair 1

E1R1: 5'- GATCACCAAGATAGTGC -3'

E1GET!: 5'- GAGACACGTCCGAAGTG -3'

Annealing temperature = 50 °C

###### Pair 2

E1R2<sup>nd</sup>RINFO: 5'- GATCTGGATGCGATCCG -3'

E1F1: 5'- CCTCATGCAGCACATCG -3'

Annealing temperature = 54 °C

##### Primers used to Create a Probe for the Analysis of *cnxG*

3FEND: 5'- GGCATGTCGAGATTGAG -3'

3REND: 5'- CCGTTACAGATCGGTTC -3'

Annealing temperature = 52 °C

The reaction mix contained the same amounts of the various components as described in section 2.22.2. As previously, the reaction volume was made up to 100  $\mu$ l with sterile distilled H<sub>2</sub>O and the sample overlaid with mineral oil. Again, a 'no template DNA' test for contamination was performed. The PCR conditions were also as described in section 2.22.2, however using the appropriate annealing temperature (as indicated above) for each pair of primers.

10  $\mu$ l of each PCR product was analysed on a 0.8 % agarose gel. The rest of the product (90  $\mu$ l) could then be purified (see section 2.23) and the appropriate volume used in the desired hybridisation reaction(s).

### **2.23 Purification of PCR Products**

PCR products were purified using the Glassmax DNA Isolation Spin Cartridge System (GibcoBRL) according to the manufacturer's instructions. The procedure followed by this kit is based on that originally described by Vogelstein and Gillespie (1979).

### **2.24 Construction and Detection of Hybrid Plasmids**

#### **2.24.1 Ligation of DNA Fragments into Vectors**

Both the DNA clone/PCR fragment and the vector DNA were digested with the appropriate restriction enzyme(s). The specific DNA fragment to be ligated was cut from an agarose gel following electrophoresis, taking care to protect the DNA from UV irradiation by placing the gel on plastic sheeting while removing the desired band. The fragment was then purified using the Gene Clean Kit (Pharmacia) as detailed in section 2.10. The digested vector DNA was heated to 65 °C for 15 min to inactivate the restriction enzyme and then treated with alkaline phosphatase (if applicable) to remove 5'-phosphate groups and suppress self-ligation/recircularisation of the vector. In a typical 10  $\mu$ l volume, 0.4

units shrimp alkaline phosphatase (SAP) (United States Biochemical, Amersham) and 1  $\mu$ l 10 x SAP buffer were added and incubated for 1 h at 37 °C. The enzyme was subsequently heat-inactivated at 65 °C for 15 min. The vector DNA and the purified DNA fragment were both dialysed for 1 h before proceeding to the ligation step.

For each ligation attempt, 4 units bacteriophage T4 DNA ligase (GibcoBRL) and 1  $\mu$ l 10 x ligase buffer were added to three different concentrations of vector:insert DNA (200 ng:50 ng, 200 ng:200 ng and 50 ng:200 ng), in a total volume of 10  $\mu$ l. Control ligations of SAP treated and non-SAP treated linearised vector, with and without ligase/ligase buffer were also performed. The ligation reactions were left for 4 h at room temperature, then transferred to 4 °C overnight. After heating for 15 min at 65 °C the total volume of each ligation reaction was used to transform 100  $\mu$ l DH5 $\alpha$  competent cells.

#### **2.24.2 Detection of Recombinant Plasmids by $\alpha$ -complementation**

Many vectors, including the pUC series, carry a DNA sequence which contains the regulatory sequences and the coding information for the first 146 amino acids (the amino-terminal fragment) of the  $\beta$ -galactosidase gene (*lacZ*), within which a polycloning site has been engineered. When a vector containing such a sequence is held within an *E. coli* strain which codes for the carboxy-terminal portion of  $\beta$ -galactosidase (*e.g.* *E. coli* strain DH5 $\alpha$ ), the two protein fragments can associate to form an enzymatically active  $\beta$ -galactosidase protein, in a process termed  $\alpha$ -complementation (Ullmann *et al.*, 1967).

$\alpha$ -complementation can be recognised by the formation of blue colonies resulting from the cleavage of X-Gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside) to galactose and 5-bromo-4-chloro-2-indole. The insertion of a foreign DNA sequence into the polycloning site of the vector, almost

invariably results in the production of an amino-terminal fragment which is incapable of  $\alpha$ -complementation, and bacteria harbouring such recombinant plasmids are therefore recognised by the formation of white colonies.

According to the rationale discussed above, 40  $\mu$ l X-Gal (20 mg/ml in dimethylformamide) and 4  $\mu$ l IPTG (isopropylthio- $\beta$ -D-galactoside)(200 mg/ml in H<sub>2</sub>O)(NBL Gene Sciences) were spread over the surface of each selective LB plate (onto which the transformed DH5 $\alpha$  cells were plated) to provide a white/blue colour indication of colonies containing recombinant plasmids. Any resulting white colonies were removed and suspended in 5  $\mu$ l H<sub>2</sub>O. Two blue colonies (containing vector DNA only) were also picked and suspended in H<sub>2</sub>O to act as controls. 0.5  $\mu$ l of each bacterial suspension was streaked onto an LB plate and incubated overnight at 37 °C to produce stock cultures. 15  $\mu$ l cracking buffer was added to the remaining volume in each suspension and mixed well with the pipette tip. Samples were subsequently loaded on a dry agarose gel and run for 10 min at 50 V. The gel was then covered with 1 x TAE and electrophoresis continued for 2 h at the same voltage. Any colonies containing extra plasmid DNA (in comparison to the two vector-only controls) were identified and plasmid preparations performed as required using material present on the stock plates.

### Cracking Buffer (1 ml)

<u>Component</u>	<u>Volume Added (ul)</u>
EDTA	50
10 % SDS	50
0.2 M NaOH	50
*RNase (10 mg/ml)	20
H <sub>2</sub> O	830

\*RNase = Ribonuclease A from bovine pancreas (Sigma)

To this a few crystals of bromocresol green were added.

### 2.24.3 Detection of Recombinant Plasmids by *in situ* Hybridisation

When ligating a DNA fragment into a vector which did not contain the partial amino-terminal *lacZ* gene sequence (*i.e.* the protein expression plasmid pET21a), it was necessary to screen transformed colonies for the presence of recombinant clones by *in situ* hybridisation. In such a case the transformed DH5 $\alpha$  competent cells were plated onto selective LB medium (without IPTG and X-Gal). Circular (82 mm) Electran optimised nylon filters (BDH) were applied to the surface of the plates containing the resulting transformed colonies and orientation holes were marked with a needle through the filter and agar. Each filter was then removed, turned over, and placed (bacteria side up) onto pieces of 3 MM Whatmann paper saturated with the appropriate solution for the times shown below.

<u>Solution</u>	<u>Contact Time with Filter</u>
Denaturing solution	10 min
Neutralising solution	5 min

After drying in air on a piece of Whatman 3 MM paper the bacterial DNA was fixed to the filters using an XL-1500 UV crosslinker on 'optimal crosslink'. The filters were then hybridised with the insert DNA which was used to construct the subclone, thus allowing any positive colonies containing recombinant plasmid clones to be identified. Prehybridisation, hybridisation, filter washing and autoradiography were carried out in an identical manner to that detailed in section 2.15. Alignment of the developed film with the original plate meant that growth from any positive colonies could be removed and plasmid DNA prepared as required.

## 2.25 DNA Sequence Analysis

### 2.25.1 Sequence Analysis of Plasmid DNA

DNA was sequenced by means of the chain-termination method (Sanger *et al.*, 1977) using the Sequenase version 2.0 kit with T7 Sequenase version 2.0 DNA Polymerase (United States Biochemical, Amersham). Sequencing reactions were carried out as detailed in the kit instructions using primers and denatured template DNA which were prepared as detailed below/overleaf. DNA was radiolabelled using  $\alpha$ -<sup>35</sup>S-dATP (ICN Pharmaceuticals).

#### Primers

The universal forward (-40) primer (5'- GTTTTCCAGTCACGAC -3') was obtained from United States Biochemical, Amersham. The universal reverse primer (5'- GGAAACAGCTATGACCATG -3') was synthesised in St Andrews University. All other primers were either synthesised in St Andrews University or were obtained from the Cruachem Primer Service.

Primers used to Sequence *cnxG*

3R1:	5'- GGAGGTCTGATCGATTC -3'
3F1:	5'- GCAAGTGGCGCGCATAG -3'
3F2:	5'- CTCAGCCGTTGCACACG -3'
3F3:	5'- CTCAGGATTGACTGGAG -3'
3F4:	5'- GACGATTGCTCCTCGAG -3'
RGEN1:	5'- GTCACGTGTATTTATCC -3'
RGEN2:	5'- CTACCCGATTCGAAGCC -3'
RGEN3:	5'- CTCGAGGAGCTTATCTG -3'
RGEN4:	5'- CGTGTGCAACGGCTGAG -3'
RGEN5:	5'- GTGGCGTAGAGTCGTTG -3'
FMUT1:	5'- TTCTACCAGGTACTCTG -3'
FMUT2:	5'- GAGTCTGCAGCGCGTTG -3'
FMUT3:	5'- CTGAAGTCCATTCAGAC -3'
FMUT4:	5'- GCGAGAACCGCAAGACG -3'

Primers used to Sequence *cnxE*

E1R1:	5'- GATCACCAAGATAGTGC -3'
E1R2:	5'- GTTATCACTCTGCCCGG -3'
E1R3:	5'- GCACATACAACGCCTTC -3'
E1R4:	5'- GGACACAGTACAAAGGG -3'
E1R5:	5'- CAATGAGAACGTGCGGC -3'
E1R6:	5'- CCTCGATCCTATCTTGC -3'
E1R7:	5'- GCTCGATCTGCTAAAGC -3'
E1R8:	5'- CATTGTCACATGCTTTC -3'
E1R9:	5'- CGAGGCACTGATGATGG -3'
E1F1:	5'- CCTCATGCAGCACATCG -3'
E1F2:	5'- GACGATGGCCGCTGTCG -3'
E1F3:	5'- CCTCGTGCACGTTGTAC -3'

E1F4: 5'- GGGCACGATGGTTGCGC -3'  
E1F5: 5'- GAGCCGGTATAGACCTC -3'  
E1F6: 5'- GAAGGAGAATACGACTG -3'  
E1F7: 5'- GTGACGGCGTGGTAGAC -3'  
E1F8: 5'- CCTGGGTTTCGAGATGG -3'  
K18F2: 5'- CCAGGCTGAAGCACCCAG -3'  
K18F3: 5'- GACCAGTGCTGAGGCTG -3'  
K18F4: 5'- CTCCACGAAGAGCATCC -3'  
K18F5: 5'- GTAGCTGCCAGCAGGCC -3'  
K18F7: 5'- GCCTCAGTGCTTCGTCC -3'  
K18F8: 5'- CTCCAGCGTGGAGCGTC -3'  
K18F9: 5'- GTGTTATCCTTAACCGC -3'  
E1FUstream: 5'- GGAATGCAGTATCTGAC -3'

The dried primer pellet was resuspended in 100  $\mu$ l distilled H<sub>2</sub>O. 5  $\mu$ l of this was diluted in 995  $\mu$ l H<sub>2</sub>O and used in the spectrophotometric determination of the primer concentration (as detailed in section 2.7). The primer was then diluted to a working concentration of 1 pmol/ $\mu$ l.

#### Denaturation of Template DNA

5  $\mu$ g template DNA was denatured per sequencing reaction. 0.1 vol 2 M NaOH, 2mM EDTA was added to the DNA and incubated at 37 °C for 30 min. 0.1 vol 3 M sodium acetate (pH 5.2) and 3 vol ethanol (-20 °C) were then added and incubated at -20 °C for 10 min. The sample was then spun for 10 min (14000 rpm, bench centrifuge), the pellet washed with 70 % ethanol, air dried and resuspended in an appropriate volume of distilled H<sub>2</sub>O (7  $\mu$ l H<sub>2</sub>O per sequencing reaction).

### Sequencing Gel Preparation

Sequi-Gen sequencing gel rigs (BIO-RAD) in two sizes (large rig = 38 x 50 cm /small rig = 21 x 50 cm) were used during the electrophoresis of samples. The rigs were assembled and used according to the manufacturers instructions.

The bottom of an appropriately sized sequencing gel rig was sealed by adding 40 ml sequencing gel /200  $\mu$ l 25 % ammonium persulphate (APS) /200  $\mu$ l TEMED (small rig) or 80 ml sequencing gel /400  $\mu$ l 25 % APS /400  $\mu$ l TEMED (large rig) to the pouring support and allowing to set for 15 min. The gel was then poured using 60 ml sequencing gel /60  $\mu$ l 25 % APS /60  $\mu$ l TEMED (small rig) or 100 ml sequencing gel /100  $\mu$ l 25 % APS /100  $\mu$ l TEMED (large rig) applied with a large syringe with the rig at a 30 ° tilt. One (small rig) or two (large rig) 24-toothed 'sharks-tooth' combs (0.4 mm thick) were inserted between the rig plates at the top of the gel to a depth of 5 mm, and clamped in place with bulldog clips. The gel was allowed to set overnight. The combs were then removed and replaced with the teeth just entering the surface of the gel. 400 ml 1 x TBE buffer was added to the bottom reservoir and 600 ml (small rig) or 1600 ml (large rig) was poured into the gel rig buffer tank behind the sequencing gel. The gel was pre-run at 2 kV until the gel temperature reached 50 °C.

### Electrophoresis of Samples

Immediately before loading onto the sequencing gel samples were heated to 75 °C for 2 min and then quenched on ice. A hypodermic syringe containing 1 x TBE was used to flush out urea which had accumulated in the sequencing wells. 3  $\mu$ l of each sample was loaded onto the gel and electrophoresis continued for the time required at 50 °C. Following electrophoresis, the gel was transferred to Whatman 3 MM

paper and dried under vacuum (BIO-RAD Model 583 gel drier) at 80 °C for 2 h.

### Autoradiography

The dried gel was exposed to X-OMAT S film (Kodak) at room temperature in an autoradiography cassette without intensifying screens. Films were processed using a Fuji RG II X-ray film processor. Exposure times varied according to the intensity of the signal expected.

### Computer Sequence Analysis

The Seqnet computer program (Daresbury Laboratory) was used for the basic analysis and manipulation of the DNA sequences obtained and to produce multiple sequence alignments. The BLASTX computer program (Altschul *et al.*, 1990; Gish and States, 1993) allowed comparison of the DNA sequences obtained with the protein sequences of previously cloned genes held within the Swiss-Prot and updated Swiss-Prot data bases.

### Sequencing Gel (500 ml)

<u>Component</u>	<u>Volume Added</u>
Urea	210 g
10 x TBE	50 ml
40 % Acrylamide*	75 ml

The volume was made up to 500 ml with distilled H<sub>2</sub>O. The sequencing gel was stored at 4 °C.

\*Acrylamide:bisacrylamide (40 %: 2.105 %) (Scotlab EASIGel)

## 10 x TBE (1 litre)

<u>Component</u>	<u>Volume Added</u>
Tris-base	109.0 g
Boric acid	55.6 g
EDTA	7.4 g

The volume was made up to 1 litre with distilled H<sub>2</sub>O.

### 2.25.2 Sequence Analysis of PCR Products

Purified PCR products containing *cnxE* gene sequences from the wild-type strain and from several mutant strains were sequenced by automated DNA sequencing using an ABI 373 A fluorescent sequencing apparatus and the PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems). In each reaction, 100-200 ng purified PCR product was used as template with the primers present at a concentration of 10  $\mu$ M. The sequences derived from the mutant strains were compared with the wild-type using the Sequencher computer program (Gene Codes Corporation).

### 2.26 Isolation of *A. nidulans* RNA

#### 2.26.1 Total RNA Isolation

A suspension of approximately  $5 \times 10^6$  conidia in 20 ml saline Tween 80 solution was used to inoculate 250 ml minimal medium containing the appropriate nitrogen source and supplements. Cells were grown for 16 h at 26 °C, shaking at 200 rpm and then harvested over sterile Miracloth. The cell material was washed with 50 ml distilled H<sub>2</sub>O, then pressed dry before freezing immediately in liquid N<sub>2</sub>. Mycelium was stored at -70 °C until required.

The method used for the isolation of total RNA from *A. nidulans* cells was based on that of Cathala *et al.* (1983). Reference was made to Davis *et al.* (1986) for guidance on general RNA methods and solutions. To avoid or minimise any possible RNase contamination, extreme care was taken throughout the isolation, quantification and manipulation of the RNA. Clean gloves were worn throughout the procedure. Prior to use all glass/Pyrex equipment was siliconised, then washed thoroughly, first with detergent (Decon), then with distilled H<sub>2</sub>O before baking twice at 180 °C for 8 hours. All other equipment such as plasticware and magnetic fleas were double autoclaved before use. All solutions were treated with 0.1 % diethyl pyrocarbonate (DEPC) by stirring for at least 1 h and then autoclaved. Exceptions were made with solutions containing Tris.HCl, which were instead made up in DEPC-treated, autoclaved, distilled H<sub>2</sub>O.

4 g pressed wet weight of mycelium was ground to a fine powder in liquid N<sub>2</sub>. 12 ml GuSCN buffer and 1.2 ml β-mercaptoethanol (Sigma) were added to the ground mycelium and stirred vigorously, before homogenising through first a 19 gauge, then a 21 gauge hypodermic needle (TERUMO). 60 ml 4 M LiCl was added with stirring and the sample subsequently incubated at room temperature for 20 min, then at 4 °C overnight. The sample was then inverted to mix, spun for 5 min, 2500 rpm at 4 °C in a Sorvall HB-4 swing-out rotor (Sorvall RC-58 refrigerated superspeed centrifuge) to remove mycelial debris, then the supernatant respun for 90 min, 8000 rpm (4 °C, swing-out rotor). The resulting pellet was resuspended in 40 ml 3 M LiCl by passing successively through 19 and 21 gauge hypodermic needles, as previously. The sample was centrifuged for 60 min, 8000 rpm (4 °C, swing-out rotor) and the pellet resuspended in 2 ml TESDS using a 21 gauge needle. The sample was then extracted with an equal volume of phenol:chloroform:isoamyl

alcohol (25:24:1) (sample vortexed then spun for 2 min at 14000 rpm on a bench centrifuge) and the resulting aqueous (upper) phase collected. The residual organic phase and interphase was back-extracted by the addition of an equal volume of TESDS buffer, followed by spinning for 2 min, 14000 rpm (bench centrifuge) and then removing the resulting aqueous phase and pooling with the first. The pooled aqueous phase was then extracted with phenol:chloroform:isoamyl alcohol until no interphase was visible, and then extracted once with chloroform:isoamyl alcohol (24:1) to remove any residual phenol. The RNA was precipitated overnight at -20 °C with 0.1 vol, 3 M sodium acetate (pH 5.2) and 2 vol absolute ethanol (-20 °C). The RNA was pelleted by spinning at 14000 rpm for 15 min, washed with 70 % ethanol, vacuum dried and resuspended in the appropriate volume of DEPC-treated H<sub>2</sub>O. The concentration and purity of the RNA was determined spectrophotometrically (see section 2.28) and the concentration was also checked by electrophoresis on a formaldehyde minigel.

#### GuSCN Buffer

5 M Guanidine isothiocyanate (Fluka)

10 mM EDTA

50 mM Tris.HCl, pH 7.5

The guanidine isothiocyanate was dissolved at 50 °C. The complete solution was filter-sterilised and stored at -20 °C.

#### TESDS

1 M Tris.HCl, pH 7.5

0.5 M EDTA, pH 8.0

10 % SDS

### 2.26.2 Messenger RNA (mRNA) Isolation

*A. nidulans* mRNA was isolated from 1 g total RNA using the Pharmacia Biotech mRNA Purification Kit according to the manufacturers instructions. This kit is based on the use of spun columns for the affinity purification of polyadenylated RNA on oligo(dT)-cellulose, as described by Aviv and Leder (1972). All precautions taken to avoid RNase contamination were carried out as previously detailed in the isolation of total RNA (see section 2.26.1).

### 2.27 Electrophoresis of RNA, Northern Transfer and Hybridisation

#### Electrophoresis of RNA

The electrophoresis of RNA was carried out on a 1 % agarose gel containing 1 x MOPS buffer, 0.66 M formaldehyde and 0.2 µg/ml ethidium bromide. To eliminate RNase contamination, gel rigs, formers and combs were washed with detergent (Decon) and then rinsed with DEPC-treated H<sub>2</sub>O before use. 3-4 µg mRNA in 5 µl DEPC-treated H<sub>2</sub>O was combined with 15 µl loading buffer and heated to 95 °C for 2 min to denature, then placed directly on ice. A suitable volume of 0.24-9.5 Kb RNA ladder (GibcoBRL), which was used to provide a molecular weight marker, was treated in the same manner. Samples were loaded onto the formaldehyde gel and electrophoresis carried out in 1 x MOPS buffer, until the dye had migrated three-fourths down the gel.

#### 10 x MOPS Buffer

0.2 M MOPS [ 3-( N-morpholino) propanesulfonic acid]

0.05 M Sodium acetate

0.01 M EDTA

The pH was adjusted to 7.0 with 5 M NaOH.

### Loading Buffer (1.5 ml)

<u>Component</u>	<u>Volume Added</u>
Deionised formamide	0.72 ml
10 x MOPS buffer	0.16 ml
Formaldehyde	0.26 ml
H <sub>2</sub> O	0.18 ml
80 % Glycerol	0.1 ml
Bromophenol blue (saturated solution)	0.08 ml

The formamide was deionised by spinning for at least 1 h with 'Amberlite' monobed mixed resin IRN-150L.

### Northern Transfer

The gel was rinsed twice, for 20 min in 500 ml 10 x SSC to remove formaldehyde from the gel. The RNA was then transferred from the gel to an Electran nylon membrane (BDH) using the method detailed by Davis *et al.*, (1986) with 10 x SSC as buffer. After overnight blotting the filter was removed and air dried on a piece of Whatman 3 MM paper at room temperature. RNA was fixed to the filter using an XL-1500 UV crosslinker on 'optimal crosslink'. The fixed blot could then be stored at room temperature between 2 sheets of 3MM paper until required.

### RNA Hybridisation

The filter was incubated at 42 °C for 3 h to overnight in an appropriate volume of hybridisation solution, shaking at 70 reciprocal strokes per min in a Grant waterbath.

### Hybridisation Solution (100ml)

<u>Solution</u>	<u>Volume Added</u>
Deionized formamide	50 ml
20 x SSPE	30 ml
100 x Denhart's solution	5 ml
Herring sperm DNA (5mg/ml)	5 ml
10% SDS	10 ml

### 100x Denhart's Solution

- 2% Gelatin
- 2% PVP (Polyvinyl-pyrrolidone)
- 2% Ficoll 400
- 1% SDS

The appropriate  $^{32}\text{P}$ -labelled DNA probe (see section 2.14) was added to the hybridisation solution and the incubation continued overnight.

### Filter Washing

Unhybridised probe was removed by washing the filter with 5 x SSC, 0.1 % SDS, 0.1 %  $\text{Na}_4\text{P}_2\text{O}_7$ , then 2 x SSC, 0.1 % SDS, 0.1 %  $\text{Na}_4\text{P}_2\text{O}_7$ , and then 0.2 x SSC, 0.1 % SDS, 0.1 %  $\text{Na}_4\text{P}_2\text{O}_7$  for 20 min periods, until background radiation on the membrane was insignificant.

### Autoradiography

After washing, the filter was sealed in thick plastic sheeting and placed in an autoradiography cassette (fitted with white intensifying screens) against Kodak X-OMAT AR fast film and placed at  $-70\text{ }^\circ\text{C}$ . The

length of exposure varied according to the strength of the signal expected. Films were developed in a Fuji II X-Ray film processor.

### Removal of Radioactive Probe from Filters

Filters were stripped of radioactive probe, if required, by boiling the membrane in a solution of 0.1 % SDS and then cooling to room temperature. This process was completed until the radiation present on the membrane was found to be negligible. The filter could then be rehybridised with a different probe.

### Verification of Uniform RNA Transfer

In order to verify that the transfer of RNA to each lane of the filter was uniform, blots were stripped of the previous probe (as detailed above) and then rehybridised with a  $^{32}\text{P}$ -labelled DNA fragment specific for the constitutively expressed *A. nidulans actA* (actin) gene (Fidel *et al.* 1988).

## **2.28 Spectrophotometric Quantitation of RNA**

The concentration of RNA preparations was estimated spectrophotometrically by measuring the optical density of an appropriate sample dilution at 260 nm, and substituting this value into the following equation.

$$\text{OD}_{260 \text{ nm}} \times \text{Sample Dilution} \times 40 = \text{RNA Concentration } (\mu\text{g/ml})$$

To determine the quality of an RNA preparation the OD of a suitable sample dilution was measured at 260 nm and 280 nm. The ratio of the  $\text{OD}_{260 \text{ nm}}/\text{OD}_{280 \text{ nm}}$  measurements allowed the purity to be

estimated, with a ratio of approximately 2.0 indicating a pure RNA preparation (Sambrook *et al.*, 1989).

### **2.29**     **Nitrite Overlay Procedure for *E. coli***

The nitrite overlay procedure for the detection of nitrate reductase activity is based on that described by Glaser and DeMoss, (1971). *E. coli* cells were streaked out on an LB plate and grown anaerobically overnight. 4 ml 1 M sodium formate (HCOONa) (in 100 mM phosphate buffer), 2 ml 20 % KNO<sub>3</sub> and 4 ml 1.5 % molten agar were combined and poured quickly and evenly over the surface of the plate. The plate was left to set for 30 min. 2 ml 4 % sulphanimide (in 25 % HCl), 1 ml 0.08% N-1-NED and 4 ml 1.5 % molten agar were combined and poured over the plate surface as before. An immediate colour change in the agar around the *E. coli* growth from colourless to pink was indicative of nitrate reductase activity.

### **2.30**     **Protein Over-Expression in *E. coli***

The *E. coli* protein expression strain BL21 (DE3) was transformed with the gene of interest cloned into the expression vector pET21a (for the structure of pET21a refer to the Appendix) and transformant colonies selected on LB/ampicillin plates. Bacteriophage T7 RNA polymerase is carried on the bacteriophage λDE3, which is integrated into the chromosome of BL21, and can be used to direct the expression of genes cloned into vectors containing the bacteriophage T7 promoter (*e.g.* pET21a) (Studier and Moffatt, 1986).

#### **2.30.1**     **Time Course of Total Cell Protein Over-Expression**

A single transformed colony of BL21 (DE3) was used to inoculate 50 ml LB/ampicillin broth and incubated at 37 °C, shaking at

250 rpm (New Brunswick Scientific orbital shaker), until the OD<sub>600</sub> reached 0.6. A 50 µl sample was removed to act as an uninduced control, before adding IPTG to a final concentration of 1 mM to induce protein expression. The culture incubation was then continued at 30 °C (250 rpm) overnight, removing 50 µl samples at appropriate time intervals. Each 50 µl sample was centrifuged (1 min, 14000 rpm, bench centrifuge ) and the cell pellet resuspended in 15 µl 5 x SDS sample buffer. The samples, along with 10 µl Sigma prestained SDS-PAGE standard solution (containing 6 prestained protein markers), were boiled for 5 min directly before loading onto an SDS-polyacrylamide gel.

#### 5 x SDS Sample Buffer (800 µl)

<u>Component</u>	<u>Volume Added (µl)</u>
Distilled H <sub>2</sub> O	500
0.5 M Tris.HCl, pH 6.8	100
Glycerol	80
20 % SDS	80
β-mercaptoethanol	40

To this a few crystals of bromophenol blue were added.

#### 2.30.2 Lysozyme/Sonication Treatment of Cells to Determine Protein Solubility

A culture of transformed BL21 (DE3) cells was grown up and induced with IPTG as described above. 500 µl samples of cells (taken before induction and from an appropriate time-point after induction) were removed, centrifuged (1 min, 14000 rpm, bench centrifuge) and the pellets resuspended in 500 µl EST buffer. A few grains of lysozyme

(from chicken egg white, lyophilised, 48000 U/mg, Sigma) were added to each sample and incubated on ice for 15 min. Cells were then sonicated (5 x 4 second bursts at 20 W with a cooling period between each burst). The samples were centrifuged (as before) and 15  $\mu$ l of the supernatant removed and combined with 5  $\mu$ l 5 x SDS sample buffer. The pellet was resuspended in the original volume (500  $\mu$ l) of EST buffer and 15  $\mu$ l removed and combined with 5  $\mu$ l 5 x SDS sample buffer. The samples (along with 10  $\mu$ l of prestained protein standards) were boiled for 5 min before loading onto an SDS-polyacrylamide gel.

#### EST Buffer

20 mM EDTA

0.5 M Sucrose

25 mM Tris.HCl, pH 7.4

### 2.31 SDS-Polyacrylamide Gel Electrophoresis of Proteins (SDS-PAGE)

#### SDS-Polyacrylamide Gel Preparation

A Mini-Protean II gel rig (BIO-RAD) was used during the electrophoresis of protein samples. The rig was assembled and used according to the manufacturers instructions. Resolving gel was added, between the gel plates, until the level reached approximately 1 cm below where the teeth of the comb would rest. The gel was immediately overlaid with a solution of saturated butanol and left to polymerise for 45 min. The overlay was rinsed off with H<sub>2</sub>O and a comb pushed down between the plates. Stacking gel was added on top of the resolving gel, until the comb teeth were completely covered, and left to polymerise for 45 min. The comb was then removed and the wells rinsed with distilled H<sub>2</sub>O before transferring the gel sandwich from the casting stand into the

buffer chamber of the rig. 1 x reservoir buffer was poured into the upper and lower buffer chambers before loading the samples into the wells. Electrophoresis was carried out at 150 V until the required protein separation had been achieved (visualised by the migration of the dye).

### Staining of Proteins with Coomassie Brilliant Blue

After electrophoresis the gel was immersed in a solution of 0.025 % Coomassie Brilliant Blue R250 (Sigma) in 40 % methanol and 10 % acetic acid. The gel was shaken gently on a plate shaker in the stain for 60 min before transferring to destaining solution (40 % methanol, 10 % acetic acid). The gel was shaken gently overnight in destaining solution to allow a thorough removal of any unbound stain.

The gel could then be stored indefinitely in distilled H<sub>2</sub>O, in a sealed container, without any diminution in the intensity of staining.

### Resolving and Stacking Gels (10 ml)

<u>Stock Solution</u>	<u>Stacking Gel</u> (Volume Added)	<u>Resolving Gel</u> (Volume Added)
30 % Acrylamide*	1.3 ml	3.32 ml
Stacking Gel Buffer	2.5 ml	—
Resolving Gel Buffer	—	2.5 ml
10 % SDS	100 µl	100 µl
10 % APS	50 µl	50 µl
TEMED	10 µl	10 µl
H <sub>2</sub> O	6.04 ml	4.02 ml

\* Acrylamide:bisacrylamide (30 %: 0.8 %) (Scotlab EASIGel)

Resolving Gel Buffer: 1.5 M Tris.HCl, pH 8.8

Stacking Gel Buffer: 0.5 M Tris.HCl, pH 6.8

10 x Reservoir Buffer

<u>Component</u>	<u>Amount per 500 ml</u>
Tris.HCl base	15.15 g
Glycine	72 g
SDS	5 g

The volume was made up to 500 ml with distilled H<sub>2</sub>O.

**2.32 HPLC Analysis of Molybdopterin Form A Dephospho and Compound Z Levels in *A. nidulans* Oxidised and Dephosphorylated Cell Free Extracts**

Form A dephospho and compound Z levels were analysed according to the method described by Johnson and Rajagopalan (1987a,b) with a few modifications.

Extraction of Molybdopterin Form A Dephospho and Compound Z from *A. nidulans* Cells

A conidial suspension in 20 ml saline Tween 80 (of approximately  $5 \times 10^6$  conidia per ml) was used to inoculate 400 ml minimal medium (plus vitamins) supplemented with the appropriate nitrogen source at 10 mM. Cells were grown for 16 hr at 30 °C and 250 rpm, and harvested by filtration over sterile Miracloth. After washing with 50 ml sterile distilled H<sub>2</sub>O the cells were pressed dry and frozen in liquid nitrogen. 1.5 g frozen cell material was homogenised by

sonication (5 x 15 second bursts at the maximum output for the tip with a 10 sec cooling period between bursts) in 3 ml 100 mM Tris.HCl buffer (pH 7.2) and the cell debris removed by centrifugation (20 min at 14000 rpm in a Sorvall RC-58 refrigerated superspeed centrifuge with an SS-34 rotor for 15 min at 4 °C). A 50 µl sample of the resulting supernatant was removed for protein determination. 1 ml of the supernatant was combined with 125 µl I<sub>2</sub>/KI (1%/2%) in 1 M HCl and then left in darkness for 10 h at room temperature. 138 µl 1% ascorbic acid and 0.5 ml 1 M Tris.HCl were added before centrifugation at 14000 rpm for 10 min at room temperature (bench centrifuge) to remove any particulate matter. The supernatant was mixed with 13 µl 1M MgCl<sub>2</sub> and 2 units alkaline phosphatase (from bovine intestine, Sigma), and incubated for a further 12 h in darkness at room temperature. The samples were subsequently applied to QAE Sephadex A-25 (Sigma) columns (acetate form, 0.5 ml bed volume) and washed with 5 ml distilled water. Form A dephospho was eluted with 5 ml 10 mM acetic acid, adjusted to pH 7.0 with NH<sub>4</sub>OH and stored frozen until HPLC analysis. Compound Z was eluted with 8 ml 10 mM HCl. The HCl eluates were applied to Florisil (Sigma) columns (250 mg Florisil mesh 100-200, washed with 12 ml 10 mM HCl, bed volume 0.6 ml). The columns were then washed with 1 ml 10 mM HCl and compound Z was eluted with 1.7 ml 22.5 % acetone. The acetone eluates were roto-evaporated until dry. The compound Z samples were dissolved in 300 µl H<sub>2</sub>O, of which 100 µl was injected for HPLC analysis.

### HPLC Analysis

Analysis of Form A dephospho and compound Z by reversed phase HPLC was performed using a Hypersil ODS column (250 x 4.6 mm, 5mm). Form A dephospho (500 µl of the acetic acid eluate) was eluted with 10 % methanol, 50 mM ammonium acetate (pH 6.7, 1 ml/min)

and detected using a Shimadzu RF-551 fluorescence detector, set to 370/450 nm (emission/excitation). Compound Z was eluted with 5 % methanol, 50 mM triethylammonium acetate (pH 7.0, 1 ml/min) and detected by fluorescence at 350/450 nm (emission/excitation).

### Determination of Protein Concentration

The concentration of protein (total) was estimated using the BCA (bicinchoninic acid) Protein Assay Kit (Pierce), following the manufacturer's instructions, with bovine serum albumin as standard. This system combines the biruet reaction (of protein reducing  $\text{Cu}^{2+}$  in an alkaline medium to produce  $\text{Cu}^{1+}$ ) with a detection reagent for  $\text{Cu}^{1+}$ , namely bicinchoninic acid. The purple reaction product formed by the intercalation of two molecules of BCA with one cuprous ion ( $\text{Cu}^{1+}$ ) is water soluble and exhibits a strong absorbance at 562 nm, allowing the spectrophotometric quantitation of protein in aqueous solution.

### 2.33 Source of Plasmids

The cloning vector pUC19 was purchased from Pharmacia. Dr. S.E. Unkles, Monash University, provided the protein expression vector pET21a. The autonomously replicating plasmids, pHELP1, pDHG25 and ARp1 used to increase transformation frequency in *A. nidulans* were supplied by Dr. A.J. Clutterbuck, Glasgow University. The *A. nidulans* *argB* bank constructed in plasmid pFB39, which carries the *argB* gene as a *SalI* partial digest in pUC8, was donated by Dr. M. Penalva, Centro de Investigaciones Biológicas, Madrid. Plasmids pTPR3 (*mogA*-containing plasmid), pJR3 (*moeA*-containing plasmid) and pEM345 (*moaD*-containing plasmid) were obtained from Professor D.H. Boxer, University of Dundee.

The structures of plasmids pUC19, pET21a, pHELP1, pDHG25 and ARp1 used in this study are detailed in the Appendix.

#### **2.34**      **Chemicals**

Unless otherwise stated chemicals and solvents were obtained from either BDH or Sigma, and were AnalaR grade. For use in media preparation, peptone from casein was obtained from Fluka, yeast extract was purchased from Amersham and casein hydrolysate and bactotryptone were supplied by Oxoid. The chemicals methanol, glacial acetic acid, acetone, triethylamine and ammonium acetate used in HPLC were obtained from BDH and were HiperSolv grade.

#### **2.35**      **Containment and Safety**

All experimental procedures undertaken during this research were carried out in accordance with the Health and Safety Guidelines issued by the School of Biological and Medical Sciences, University of St Andrews and the Genetic Manipulation Advisory Council (GMAC).

## CHAPTER 3

### ISOLATION AND MOLECULAR CHARACTERISATION OF THE *cnxG* GENE

#### 3.1 Brief Introduction

This chapter deals with the isolation of the *cnxG* gene from the *A. nidulans* genome and the characterisation of this gene at the molecular level. The over-expression of the CnxG protein is also described. No temperature sensitive *cnxG* mutants have been identified (MacDonald and Cove, 1974) and as a result of this and the lack of any other data pertaining to *cnxG* mutant strains, the possible function of the CnxG protein is totally unclear.

#### 3.2 Initial Isolation of *cnxG* from the *A. nidulans* Genome

The strategy employed to isolate the gene *cnxG* was to use genomic wild-type *A. nidulans* DNA to phenotypically complement an *A. nidulans* *cnxG* mutant strain. The mutant strain *cnxG4* was tested in transformation experiments using either *Sau3AI* partially digested wild-type genomic DNA (average fragment size around 10 kb) or an *A. nidulans* *argB* genomic DNA bank (undigested), constructed in the pUC8-based pFB39, which carries the *argB* marker gene and contains an average insert size of 5.8 Kb (M. Penalva, unpublished). The genomic DNA was cotransformed with one of the following autonomously replicating vectors in each experiment: pHELP1 (*Bam*HI endonuclease digested); ARp1 (undigested); pDHG25 (undigested). These plasmids were added in an attempt to increase the frequency of transformation and in the case of the partially digested wild-type DNA, to create an 'instant gene bank' (Gems *et al.*, 1993). Transformant colonies were selected by

their ability to utilise nitrate as sole nitrogen source due to reconstituted nitrate reductase activity.

Approximately 1  $\mu\text{g}$  genomic DNA (*Sau3A* partial digest or *argB* bank) and 4  $\mu\text{g}$  of the chosen autonomously replicating vector DNA were used in the initial transformation experiments with strain *cnxG4*. Along with each experiment, a negative control was provided by performing a transformation without any DNA addition or with the addition of pHELP1 alone. Any colonies appearing on the selection plates were transferred to minimal medium plates, containing nitrate as sole nitrogen source to maintain the selective conditions. Total DNA (genomic and plasmid) was extracted from each transformant and approximately 0.5-3.0  $\mu\text{g}$  used to transform competent *E. coli* DH5 $\alpha$  cells in order to 'rescue' any plasmids present within the sample. Plasmid DNA was subsequently prepared from any transformed (ampicillin resistant) *E. coli* colonies. Where a large number of *E. coli* transformants (40-200 colonies) resulted, approximately 20 representative colonies were picked at random to prepare plasmid DNA. The 'rescued' plasmids were used to retransform *A. nidulans* strain *cnxG4*, either individually or in pools. Typically, 0.5-2.0  $\mu\text{g}$  'rescued' plasmid DNA and 4  $\mu\text{g}$  pHELP1 were used in each retransformation experiment. Table 1 details the results of the experiments carried out to isolate *cnxG*, by transformation and phenotypic complementation of *A. nidulans* strain *cnxG4*.

In only one experiment was a 'rescued' plasmid able to repeatedly complement the *cnxG4* mutant strain at high frequency. This plasmid, designated plasmid K (pK), resulted from an original transformation using the undigested genomic DNA bank (containing the *argB* marker) cotransformed with *Bam*HI digested pHELP1. The following number of transformant colonies were recorded in 4 separate repeated experiments using 2  $\mu\text{g}$  pK and 4  $\mu\text{g}$  pHELP1 in each case:

experiment 1, 363 colonies (Figure 7); experiment 2, 68 colonies; experiment 3, 287 colonies and experiment 4, 94 colonies. (Refer to section 3.6 for PCR analysis of several pK transformant colonies). Several (*i.e.* 2-20) colonies were noted in four of the other retransformation experiments (Table 1) however, these results could not be repeated in subsequent experiments with the same 'rescued' DNA samples.

The generation of 'rescued' plasmids for retransformation was carried out by J. Nieder, Braunschweig Technical University, Germany.

### **3.3 Molecular Analysis of pK**

Plasmid K was found to be approximately 18.0 kb in length. Southern analysis was carried out to determine which (if any) fragments of pK contained genomic *A. nidulans* DNA, free of vector sequences. Various restriction digests of pK (*Bam*HI, *Eco*RI, *Hind*III, *Kpn*I, *Sal*I and *Xba*I) were probed with the *A. nidulans argB* gene sequence, the bacterial AMA1 sequence and the bacterial pUC19 sequence. These 3 probes covered all possible vector sequences which could have been incorporated (from the *argB* bank or pHELP1) into the cointegrate pK by recombination events during the transformation procedure. A 2.2 kb *Bam*HI fragment of pK did not hybridise to any of the vector probes and thus was assumed to contain only genomic *A. nidulans* DNA derived from the *argB* bank. All of the other pK fragments hybridised to at least one of the probes.

### **3.4 Molecular Analysis of the 2.2 kb BamHI Fragment of pK**

It was necessary to determine whether the 2.2 kb *Bam* HI fragment of pK contained whole, part or none of the gene of interest *i.e.* *cnxG*. The 2.2 kb fragment did not appear to contain the entire gene as a

Table 1: Isolation of *cnxG* by transformation of mutant strain *cnxG4* with *A. nidulans* genomic DNA

Experiment Number	DNA Material Used <sup>a</sup>	Number of Initial <i>cnxG4</i> Transformant Colonies	<i>cnxG4</i> Transformant Colony Designation	Number of <i>E. coli</i> Transformant Colonies <sup>b</sup>	Number of <i>cnxG4</i> Retransformant Colonies <sup>bc</sup>
1	Genomic DNA ( <i>Sau3AI</i> ) + pHELP1	6	1.1	-	N.A.
			1.2	15	-
			1.3	50	(2)
			1.4	-	N.A.
			1.5	120	(10)
			1.6	40	(20)
2	Genomic DNA ( <i>Sau3AI</i> ) + pHELP1	4	2.1	-	N.A.
			2.2	-	N.A.
			2.3	100	-
			2.4	43	-
3	Genomic DNA ( <i>Sau3AI</i> ) + ARp1	3	3.1	-	N.A.
			3.2	200	-
			3.3	150	(3)
4	<i>argB</i> bank + pHELP1	3	4.1	1	200
			4.2	5	-
			4.3	7	-
5	<i>argB</i> bank + pDHG25	1	5.1	17	-

<sup>a</sup>The structures of plasmids pHELP1, ARp1 and pDHG25 are detailed in the Appendix.

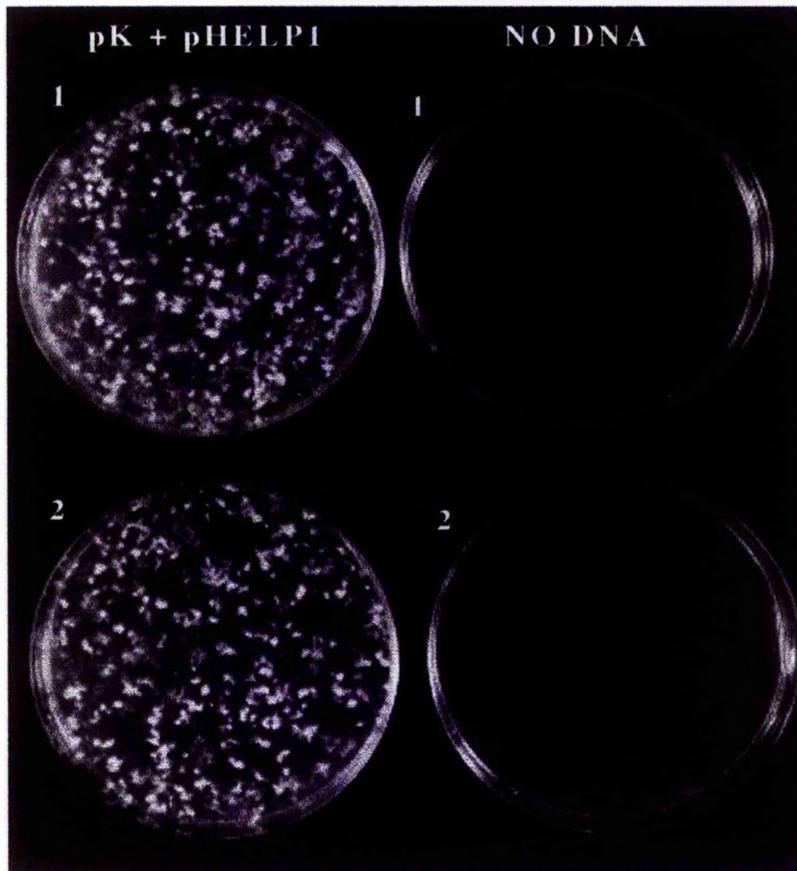
The *argB* bank is described in section 3.2.

<sup>b</sup>A straight line (-) indicates that no transformant/retransformant colonies were obtained.

<sup>c</sup>N.A. refers to not applicable.

The numbers in brackets refer to retransformant colonies which were observed in initial retransformation experiments but which were not present in repeated experiments using the same 'rescued' DNA samples.

Figure 7: Complementation of strain *cnxG4* with pK



Protoplasts prepared from strain *cnxG4* were transformed with 2  $\mu\text{g}$  pK and 4  $\mu\text{g}$  pHELP1. The transformed protoplasts were divided into two, duplicate aliquots and then plated on protoplast regeneration medium containing 10 mM nitrate as sole nitrogen source to provide the selection of only those cells with reconstituted nitrate reductase activity. Strain *cnxG4* protoplasts without any DNA addition were treated in the same manner to provide a negative control. Plates 1 and 2 represent duplicate samples. Incubation of plates was carried out for 4 days at 37  $^{\circ}\text{C}$ .

*Bam*HI digest of pK failed to complement the *cnxG* mutant strain *cnxG4* in subsequent transformation experiments using 1  $\mu$ g *Bam*HI digested pK + 4  $\mu$ g undigested pHELP1.

The *Bam*HI fragment was used to probe a Northern blot of mRNA isolated from wild-type cells and the result compared with an identical Northern blot probed with linearised pK. The 2.2 kb *Bam*HI fragment hybridised to the same two bands as the entire pK molecule, a strongly hybridising 1.6 kb band and a weakly hybridising 7.5 kb band in each case. It therefore appeared that the *Bam*HI fragment derived from pK contained part of the *cnxG* gene.

### **3.5**      Isolation of *cnxG* Cosmid Clones

Due to the possibility that DNA rearrangements may have occurred during the formation of the cointegrate pK so as to alter the *cnxG* gene sequence (without affecting the gene function) it was deemed necessary to isolate an independent genomic clone for sequencing purposes. The 2.2 kb *Bam*HI fragment of pK was used as a probe to screen the chromosome VI portion of the *A. nidulans* cosmid library, in an attempt to isolate an independent genomic cosmid *cnxG* clone (or clones). The gene *cnxG* has been mapped to chromosome VI by previous classical genetic studies, therefore only this portion of the cosmid library, containing 162 pWE15 and 114 LORIST2 cosmid clones, was screened. Five cosmid clones were found to hybridise to the probe (overleaf).

<u>Cosmid Clone Designation</u>	<u>Cosmid Number</u>
1	L9G12
2	L32D09
3	W16D10
4	W19D10
5	W28E11

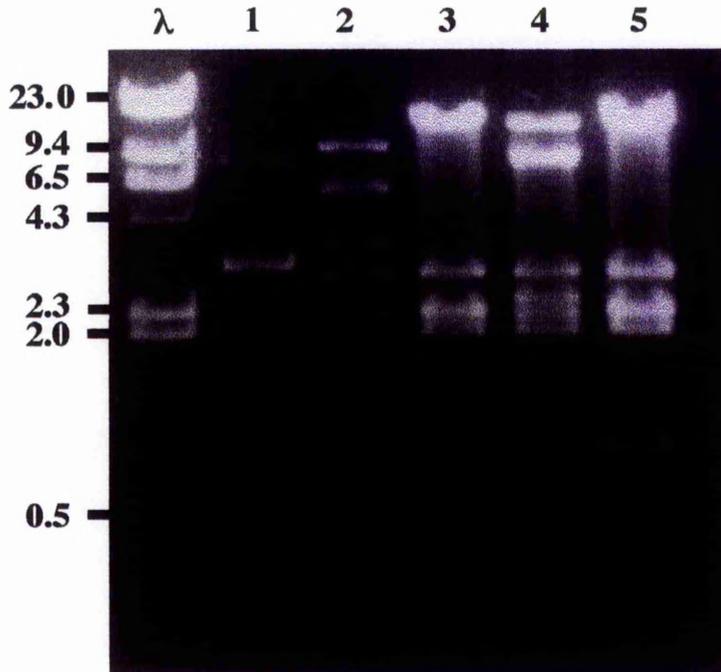
L = LORIST2 clone

W = pWE15 clone

Clones 2, 3, 4 and 5 produced similar restriction patterns when digested with the same restriction enzyme (Figure 8) suggesting an amount of overlap between the cosmids. Clone 1 was found to contain a single 3 kb band, much smaller than the approximately 30-55 kb inserts present in the other clones.

Cosmid clones 2, 3, 4 and 5 appeared able to complement the *cnxG4* mutant strain however, the transformation frequencies obtained were low in comparison to those achieved with pK. Frequencies of 1-11 transformant colonies per 2 µg cosmid DNA + 4 µg pHELP1 were recorded in repeated experiments using each of clones 2, 3, 4 and 5. Repeated transformation using 2 µg cosmid clone 1 + 4 µg pHELP1 consistently failed to produce any nitrate-utilising colonies. Various restriction digests of cosmid clones 2, 3, 4 and 5 displayed the same very low complementation frequency or no complementation when cotransformed with undigested pHELP1 and it proved difficult to determine which, if any, of the cosmid digests were definitely able to complement the *cnxG4* mutation.

Figure 8: Restriction digest pattern of the *cnxG* cosmid clones



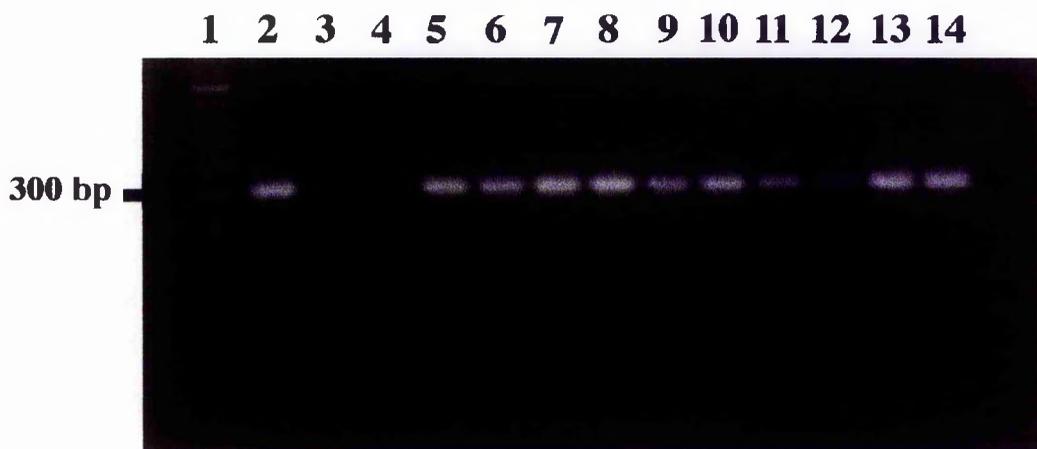
Cosmid clones 1 (L9G12), 2 (L32D09), 3 (W16D10), 4 (W19D10) and 5 (W28E11) were each digested with the restriction enzyme *Bam*HI and the resulting samples electrophoresed in a 0.8 % agarose gel. The lane containing each cosmid digest is represented by its respective clone number. The symbol  $\lambda$  indicates the lane containing lambda (digested with *Hind*III) markers. The molecular sizes (kb) of the markers are shown.

### **3.6 PCR Analysis of *A. nidulans* Colonies Transformed by pK and Cosmid Clone DNA**

To determine that the colonies resulting from transformation with pK and cosmid clone DNA were *bona fide* transformants *i.e.* that foreign vector DNA had become incorporated into the cell (either held as an autonomously replicating molecule within the cell or incorporated into the chromosomal DNA), a random sample of such colonies were tested for the presence of the *E. coli* ampicillin resistance gene *bla* (encoding  $\beta$ -lactamase) which is used for selection in a number of vectors including pHELP1, the pUC-based *argB* bank vector pFB39 (both used to create pK) and the pWE15 cosmid vector. 4 x pK, 3 x cosmid clone 4, and 3 x cosmid clone 5 transformant colonies were removed from the selection plates and total DNA prepared from each. Cosmid clones 4 and 5 were selected for further study as they both contain *A. nidulans* DNA inserted into a pWE15 vector, which contains the *E. coli* ampicillin resistance gene *bla*. DNA was also prepared from the *A. nidulans* wild-type strain to act as a negative (untransformed) control.

Each DNA sample was tested in PCR with primers specific for sequences within the *E. coli* ampicillin resistance gene. A positive control was provided by using pUC19 DNA as template. A 'no template DNA' negative control was also performed. It was calculated that successful PCR would result in the amplification of a 312 bp DNA fragment corresponding to part of the *bla* ampicillin resistance gene sequence (positions 2165-2477 in the cloning vector pUC19, accession number X02514). As can be seen from the data presented in Figure 9, each of the transformant colony DNA samples produced a single band corresponding to the expected 312 bp fragment size, with the wild-type sample and the 'no template DNA' negative controls lacking the presence of an equivalent DNA band. It therefore appeared that the transformant colonies resulting

Figure 9: PCR analysis of plasmid K, cosmid clone 4 and cosmid clone 5 *A. nidulans* transformant colonies



PCR was used to identify the *E. coli* ampicillin resistance gene (*bla* encoding  $\beta$ -lactamase) in DNA extracted from putative *A. nidulans* transformant colonies. A positive control was provided by using pUC19 DNA as template. PCR was performed without template DNA and with DNA extracted from the *A. nidulans* wild-type strain to provide negative controls. Following PCR, 10  $\mu$ l of each sample was electrophoresed in a 2% agarose gel, with 100 bp marker DNA for size determination. The position of the 300 bp marker is indicated. The samples present in each track are as follows: lane 1, 100 bp markers; lane 2, positive control (pUC19 DNA); lane 3, negative control 1 (no template DNA); lane 4, negative control 2 (wild-type DNA); lanes 5-8, DNA from pK transformants; lanes 9-11, DNA from cosmid 4 transformants; lanes 12-14, DNA from cosmid 5 transformants.

from pK, cosmid clone 4 and cosmid clone 5 complementation were 'real' in that they contained foreign vector DNA, and were not mutant revertants to wild-type or mere contaminants.

### 3.7 Isolation of *cnxG* cDNA clones

The 2.2 kb *Bam*HI fragment of pK was also used to probe an *A. nidulans* cDNA library in an attempt to isolate a *cnxG* cDNA clone(s). Using the equation below it was calculated that for the *A. nidulans* cDNA bank, with an average insert size of 2 kb, approximately 30958 plaque forming units (pfu's) would need to be screened in order to have a 99% probability of including any one gene.

$$N = \log (1 - P) / \ln (1 - f)$$

N = Number of recombinant clones (colonies or plaques/pfu's)

P = Desired probability, for 99% probability use 0.99

f = Average insert size (in bp) ÷ Genome size (31 x 10<sup>6</sup> bp for *A. nidulans*)

40000 pfu's were screened to ensure that an excess of cDNA clones were tested. Six cDNA clones were isolated and found to contain DNA inserts of different lengths ranging from 1.3 kb to 4.2 kb (overleaf).

<u>cDNA Clone Designation</u>	<u>Insert Length (kb)</u>
1	3.9
2	4.2
3	1.7
4	1.3
5	2.5
6	1.3

### **3.8 Nucleotide Sequence Analysis of the *cnxG* cDNA Clones**

cDNA clone 3 (1.7 kb) was chosen at random and completely sequenced in one DNA strand starting with the universal reverse and forward (-40) primers and working along the sense and antisense strands until the sequences met. The clone was found to contain an open reading frame (ORF) of 1235 bp which did not extend beyond the 5' end of the cDNA sequence indicating that the whole coding region of the gene was most likely included on this clone. The first methionine residue within the correct reading frame was identified as the putative translational start codon for the CnxG protein. At the 3' end of the cDNA clone a poly(A) tail consisting of 19 adenine residues was identified, 33 bp downstream of the putative translational stop codon.

After sequencing from the 3' poly (A) tail end using the universal reverse primer, the five other cDNA clones were found to be identical to clone 3 in this region. cDNA clones 4 and 6 which were smaller than clone 3 were sequenced from the 5' end with the universal forward primer. The sequence at the 5' end of the two smaller clones was identical to that contained on the equivalent portion of clone 3, thus clones 4 and 6 appeared to be truncated versions of clone 3. A primer taken from the 5' end of clone 3 was used in sequencing reactions with the three larger cDNA clones. In each of the clones the resulting sequence

was found to match that obtained for clone 3. This suggested that clones 1, 2 and 5 contained extended versions of clone 3 stretching further up past the end of clone 3 in the 5' direction.

Approximately 1 kb upstream of the 5' end of cDNA clone 3 was sequenced (in 1 strand) on each of the longer cDNA clones. cDNA clone 5 was shown to terminate towards the end of the sequenced region with clones 1 and 2 proceeding further in the 5' direction. The 1 kb upstream region contained no open reading frame and displayed no significant similarity to any other protein sequence held within the Swiss-Prot and updated Swiss-Prot data bases (determined using the BLASTX computer sequence comparison program).

A 600 bp fragment from within the sequenced 1 kb upstream region was amplified by PCR, and used as a probe in Northern analysis with a Northern blot containing mRNA extracted from wild-type *A. nidulans* cells. No RNA message was detected after incubation of the hybridised blot with fast film for 3 weeks at -70 °C.

The upstream region present on the long cDNA clones does not appear to code for any gene transcript. It therefore seems likely that the long cDNA clones can be attributed to an artefact, however it is not possible to speculate on what is contained within the further 5' upstream regions of clones 1 and 2 because these stretches have not been investigated.

### **3.9 Construction of a *cnxG* Genomic Subclone**

The 1.7 kb insert of cDNA clone 3 was used as a probe in Southern analysis in an attempt to isolate a cosmid (genomic) fragment which held the entire *cnxG* gene. Sequencing directly from the large cosmid clones would have been difficult requiring a large amount of template cosmid DNA (approximately 10-15 µg cosmid DNA per

sequencing reaction), so it was necessary to obtain a smaller subclone. It was important to obtain a genomic sequence for *cnxG* to identify the presence of any introns and to characterise the 5' untranslated region of the *cnxG* gene. Various restriction digests of one of the *cnxG* containing cosmids (cosmid clone 5) were carried out and pK DNA was also digested with the same enzymes and run alongside each of the corresponding cosmid digests to act as a positive control. The resulting gel was used to perform a Southern blot, the filter of which was hybridised with the 1.7 kb insert sequence contained within cDNA clone 3. The result of this experiment is shown in Figure 10.

The probe hybridised to at least one band of the same size within each of the pK and cosmid 5 digests, providing more evidence to suggest that pK and cosmid 5 contain equivalent pieces of genomic *A. nidulans* DNA. The labelled probe hybridised to more than one band in the *Bam*HI, *Not*I, *Bam*HI-*Not*I, *Hind*III, *Pst*I and *Xho*I cosmid (and pK) digests and to a single band in the each of the remaining cosmid 5 (and pK) digests. The following single cosmid DNA bands were highlighted: the 6.8 kb *Eco*RI fragment; the 8.2 kb *Sal*I fragment; the 3.0 kb *Eco*RI-*Sal*I fragment; the 7.4 kb *Sma*I fragment and the 14.0 kb *Xba*I fragment.

The smallest of these single hybridising bands *i.e.* the 3.0 kb *Eco*RI-*Sal*I fragment was selected and ligated into vector pUC19. The resulting genomic subclone (designated pLJM18), was able to phenotypically complement the *cnxG4* mutant strain at a high frequency comparable to that obtained with pK (see Figure 11). In a single experiment 72 nitrate-utilising colonies resulted following transformation using 1  $\mu$ g pLJM18 and 4  $\mu$ g pHELP1, compared to 20 nitrate-utilising colonies using 1  $\mu$ g pK and 4  $\mu$ g pHELP1, with the negative 'no DNA' control completely devoid of growth. pLJM18 therefore contained a

Figure 10: Identification of a *cnxG* cosmid fragment suitable for subcloning

Various different restriction digests of cosmid 5 DNA were performed and electrophoresis carried out in a 0.8% agarose gel. To act as a positive control, pK DNA was cut with the same restriction enzymes as cosmid clone 5, and run alongside each corresponding cosmid DNA sample during electrophoresis. The resulting gel was used to produce a Southern blot which was hybridised with the 1.7 kb insert of cDNA clone 3. Autoradiography was carried out for 2 h at room temperature. The following abbreviations are used: B = *Bam*HI; N = *Not*I; B/N = *Bam*HI/*Not*I double digest; H = *Hind*III; E = *Eco*RI; Sa = *Sal*I; E/Sa = *Eco*RI/*Sal*I double digest; P = *Pst*I; Sm = *Sma*I; Xb = *Xba*I; Xh = *Xho*I. In each of the restriction digests, 1 = pK DNA and 2 = cosmid 5 DNA. The molecular sizes (kb) of the lambda (digested with *Hind*III) markers are indicated. The 3.0 kb *Eco*RI/*Sal*I fragment which was picked for subcloning is highlighted by an arrow.

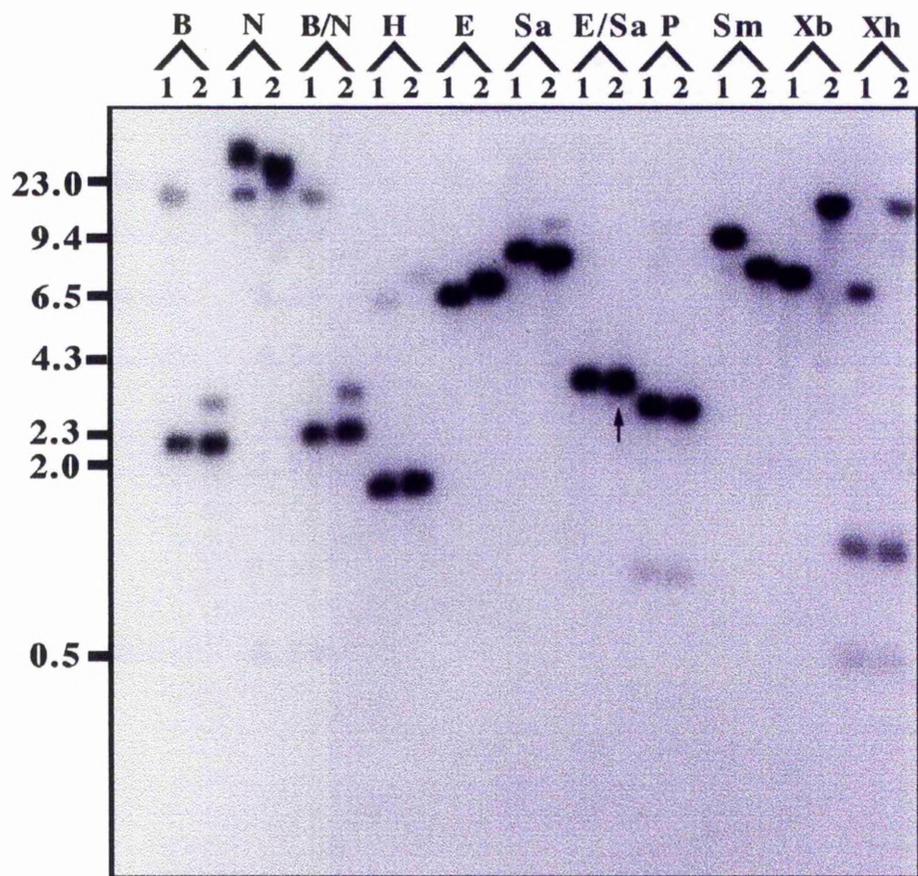
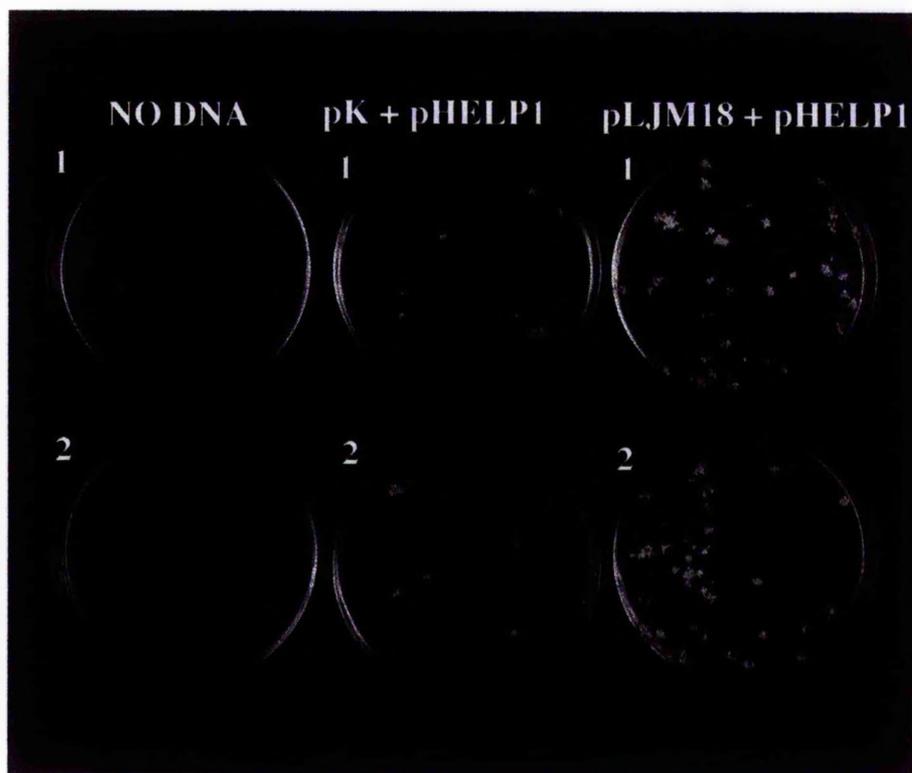


Figure 11: Complementation of strain *cnxG4* by pLJM18



Protoplasts prepared from *A. nidulans* strain *cnxG4* were transformed with 1  $\mu\text{g}$  pLJM18 and 4  $\mu\text{g}$  pHELP1. The transformed protoplasts were divided into two, duplicate portions and plated on protoplast regeneration medium containing 10 mM nitrate as sole nitrogen source to provide the selection of only those cells with reconstituted nitrate reductase activity. Strain *cnxG4* protoplasts were transformed with 1  $\mu\text{g}$  pK and 4  $\mu\text{g}$  pHELP1 to provide a positive control. Strain *cnxG4* protoplasts transformed without any DNA addition provided a negative control. Plates 1 and 2 represent duplicate samples. Incubation of plates was carried out for 4 days at 37  $^{\circ}\text{C}$ .

3.0 kb fragment of genomic *A. nidulans* DNA able to complement the *cnxG4* lesion.

### **3.10**     **Structure of the *cnxG* Gene**

The genomic subclone pLJM18 was sequenced using primers which were (in some cases) deduced from the cDNA clone. Primers were chosen to allow sequence analysis of the opposite DNA strand to that already obtained from the cDNA clone. The subclone was sequenced, upstream of the ORF proposed to be *cnxG*, until the *EcoRI* restriction site of the vector was reached. No other ORF was identified in this region. Figure 12 details the *cnxG* genomic sequence obtained.

Comparison of the genomic and cDNA clone sequences allowed the identification of two introns, a 52 bp intron present in the 5' untranslated region of the gene (starting at nucleotide position -302, where A in the translation initiation codon = +1) and a second 45 bp intron located within the *cnxG* gene coding region (starting at nucleotide position +162). Both introns display good consensus sequences for 5' and 3' splice sites. The intron present in the 5' non-coding region has a potential internal consensus sequence for lariat formation *i.e.* AACTGAC (starting at nucleotide position -296 bp). An obvious motif for lariat formation could not be identified in the other intron. Although few fungal introns follow the yeast consensus sequence for lariat formation (TACTAAC) exactly, most (92%) have CT at positions 3 and 4 of the element and A at position 6 (Unkles, 1992 and references therein).

The *cnxG* coding region (minus the intron) was found to be 396 amino acids in length with a predicted protein weight of 43.56 kDa calculated using the equation indicated overleaf.

Figure 12: Nucleotide sequence and inferred amino acid sequence of the *cnxG* gene

The genomic *cnxG* clone pLJM18 was sequenced in a single strand, opposing that already obtained from sequence analysis of the *cnxG* cDNA clone. Introns and non-coding regions present on the genomic clone were sequenced in both directions.

The numbers to the left of the sequence indicate nucleotides and the numbers to the right of the sequence indicate amino acids. Introns are presented in lower case letters. Italics have been used to denote the 5' and 3' splice sites and the site of lariat formation within the introns. The possible candidates for CAAT boxes are overlined. The pyrimidine-rich sequences are underlined. The translational stop codon is indicated by an asterisk. The single arrow head distinguishes the position of poly(A) tail addition and the position of the stop codon which defines the reading frame is indicated by three arrowheads. (Note that Figure 12 is contained on two sequential pages).

-572 GAATTCAGATTGTTTCGTCAGGACAAGGTCGTGCAGCTACTAGCATTCCCTCAACGACTTTA  
-512 GTCATGGCAAATGCTTGAACCTTCCTCCTGAAGAGCACCGACACTTTCGAAGCTTTC AAC  
-452 GCTCCGAAAATACGGAGTCCGCAITGTTGACGCCAAGTTTCGCGTTGCCAAAGACCGATG  
-392 ATGACCCCTCTTCCTTTTCTCAATCTCGACATGCCGGAGTATCCTAGCGAACACGACG  
-332 ATATTACCATTCCTTTTGATACGGAGGCTGgtgagtgtggtgacgaaccataagact  
-272 taaaaotgacgacagagatcagATCGATTCAACTTCCAATCAGCAGCGCCTGCCTCCGTC  
-212 CGCGAGCCCTCACGAATCGCGTCCCTGCGGCGGTAGAGGTTTAGAATCAATCAGTTGCTG  
-152 GAATTGCTAGCAGGCTTGTCTCTCGAAGCTCGGTGGCTCGTCGGGCACGCAATCTCCA  
-92 CTGGACCCGCCGGTACGCTGATTTATCCCTCATTATGAACCGTCTTCGGGTTCTCGCC  
-32 CTCTTCCCTCTTTAACCCACAGTCAAACGCAATGCTGCCAGAAGGCCAGGACGTTGGCG  
M L P E G Q D V G A 10  
28 CCTCGACGCGCTCACTGCATGCAGACGATGTGCTAAATGTCGTCACCTGATGTTGCCCTC  
S T R S L H A D D V L N V V T D V A P P 30  
88 CTATTTCATGTTGCGACCACCTTCCGCTATTCCGATGACCCGCCAAGTTGATGCCGGCGG  
I H V A T T F R Y S D D P A K L M P A A 50  
148 CAGATCTCAATGGGgtaagtatccgaaaatcattccttttcagagggttaacogcgagA  
D L N G T 55  
208 CCCTCGAGGAGCAATCGTCATATATCTACTCGCGTCTCACTGCGCCAACCTACTACCCGAT  
L E E Q S S Y I Y S R L T A P T T T R F 75  
268 TCGAAGCCATTCCTTCGTCCTCCCAACGGCCAACCATCAGTTACTCATCGGGGCTGT  
E A I L S S L L N G Q T I S Y S S G L S 95  
328 CGGCATTCCATGCGGCTCTCACCTTCTAAACCTCGTAGGATCTCGATCGGCAAAGGAT  
A F H A A L T L L N P R R I S I G K G Y 115  
388 ATCATGGATGCCACGGTGTGATTGACATTTTCAACCGTCTGAATGGACTTCAGAAGCTGG  
H G C H G V I D I F N R L N G L Q K L D 135  
448 ATTTGGAITGCCAGCCGAACAGCTGGAGGCTGGGGATGTCATACACTTGGAACTCCAG  
L D C P A E Q L E A G D V I H L E T P V 155  
508 TCAATCCTGAGGGGACCTCTTTCAATATTGAGCAATATGCCAAGAAAGCACACTCTCGAG  
N P E G T S F N I E Q Y A K K A H S R G 175  
568 GAGCTTATCTGATAGTTGACGGAACTTTTCGCCCGCCGCTCTGCAAGACCCGTTCAAAT  
A Y L I V D G T F A P P P L Q D P F K F 195  
628 TTGGTGTGACCTGGTGTGCTGCACTCTGGATCCAAATACTTTGGAGGGCACAGGATGTCT  
G A D L V L H S G S K Y F G G H S D V L 215  
688 TGTGCGGAGTTCCTGGCCACGCAGAACAAGGACTGGGCGCAGCAATTGCTCAGAGACCCGG  
C G V L A T Q N K D W A Q Q L L R D R V 235  
748 TTTTCCTTGGTGGTGTTCATGGGCAATCTGGAGGGCTGGCTCGGCGTCCGGAGTTTCGGGA  
F L G G V M G N L E G W L G V R S L R T 255  
808 CTCTGGAGGTTTCGTGTGCAACGGCTGAGCCAGAACGCAACGAATTTGGTCTTGTGGCTTC  
L E V R V Q R L S Q N A T N L V L W L H 275  
868 ACAACGCGCTGCAGACTCCAAGCCCCGACCCGGCAGCGATGAAGAGGGCAGCGAGAAAG  
N A L Q T P S P A P G S D E E A T Q K V 295  
928 TGCTGGAGCAGGTTTTCACCTCCAGCCTTCAAAGGACGACGAGTCTGGCTCCTGAAGC  
L E Q V F H S S L Q K D D E S W L L K Q 315

Figure 12 continued

988 AGATGCCGAACGGGTTTCGGGCCTGTCTTCTCGATCACCATGAGAGAAGAAGACTATGCCG 335  
M P N G F G P V F S I T M R E E D Y A R

1048 GCCACTTGCCAAGCAAGCTTGCCCTGTTCAGCACGCGACCAGTCTGGGTGGCGTAGAGT 355  
H L P S K L A L F Q H A T S L G G V E S

1108 CGTTGATTGAATGGCGGACAATGTCGGACAAGACGGTGGACCGACGGTTGCTGCGAGTTA 375  
L I E W R T M S D K T V D R R L L R V S

1168 GCATTGGGTTGGAGAACTGGGAGGATTTGAAGCGGGATCTGGTGAATGCTTTTAGAGCAC 395  
I G L E N W E D L K R D L V N A F R A L

1228 TGGTTTAGATACACCTATATATGAATGACTTACGTCGCGTACCGGATCTTGAGTAAACAG 396  
V \* ^

1288 AGTACCTGGTAGAAAGTCTACTAAGATAAGATTTCCCAATGAAGATCCCCACACCCCACT

1348 TTACCCAACCT

Number of amino acids x 110 = Predicted protein molecular  
weight (in Daltons)

Where 110 is the calculated average molecular weight of an amino acid residue (in Daltons).

As previously described, the translation initiation codon was identified as the first methionine within the open reading frame. The presence of a purine at nucleotide position -3 (*i.e.* GCAATG where A in the translation initiation codon = +1) is in agreement with most (84%) fungal translational start sites (Unkles, 1992).

It was apparent that the sequences present in cDNA clones 1, 2 and 5 extended past the 5' end of the genomic sequence contained within pLJM18. This result was unexpected because the cDNA clones should begin at a transcriptional start site downstream of promoter sequences contained only on the genomic clone.

Without formal identification of the transcriptional start site, by primer extension analysis, it is difficult to speculate on the location of promoter elements for the *cnxG* gene. There are three possible candidates for CAAT boxes located within the 572 bp 5' upstream non-coding region contained on pLJM18. These are located at nucleotides -168 (TCAAT), -237 (CCAAT) and -371 (TCAAT). No TATA motifs similar to the conserved sequence TATAAA can be found within the nucleotide stretch of the 5' non-coding region which has been sequenced, however several other examples of fungal genes which lack a conserved TATA sequence have been identified including the *niaD* gene from *A. niger* and *ura3* and *ura5* from *T. reesei* (Unkles, 1992). CT boxes or pyrimidine-rich sequences have been described in the promoters of yeast and filamentous fungal genes (Unkles, 1992) however the significance of these

sequences is unclear. These CT-rich motifs often occur immediately before the major transcription start point but can occur at any point in the 5' upstream region. There are two pyrimidine-rich tracts within the 5' non-coding region which has been sequenced, one of 22 bp starting at nucleotide -40 and one of 19 bp starting at nucleotide -388.

The possibility that the promoter elements for *cnxG* are located further upstream past the 5' limits of pLJM18 seems unlikely. The fact that the subclone complements the *cnxG* mutant suggests that the complete gene including promoter elements is contained within this DNA segment.

In the 3' non-coding region the position of the poly(A) tail was identified, 33 nucleotides downstream of the translation stop codon. A polyadenylation signal (AATAAA) could not be identified between the translational stop codon and the point of poly(A) tail addition, however this is not unusual because many fungal genes lack this consensus sequence (Unkles, 1992).

Sequencing of the *cnxG* gene was carried out in collaboration with J. Nieder, Braunschweig Technical University, Germany, who subcloned recomplementing fragments of pK into pBluescript and carried out retransformation experiments which restricted the sequence to a 2.1 kb *EcoRV* fragment (J. Nieder, personal communication). Subsequent analysis of this 2.1 kb *EcoRV* fragment by J. Nieder revealed that it contained a sequence which matched that produced from experiments carried out in this study.

### **3.11 CnxG Protein Sequence Similarities**

The full length *cnxG* genomic sequence was assessed for similarities to protein sequences held within the Swiss-Prot and updated Swiss-Prot data bases using the BLASTX computer sequence comparison program. Table 2 displays the various protein similarities to the CnxG

Table 2: Protein similarities to *cnrG*

Position	Organism	Protein	Known or Postulated Protein Function	% Identity	% Similarity	No. of Amino Acids <sup>a</sup>	Reference(s)
1	<i>Saccharomyces cerevisiae</i>	Yhr2	Hypothetical lyase Similarity to trans-sulfuration enzymes Specific function unknown	48.0	68.6	378	Johnston <i>et al.</i> , (1994)
2	<i>Escherichia coli</i>	MetB	Cystathionine Gamma-Synthase (O-succinylhomoserine (Thiol)-Lyase)	33.6	54.4	386	Duchange <i>et al.</i> , (1983)
3	<i>Helicobacter pylori</i>	MetB	Cystathionine Gamma-Synthase (O-succinylhomoserine (Thiol)-Lyase)	33.4	58.4	380	Tomb <i>et al.</i> , (1997)
4	<i>Pseudomonas putida</i>	MegL	Methionine Gamma-Lyase	33.3	58.5	398	Inone <i>et al.</i> , (1995); Nakagana <i>et al.</i> , (1988)
5	<i>Saccharomyces cerevisiae</i>	Cys3	Cystathionine Gamma-Lyase	32.5	55.5	393	Ono <i>et al.</i> , (1992); Yamagata <i>et al.</i> , (1993); Barton <i>et al.</i> , (1993); Onellette <i>et al.</i> , (1993); Ono <i>et al.</i> , (1993)
6	<i>Herpetosiphon aurantiacus</i>	MetB	Probable Cystathionine Gamma-Synthase (O-succinylhomoserine (Thiol)-Lyase)	32.0	54.9	319	Duesterhoeft <i>et al.</i> , (1991)
7	<i>Arabidopsis thaliana</i>	MetB	Cystathionine Gamma-Synthase (O-succinylhomoserine (Thiol)-Lyase)	31.8	57.0	563	Le Guen <i>et al.</i> , (1994)
8	<i>Mycobacterium leprae</i>	MetB	Cystathionine Gamma-Synthase (O-succinylhomoserine (Thiol)-Lyase)	31.4	54.4	388	Smith, D.R. and Robinson, K., (1994) -unpublished
9	<i>Haemophilus influenzae</i>	MetB	Cystathionine Gamma-Synthase (O-succinylhomoserine (Thiol)-Lyase)	30.4	57.1	369	Fleischmann <i>et al.</i> , (1995)
10	<i>Saccharomyces cerevisiae</i>	Ygt4	Hypothetical lyase Similarity to trans-sulfuration enzymes Specific function unknown	30.4	52.3	465	Cogliovina <i>et al.</i> , (1997)
11	<i>Homo sapiens</i>	Cgl	Cystathionine Gamma-Lyase	29.9	50.8	405	Lu <i>et al.</i> , (1992)

Position	Organism	Protein	Known or Postulated Protein Function	% Identity	% Similarity	No. of Amino Acids <sup>a</sup>	Reference(s)
12	<i>Saccharomyces cerevisiae</i>	Met17	O-acetylhomoserine Sulphydrylase	29.6	51.9	443	Kerjan <i>et al.</i> , (1986); Ono <i>et al.</i> , (1993); Norbeck and Blomberg, (1996)
13	<i>Caenorhabditis elegans</i>	Cgl	Putative Cystathionine Gamma-Lyase	28.9	55.6	392	Wilson <i>et al.</i> , (1994)
14	<i>Streptomyces coelicolor</i>	CysA	Putative Cystathionine Gamma-Lyase	28.8	53.4	392	Li and Strohl, (1996)
15	<i>Pseudomonas aeruginosa</i>	MetZ	O-succinylhomoserine Sulphydrylase	28.5	54.2	403	Fogliano <i>et al.</i> , (1995)
16	<i>Aspergillus nidulans</i>	CysD	O-acetylhomoserine Sulphydrylase (Homocysteine Synthase)	28.2	50.1	437	Sienko, M., Topczewski, J. and Paszewski, A., (1995) -unpublished
17	<i>Rattus norvegicus</i>	Cgl	Cystathionine Gamma-Lyase	28.0	51.1	398	Nishi <i>et al.</i> , (1994); Erickson <i>et al.</i> , (1990)
18	<i>Bordetella avium</i>	MetC	Cystathionine Beta-Lyase	26.6	50.5	395	Gentry-Weeks <i>et al.</i> , (1993)
19	<i>Arabidopsis thaliana</i>	MetC	Cystathionine Beta-Lyase	26.4	49.5	464	Ravenel <i>et al.</i> , (1995)
20	<i>Escherichia coli</i>	MetC	Cystathionine Beta-Lyase	25.8	49.9	395	Belfaiza <i>et al.</i> , (1980); Clausen <i>et al.</i> , (1996)
21	<i>Saccharomyces cerevisiae</i>	MetC	Putative Cystathionine Beta-Lyase	25.5	51.1	340	Murakami <i>et al.</i> , (1995); Eki <i>et al.</i> , (1996)
22	<i>Salmonella typhimurium</i>	MetC	Cystathionine Beta-Lyase	24.6	45.9	395	Park and Stauffer, (1989)
23	<i>Haemophilus influenzae</i>	MetC	Cystathionine Beta-Lyase	24.0	47.9	396	Fleischmann <i>et al.</i> , (1995)
24	<i>Saccharomyces cerevisiae</i>	MetX	Putative Cystathionine Gamma-Synthase (O-succinylhomoserine (Thio)-Lyase)	20.7	44.2	649	Gentles, S., Bowman, S., Barrell, B.G. and Rajandream, M.A., (1994) -unpublished

The protein sequences are listed in order of decreasing % identity. The % similarity values for each protein are also indicated.

<sup>a</sup>Number of inferred amino acids given for each protein.

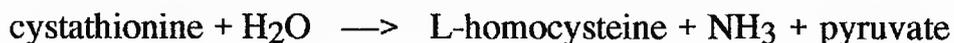
amino acid sequence in decreasing order of % identity (designated 1 to 24, where 1 represents the most identical sequence). The % identity and % similarity values were obtained by comparing each of the protein sequences with the deduced CnxG amino acid sequence using Bestfit alignments on the Seqnet computer program.

The protein Yhr2 from the yeast *Saccharomyces cerevisiae* was most similar to CnxG with 48.0% identity (68.6% similarity) exhibited between the 2 sequences. The specific function of Yhr2 is unknown however it is thought to be a lyase with similarity to enzymes involved in the catalysis of trans-sulphuration reactions (Johnston *et al.*, 1994). The protein Ygt4 from *S. cerevisiae*, which also has similarity to transsulphuration enzymes and an unknown specific function, was less similar to CnxG exhibiting only 30.4% identity (52.35% similarity). The MetB proteins from *Escherichia coli*, *Helicobacter pylori*, *Herpetosiphon aurantiacus*, *Arabidopsis thaliana*, *Mycobacterium leprae* and *Haemophilus influenzae* and the protein MetX from *S. cerevisiae* all displayed similarity to CnxG. MetB from *E. coli* and MetB from *H. pylori* were placed as second and third most similar to CnxG with identity values of 33.6 and 33.4% respectively. By contrast, the *S. cerevisiae* MetX protein produced only 20.7% identity to CnxG, which was the lowest value, placed at position 24 in Table 2. These proteins all encode the enzyme cystathionine gamma-synthase or o-succinylhomoserine (thiol)-lyase which catalyses the following reaction:



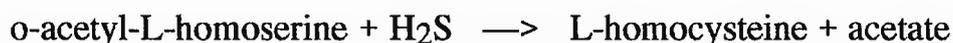
This reaction occurs during methionine biosynthesis (and catalyses the transfer of sulphur derived from L-cysteine to o-succinyl-L-homoserine to produce cystathionine, a precursor of methionine. In many organisms

the protein MetC catalyses the subsequent step in methionine biosynthesis *i.e.* the conversion of cystathionine to L-homocysteine, as detailed in the equation below:



MetC proteins from *Bordetella avium*, *Arabidopsis thaliana*, *Escherichia coli*, *Saccharomyces cerevisiae*, *Salmonella typhiurium* and *Haemophilus influenzae* produced similarities to CnxG which ranged from 26.6% identity (*B. avium*) to 24.0% identity (*H. influenzae*). Occupying positions 18-23 in Table 2, the MetC proteins were collectively, least similar to CnxG. MetZ from *Pseudomonas aeruginosa*, which catalyses the one-step conversion of O-succinylhomoserine to homocysteine during methionine biosynthesis, in a reaction equivalent to that produced by both MetB and MetC in other organisms, was found to be 28.5% identical (54.2% similar) to CnxG.

CysD or O-acetylhomoserine sulfhydrylase or homocysteine synthase from *Aspergillus nidulans* was 28.2% identical (50.1% similar) to CnxG. CysD is involved in the conversion of O-acetylhomoserine to homocysteine during methionine biosynthesis as detailed in the following equation:



The equivalent enzyme Mt17 or o-acetylhomoserine sulfhydrylase from *S. cerevisiae* exhibited 29.6% identity (51.9% similarity) to CnxG. In *S. cerevisiae* Mt17 catalyses the conversion of o-acetylserine to cysteine (see equation below) as well as o-acetylhomoserine into homocysteine.



MegL from *Pseudomonas putida* which functions as methionine gamma-lyase, was fourth most similar to CnxG showing 33.3% identity (58.5% similarity). Methionine-gamma lyase catalyses the following reaction:



Cys3 from *S. cerevisiae*, Cgl from *Caenorhabditis elegans*, rat (*Rattus norvegicus*) and human (*Homo sapiens*) and CysA from *Streptomyces coelicolor* are all cystathionine-gamma lyases. These proteins exhibited similarities to CnxG which ranged from 32.5% identity (*S. cerevisiae*) to 28.0% identity (rat). Cystathionine-gamma lyase is responsible for the cleavage and deamination of cystathionine to produce cysteine during cysteine biosynthesis. The reaction catalysed by cystathionine-gamma synthase is detailed below:

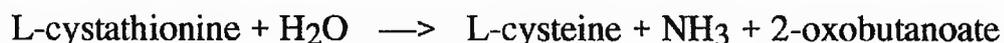


Figure 13 shows an alignment of CnxG with the top five most similar proteins, Yhr2 (*S. cerevisiae*), MetB (*E. coli*), MetB (*H. pylori*), MegL (*P. putida*) and Cys3 (*S. cerevisiae*), performed using Pileup and Prettybox functions on the Seqnet computer program. The similarity to all five enzymes stretches along the entire length of the CnxG sequence.

### **3.12 Sequence Analysis of the *cnxG4* Mutant Strain**

Genomic DNA was extracted from strain *cnxG4* and digested with the restriction enzyme *EcoRI* which was known to leave the *cnxG*

Figure 13: Alignment of the deduced amino acid sequence of *A. nidulans* CnxG with the proteins Yhr2 from *S. cerevisiae*, MetB from *E. coli*, MetB from *H. pylori*, MegL from *P. putida* and Cys3 from *S. cerevisiae*.

Gaps (.) have been inserted to produce maximum alignment. The consensus sequence for the five proteins is shown.

1 50  
 CnxG *A. nidulans* MLPEGQDVGA STRSLHAD.D VLNVVTDVAP PIHVATFFRY SDDPAKLMPA  
 Yhr2 *S. cerevisiae* .....MVDL STALIHGD.D KDNRVTDVAP PINVSTFFRY DDDDLIPWTE  
 MetB *E. coli* .....MTRKQ ATIAVRSGLN DDEQYGCVVP PIHLSSYTFN FG.....F.  
 MetB *H. pylori* .....MRM QTKLIHGGS EDATGAVSV PIYQTSTYRQ DA.....I.  
 MegL *P. putida* MHGSNKLPGF ATRAIHHGYD PQDHGGALVP FVYQTATFTF PT.....VE  
 Cys3 *S. cerevisiae* ..TLQSDKF ATKAIHAGEH VDVH.GSVIE PISLSTTFKQ SS.....  
 Consensus M-----V-F AT-AIH-G-D -D-HVG-VVP PI--STTFR--DD-----E

51 100  
 CnxG *A. nidulans* ADLNGTLEEQ SSIYISRLTA PTTTRFEAIL SLLNGQ.TI SYSSGLSAFH  
 Yhr2 *S. cerevisiae* RENLDFMEKK P..VYSRLAH PNSTRLESIF SEILDGY.AV IYSSGLAIFY  
 MetB *E. coli* .....NEP RAHDYSRRGN PTRDVVQRAL AELEGGAGAV LNTMGSAIH  
 MetB *H. pylori* .....GRH KGYEYSRSGN PTRFALEELI ADLEGGVKGAF AFASGLAGIH  
 MegL *P. putida* YGAACFAGEQ AGHFYSRISN PTLNLEARM ASLEGEAGL ALASGMCAIT  
 Cys3 *S. cerevisiae* ....PANPI GTYEYSRSON PNRENLERAV AALENAQYGL AFSSG.SATT  
 Consensus -----F--EQ -GYEYSR-GN PTR-RLS--L AELEGGQ-G- A-SSGLSAIH

101 150  
 CnxG *A. nidulans* AALTL...N PRRISIGKGY HGCHGVIDIF NRLNGLQKLD LDCPAEQL..  
 Yhr2 *S. cerevisiae* AAMVHY...N PKKIFIGQSY HGVRAIANIL TRNYGIEQHP LEDIEKCA..  
 MetB *E. coli* LVTTVFLKPG DLLVAPHDCY GGSYRLFDSL AKRGCYRVLV VDQGDQALR  
 MetB *H. pylori* AVFS.LLQSG DHVLLGDDVY GGTFRLEFNOV LVKNGLSCTI IDTSDISQK  
 MegL *P. putida* STLWTLRPG DEVLLGNTLY GCTFALHGG IGEFGVKLRH VDMADLQALE  
 Cys3 *S. cerevisiae* ATILQSLPQG SHAVSIQDVY GGTFRYFTKV ANAHGVETSF TN.DLLNDLP  
 Consensus A-IT-LL-PG DHV-LIGDVY GGTFRLE-IL AR--GVE-LF VD--D-QAL-

151 200  
 CnxG *A. nidulans* ....EAGDVI HLETPVNPEG TSFNIEQYAK KAHSRGA... .YLIVDGTFA  
 Yhr2 *S. cerevisiae* ....SEGDIV HLESPVNPYG TSSDIESLAR RAHAKGA... .LLIVDSTFA  
 MetB *E. coli* AALAEKPKLV LVESPSNPLL RVVDIAKICH LAREVGA... .VSVVDNTEFL  
 MetB *H. pylori* KAIKPNTKAL YLETPSNPLL KITDLAQCAS VAKDHGL... .LTIVDNTEFA  
 MegL *P. putida* AAMTPATRVI YFESPANPNM HMADIAGVAK IARKHGA... .TVVVDNTEYC  
 Cys3 *S. cerevisiae* QLIKENTKLV WIETPTNPTL KVTDIQKQVAD LTKKHAAGQD VLLVVDNTEFL  
 Consensus AA-KE-TKVV YLE-P-NPLL -VTDIA-VAK -A--HGA--- --L-VDNTEFA

201 250  
 CnxG *A. nidulans* PPPLQDPFKF GADLVLHSGS KYFGGHSQVLCVGLATQNKD WAQ.QLLRDR  
 Yhr2 *S. cerevisiae* SPPLQYAWNF GADIVLYSAT KYFGGHSDDL SGVLVVKKEA TSR.QLKDDR  
 MetB *E. coli* SPALQNPLAL GADLVLHSGT KYLNGHSDVV AGVVIKDPD VVT.ELAWWA  
 MetB *H. pylori* TPYIQNPLLL GADIVVHSGT KYLGGHSDVV AGLVTTNNEA LAQ.EIAFFQ  
 MegL *P. putida* TPYLQRPLEL GADLVHVSAT KYLSCHGDIT ACIVVGSQAL VDRIRLQGLK  
 Cys3 *S. cerevisiae* SPYISNPLNF GADIVVHVSAT KYINGHSDVV LGVLATNKP LYE.RLQFLQ  
 Consensus SPYLNPLN- GAD-VLHVSAT KYLGGHSDVV AGV-VT-N-- -A---L-FL-

251 300  
 CnxG *A. nidulans* VFLGGVMGNL EGWLGVRSLR TLEVRVQRLS QNATNLVLWL HNALQTPSPA  
 Yhr2 *S. cerevisiae* IYLGTVNANL ESFMLLRSLR TYEMRITKQS ENATKLVRF .....  
 MetB *E. coli* NNIGVTGGAF DSYLLLRGLR TLVPRMELAQ RNAQAIVKYL .....  
 MetB *H. pylori* NAIGGVLGPO DSWLLQRGIK TLGLRMEAHQ KNALCVAEFL .....  
 MegL *P. putida* DMTGAVLSPH DAALLMRGIK TLNLRMDRHC ANAQVLAEFL .....  
 Cys3 *S. cerevisiae* NAIGAIKPSPF DAWLTHRGLK TLHLRVRQAA LSANKIAEFL .....  
 Consensus N-IG-VLGP- DSWLLRGL- TLELR-ER-- -NA-K--EFL -----

301 350  
 CnxG *A. nidulans* PGSDEEATQK VLEQVFHSSL QKDDSWLLK QMPNGF..GP VFSITMREED  
 Yhr2 *S. cerevisiae* .SDHQSEFDK VLKTIYHSSL QTEE..PVKK QLVGGY..GP VFAITLYTKE  
 MetB *E. coli* .....QTQPL .VKKLYHPSL PENQGHETAA RQKGF..GA MLSFELDGE  
 MetB *H. pylori* .....EKHPK .VERVYYPGL PTHPNYELAK QMRGF..SG MLSFTLKND  
 MegL *P. putida* .....ARQPO .VELIHYPGL ASFPQYTLAR QMSQP..GG MIAFELKGGI  
 Cys3 *S. cerevisiae* .....AADKE NVVAVNYPGL KTHPNYDVVL QHRDALGGG MISFRIKGA  
 Consensus -----AQPK VVE-VY-P-L -T-PNYE-AK QMRGF--GG M-SFTLKGE

351 400  
 CnxG *A. nidulans* YA.RHLPSKL ALFQHATSLG GVESLIEWRT .....MSDKTVDRR  
 Yhr2 *S. cerevisiae* QC.KQLPLKL KYFHATSLG GIESLIEWRA .....MTDPYIDQT  
 MetB *E. coli* QTLRRFLGGL SLFTLAEISLG GVESLISHAA TMTHAGMAPE ARAAGISET  
 MetB *H. pylori* EAV.AFVESL KLFILGESLG GVESLVGIPA FMTHACIPKT QREAAGIRDG  
 MegL *P. putida* GAGRRFMNAL QLFSRAVSLG DAESLAQHPA SMTHSSYTP ERAHYGISEG  
 Cys3 *S. cerevisiae* EAASKFASST RLFTLAEISLG GIESLLEVPA VMTGGIYKPE AREASGVFDD  
 Consensus -A-RRF-S-L KLFTLAEISLG GVESL-E-PA -MTHAGIP-E AR-AAGI---

401 429  
 CnxG *A. nidulans* LLRVSIGLEN WEDLKRDLVN AFRALV...  
 Yhr2 *S. cerevisiae* LIRVSGCES ANDLIKDLAS ALKELQDAA  
 MetB *E. coli* LLRISTGIED GEDLIADLEN GFRAANKG.  
 MetB *H. pylori* LVRLSVGIEH EQDLLEDLEQ AFAKIG...  
 MegL *P. putida* LVRLSVGLED IDLLADVQQ ALKASA...  
 Cys3 *S. cerevisiae* LVRISVGIED TDDLLEDIKQ ALKQATN..  
 Consensus LVR-SVGIED --DLL-DLEQ A-KA-----

gene uncut. PCR was then used to amplify an appropriate fragment of genomic DNA containing *cnxG* to use as a template for sequencing. Table 3 shows the location, nucleotide sequence and amino acid sequence of the *cnxG4* mutation. The numbers specifying the position of the mutated nucleotide and amino acid residues refer to the numbering system detailed for the wild-type *cnxG* sequence (Figure 12). The *cnxG4* mutation was found to be a single base pair substitution, which resulted in a change from guanine to thymine at nucleotide position 792. This mutation would be responsible for a change in the corresponding predicted amino acid sequence from valine to phenylalanine at residue 250.

### **3.13 Northern Analysis of the *cnxG* Gene**

Northern analysis was carried out to determine the size of the *cnxG* transcript and to provide an indication of whether gene transcription was regulated by nitrate or ammonium as is the case with the nitrate reductase apoenzyme structural gene, *niaD*. Messenger RNA was extracted from wild-type *A. nidulans* cells grown with either nitrate or ammonium as sole nitrogen source. Messenger RNA from both growth conditions was used in a Northern blot and probed with an approximately 500 bp *Bam*HI-*Sph*I fragment taken from within the *cnxG* coding region of pLJM18.

The probe hybridised to two RNA messages, a strongly hybridising 1.6 kb message and a very faintly hybridising 7.5 kb message in both lanes of the Northern blot (Figure 14). These RNA transcripts were the same sizes as the two bands seen in earlier Northern blots using pK and the 2.2 kb *Bam*HI pK fragment as probes (refer to section 3.4). The intensity of hybridisation appeared to be evenly matched between each of the mRNA samples. A 1.6 kb transcript would agree with a *cnxG* gene coding region of 1.2 kb plus an approximate 0.4 kb untranslated

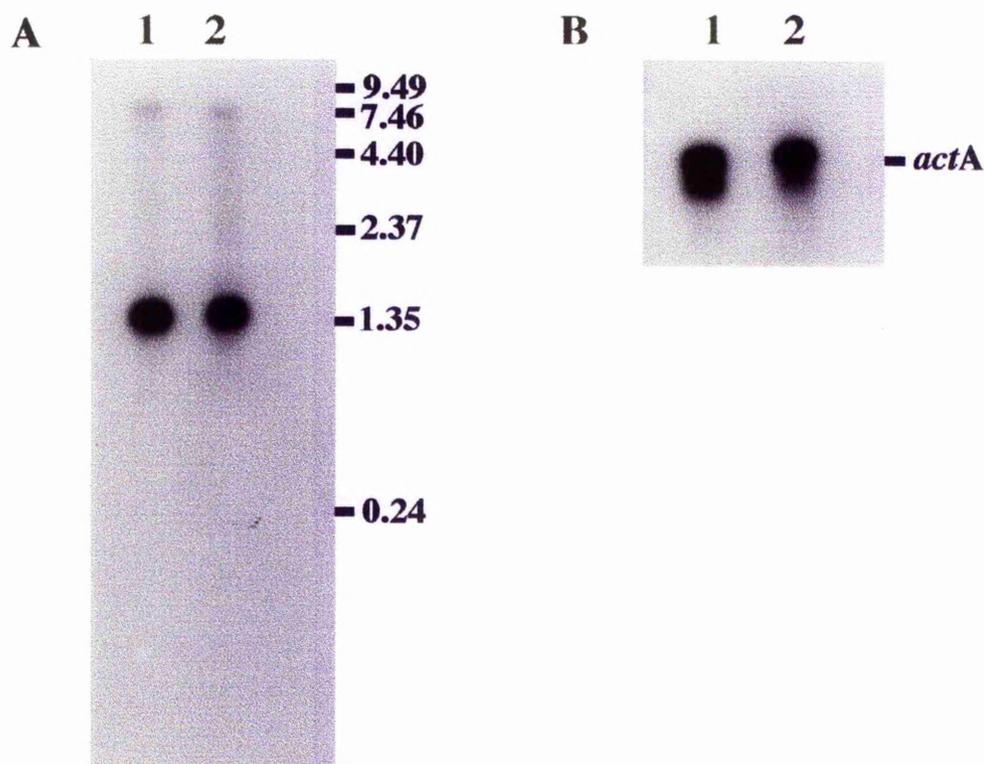
Table 3: Location, nucleotide sequence and amino acid sequence of the *cnxG4* mutation

Strain	Nucleotide Sequence <sup>ab</sup>		Amino Acid Sequence <sup>ab</sup>					
Wild-type	786	CTC GGC GTC CGG AGT	248	L	G	V	R	S
<i>cnxG4</i>		CTC GGC <u>T</u> TC CGG AGT		L	G	<u>F</u>	R	S

<sup>a</sup> Underlined letters indicate residues which are altered by the *cnxG4* mutation.

<sup>b</sup> The numbers to the left of the DNA and amino acid sequences correspond to the numbering system detailed for the *A. nidulans* wild-type *cnxG* nucleotide and amino acid sequences (refer to Figure 12).

Figure 14: Northern analysis of the *cnxG* gene



Messenger RNA was extracted from wild-type *A. nidulans* cells grown in minimal medium with either 10 mM nitrate or 10 mM ammonium as sole nitrogen source. Messenger RNA was then used to produce a Northern blot and hybridised with the 500 bp *Bam*HI-*Sph*I fragment taken from within the *cnxG* gene coding region. Panel A shows the appearance of the autorad after incubation with the hybridised Northern blot for 3 days at  $-70^{\circ}\text{C}$ . The molecular sizes (kb) of the RNA markers are indicated. The blot was stripped and rehybridised with a fragment specific for the constitutively expressed *A. nidulans actA* gene (Panel B). In both Panels, lane 1 = nitrate-grown cells and lane 2 = ammonium-grown cells.

leader sequence. The faint 7.5 kb message was unexpected. This large message may be an artefact. Alternatively, the probe may have partial homology with another transcript.

Rehybridisation with a fragment of the *A. nidulans* actin gene revealed that the distribution of mRNA between the 2 lanes of the Northern blot was reasonably equal. The expression of *cnxG* therefore does not appear to be controlled by nitrate or ammonium regulation.

### **3.14 Further Investigation of the Second *cnxG* Transcript**

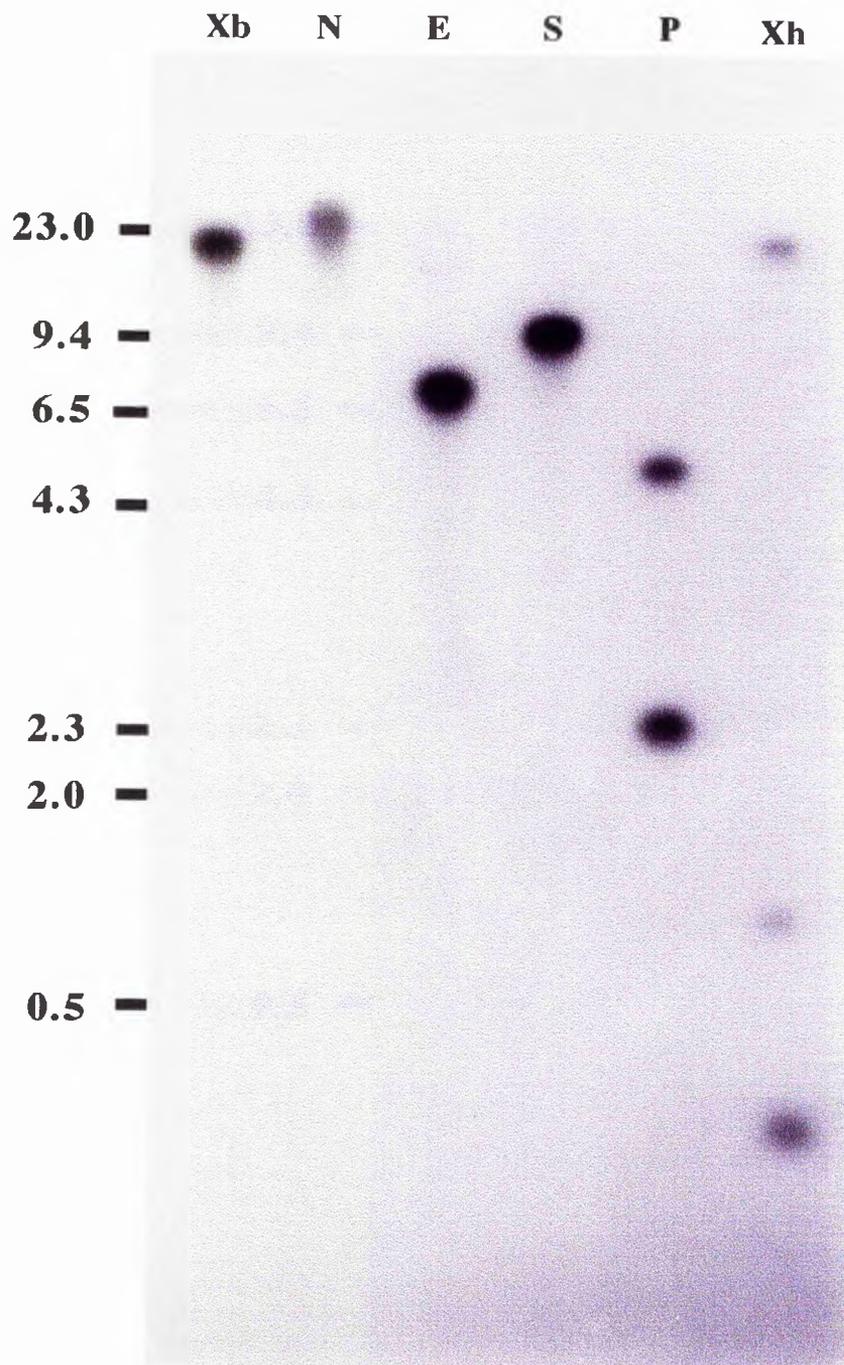
An experiment was set up to investigate whether a second gene with homology to the *cnxG* coding region was present in *A. nidulans*. This experiment involved Southern blotting and hybridisation of digested wild-type DNA with a probe taken from within the *cnxG* coding region. It was hoped that this would show if any extra hybridising bands were present in the genomic DNA which might correspond to a gene encoding the large 7.5 kb transcript.

6 µg wild-type genomic DNA was digested with several different restriction enzymes known either to: cut the probe once (*PvuII*); cut the probe twice (*XhoI*) or fail to cut the probe (*XbaI*, *NotI*, *EcoRI* and *SalI*). The digests were used in a Southern blot and probed with the same 500 bp *BamHI-SphI* coding region fragment of pLJM18 which was used in earlier Northern analysis (see section 3.12).

As can be seen from Figure 15, the probe did not hybridise to any extra DNA bands in the various genomic digests. The number of bands hybridising to the probe was in each case as predicted by the *cnxG* restriction sites. If the large, faint RNA message observed in Figure 14 was a second area of homology to the *cnxG* probe we would have expected the same homologous area to be highlighted in the genomic DNA, therefore it appears that this explanation can be ruled out. The

Figure 15: Southern blot analysis of the second *cnxG* transcript

A Southern blot containing wild-type genomic DNA digested with various restriction enzymes was hybridised with the same 500 bp *Bam*HI-*Sph*I probe (taken from within the *cnxG* coding region) which was used in previous Northern analysis. The chosen restriction enzymes were known to either: cut the probe once; cut the probe twice or fail to cut the probe. The following abbreviations are used: Xb = *Xba*I; N = *Not*I; E = *Eco*RI; S = *Sal*I; P = *Pvu*II; Xh = *Xho*I. Autoradiography was carried out for 4 days at -70 °C. The molecular sizes (kb) of the lambda (digested with *Hind*III) markers are indicated.



faint 7.5 kb band highlighted by the *cnxG* coding region probe in Northern analysis is most probably an artefact.

### **3.15 Over-expression of the CnxG Protein**

#### **3.15.1 Construction of a CnxG Protein Over-Expression Plasmid**

A CnxG protein over-expression plasmid was constructed to enable future enzyme assays to be carried out and to allow future protein purification. The strategy employed was to amplify (by PCR) a DNA fragment containing the *cnxG* coding region with engineered restriction sites to allow insertion of the gene sequence into the correct reading frame and orientation within the protein expression vector pET21a. pET21a contains a sequence adjacent to the polycloning site which codes for a peptide tag consisting of six Histidine residues (His.Tag). This His.Tag becomes incorporated into the over-expressed protein as a short C-terminal extension and allows the protein to be purified by affinity chromatography.

Primers were designed so as to anneal to the DNA sequences encoding the first 6 amino acids at the amino-terminal end (starting with the ATG methionine translational start codon) and the last 7 amino acids (discounting the translational stop codon) at the carboxy-terminal end of CnxG. An *EcoRI* restriction site was incorporated into the 5' or amino-terminal primer and a *SalI* restriction site was incorporated into the 3' or carboxy-terminal primer. These restriction sites were known to exist in the correct orientation (*i.e.* 5' to 3') in the pET21a polylinker and were absent within the *cnxG* coding region sequence. Sequences which would allow efficient cleavage of the resulting fragment during restriction (designated by double asterisks [\*\*] below) were also included in each primer. The primer sequences are detailed overleaf.

*cnxG* 5' Primer (*EcoRI* site)

5'- GCCG GAATTC ATG CTG CCA GAA GGC CAG -3'

\*\* *EcoRI* site *cnxG* coding region -first 18 bp

*cnxG* 3' Primer (*SalI* site)

5'- TGCG GTCGAC AAC CAG TGC TCT AAA AGC ATT -3'

\*\* *SalI* site *cnxG* coding region -last 21 bp

PCR was performed using cDNA clone 3 as template so as to produce the *cnxG* coding region sequence without the intron which the *E. coli* protein expression strain would not be able to remove. The fragment resulting from PCR (approximately 1.2 kb) was digested with *EcoRI* and *SalI* and subsequently ligated into *EcoRI/SalI*-digested pET21a. The resulting CnxG protein over-expression plasmid was designated pLIND-G.

The insert region of pLIND-G was completely sequenced in one strand to check that the *cnxG* gene sequence did not contain any changes and had been incorporated into pET21a in the correct orientation and position. The sequence obtained from the protein expression plasmid appeared to be identical to that of cDNA clone 3. The *cnxG* coding region sequence had been inserted within the correct reading frame, in the correct position within the vector.

### **3.15.2 Time Course and Yield of CnxG Protein Over-Expression**

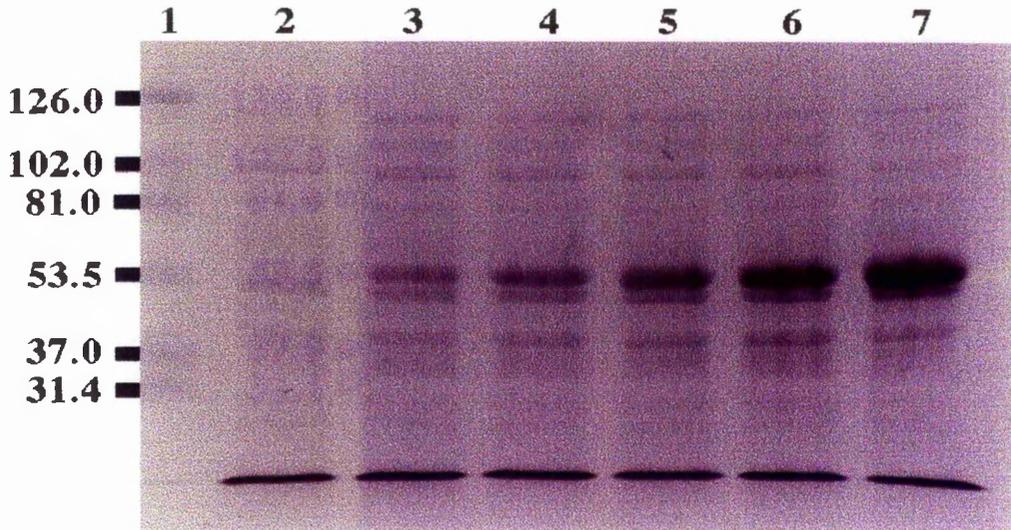
The vector pLIND-G was transformed into the *E. coli* protein over-expression strain BL21-DE3. An aliquot of transformed cells was taken from a growing culture before the induction of protein expression and further samples were taken after IPTG-induction at hourly intervals for 4 h, and following 18 h of growth. The amount of total cell protein

present in each sample was estimated by SDS-PAGE (Figure 16). A concentrated protein band with an apparent molecular weight of approximately 53.5 kDa (which was absent in the uninduced sample) was visible 1 h after induction with IPTG. This band became progressively more intense over the 4 h sample period following induction, with the greatest protein production occurring after 18 h incubation.

### **3.15.3 Determination of CnxG Protein Solubility**

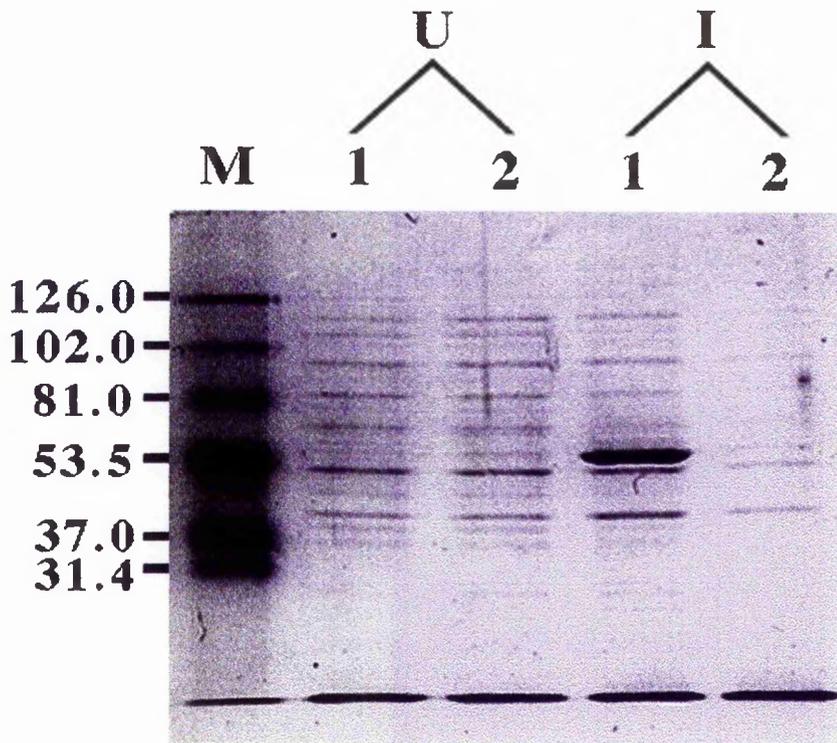
It was important to determine the solubility of the CnxG protein for future purification. To assess the solubility of the over-expressed protein it was necessary to compare the amount of protein present in the soluble and insoluble cell fractions of pLIND-G transformed BL21-DE3. Aliquots of uninduced and induced cells were removed from a growing culture of transformed BL21-DE3. The induced sample was removed after 18 h growth following IPTG induction, because this time-point appeared to produce the greatest yield of over-expressed protein during the time-course experiment (detailed in section 3.15.2). Lysozyme/sonication treatment was used to disrupt the cells and, following centrifugation, equal volumes of the supernatant (the soluble cell fraction) and the pelleted cell material (the insoluble protein fraction) were removed and examined for protein content by SDS-PAGE. Most of the induced protein was present in the supernatant sample and therefore CnxG would appear to be soluble (Figure 17). This means that the over-expressed CnxG protein will not require solubilisation prior to purification by affinity chromatography.

Figure 16: Time course and yield of CnxG protein over-expression



The vector pLIND-G was transformed into *E. coli* strain BL21 (DE3) and samples of a growing culture removed before induction, and at several time-points after IPTG-induction. The figure shows the appearance of the stained gel following SDS-PAGE of the total cell protein samples. The samples present in each track are as follows: lane 1, prestained SDS-PAGE protein markers; lane 2, uninduced cells; lane 3, cells 1 h after induction; lane 4, cells 2 h after induction; lane 5, cells 3 h after induction; lane 6, cells 4 h after induction; lane 7, cells 18 h after induction. The apparent molecular weights (kDa) of the protein markers are indicated.

Figure 17: Determination of CnxG solubility



Samples of cells taken from a growing culture of pLIND-G-transformed *E. coli* strain BL21 (DE3), before induction and 18 h after IPTG-induction, were subjected to lysozyme/sonication treatment. Equal volumes of the resulting soluble and insoluble protein fractions were removed from each sample and analysed by SDS-PAGE. The appearance of the stained gel is shown. The samples present in each track are as follows: lane M, prestained SDS-PAGE protein markers; lane U, uninduced cells; lane I, cells 18 h after IPTG induction. Within each of the uninduced (U) and induced (I) samples, 1 = soluble cell fraction and 2 = insoluble cell fraction. The apparent molecular weights (kDa) of the protein markers are indicated.

## CHAPTER 4

### ISOLATION AND MOLECULAR CHARACTERISATION OF THE *cnxE* GENE

#### 4.1      Brief Introduction

This chapter deals with the isolation of *cnxE* from the *A. nidulans* genome and the characterisation of this gene at the molecular level.

Mutations in the gene *cnxE* are phenotypically repairable on nitrate as nitrogen source by the addition of 33 mM sodium molybdate (Arst *et al.*, 1970). In this respect *cnxE* is unique among the *cnx* genes identified in *A. nidulans*. In the bacterium *E. coli*, two molybdate repairable loci have been identified, *mod* and *mog*. Strains which have mutations in either *mod* or *mog* can be repaired by the addition of 1 mM sodium molybdate to the growth medium (Miller and Amy, 1983). Both the *mod* operon (Maupin-Furlow *et al.*, 1995; Walkenhorst *et al.*, 1995) and the *mogA* gene (Reiss *et al.*, 1987; Yura *et al.*, 1992; James *et al.*, 1993) have been isolated and sequenced. The *mod* operon encodes a high-affinity molybdate uptake and transport system (Johann and Hinton, 1987; Miller *et al.*, 1987). The function of the *mogA* gene product remains unknown, however it has been suggested that MogA may be involved in the insertion of molybdenum into molybdopterin during MoCo biosynthesis (Miller and Amy, 1983).

It has been proposed that *cnxE* in *A. nidulans* may function in a similar manner to the *mogA* locus because *cnxE* mutants are still able to produce the 7.6 S form of mutant nitrate reductase (detected by cytochrome *c* reductase activity) implying that these mutants produce molybdenum-free cofactor with a structure sufficiently akin to the mature

cofactor to allow aggregation of the NR monomers (MacDonald *et al.*, 1974 and see Chapter 1, Introduction).

#### **4.2 Initial Cloning Strategy: Transformation of *A. nidulans* *cnxE* Mutant Strains with Wild-type Genomic *A. nidulans* DNA.**

The initial strategy employed to isolate the *cnxE* gene was to use genomic wild-type *A. nidulans* DNA to phenotypically complement *A. nidulans* *cnxE* mutant strains. The *cnxE* mutant strains *cnxE14* and *cnxE10* were tested in transformation experiments using either *Sau3AI* partially digested wild-type genomic DNA or an *A. nidulans* *argB* genomic DNA bank (undigested). In each experiment the genomic DNA was cotransformed with the autonomously replicating vector pHELP1 (undigested). Transformant colonies were selected by their ability to utilise nitrate as sole nitrogen source due to nitrate reductase activity. Approximately 1-4  $\mu\text{g}$  genomic DNA (*Sau3AI* partial digest or *argB* bank) and 1-4  $\mu\text{g}$  pHELP1 DNA were used in the initial transformation experiments. In parallel with each experiment, a negative control was provided by performing a transformation without any DNA addition or with the addition of pHELP1 alone.

Colonies appearing on the selection plates were transferred to fresh minimal medium plates, containing nitrate as sole nitrogen source to maintain the selective conditions. Total DNA (genomic and plasmid) was extracted from each transformant and 0.5-3.0  $\mu\text{g}$  used to transform competent *E. coli* DH5 $\alpha$  cells in order to 'rescue' any plasmids present within the sample. Plasmid DNA was subsequently prepared from any transformed (ampicillin resistant) *E. coli* colonies. In one case, where 73 *E. coli* transformants resulted, 28 colonies were picked at random as a representative number to prepare DNA. In all other cases, every *E. coli* transformant colony was used to prepare DNA. The 'rescued' plasmids

were used to retransform the *A. nidulans* *cnxE*14 mutant strain, either individually or in pools. Typically 0.5-2.0 µg 'rescued' plasmid DNA and 1-4 µg pHELP1 were retransformed in each experiment. Table 4 details the results of the experiments carried out to isolate *cnxE*, by transformation of *A. nidulans* *cnxE* mutant strains.

Although many *E. coli* ampicillin-resistant colonies resulted from transformation with total DNA extracted from several of the initial *A. nidulans* transformant colonies, in no case was a 'rescued' plasmid able to retransform the *cnxE* mutant strain.

#### **4.3**      **Alternative Cloning Strategy: Complementation of an *E. coli* *mogA* Mutant Strain with an Excised *A. nidulans* cDNA Library**

The failure of attempts to isolate the *cnxE* gene by functional complementation of two different *A. nidulans* *cnxE* mutant strains, led to the adoption of an alternative strategy. It has been suggested that *cnxE* may be involved in the incorporation of molybdenum into molybdopterin during MoCo biosynthesis (Arst *et al.*, 1970), in a step analogous to that carried out by the *E. coli* *mogA* locus. It was therefore postulated that *cnxE* might be isolated by functional complementation of an *E. coli* *mogA* mutant strain.

##### **4.3.1**      **Transformation of *E. coli* Strain JL3705 Competent Cells**

The *E. coli* *mogA* deficient strain JL3705 (J.H.L. Cheung, University of Dundee, unpublished) was transformed with 0.5 µg of an excised *A. nidulans* cDNA library (R. Williams, University of St Andrews, unpublished). The excised *A. nidulans* cDNA library contains clones which are under a bacterial promoter and contain no introns and therefore can be expressed within *E. coli*. Strain JL3705 was also

Table 4: Attempts to isolate *cnxE* by transformation of *cnxE* mutant strains with *A. nidulans* genomic DNA

Experiment Number	DNA Material Used <sup>a</sup>	<i>A. nidulans</i> Recipient Strain Used	Number of Initial <i>A. nidulans</i> Transformant Colonies <sup>b</sup>	Transformant Colony Designation <sup>c</sup>	Number of <i>E. coli</i> Transformant Colonies <sup>bc</sup>	Number of <i>cnxE14</i> Retransformant Colonies <sup>bc</sup>
1	<i>argB</i> bank + pHELP1	<i>cnxE14</i>	1	1.1	-	N.A.
2	Genomic DNA ( <i>Sau3AI</i> ) + pHELP1	<i>cnxE14</i>	-	N.A.	N.A.	N.A.
3	<i>argB</i> bank + pHELP1	<i>cnxE10</i>	2	2.1	1	-
				2.2	61	-
4	Genomic DNA ( <i>Sau3AI</i> ) + pHELP1	<i>cnxE10</i>	-	N.A.	N.A.	N.A.
5	<i>argB</i> bank + pHELP1	<i>cnxE14</i>	2	5.1	73	-
				5.2	33	-

<sup>a</sup>The structure of pHELP1 is detailed in the Appendix. The *argB* bank is described in section 3.2.

<sup>b</sup>A straight line (-) indicates that no transformant/retransformant colonies were obtained.

<sup>c</sup>N.A. refers to not applicable.

transformed with 0.5 µg pTPR3 (which contains the *E. coli mogA* gene (Reiss *et al.*, 1987) cloned into the vector pRK415 carrying the tetracycline resistance gene marker (T. Palmer, unpublished) to act as a positive control. A negative control was provided by transforming *E. coli* strain JL3705 competent cells without any DNA addition.

Transformed cells were spread onto LN (ampicillin or tetracycline) selection plates and incubated for 3 days under anaerobic conditions at room temperature. This provided the selection of only those cells which were ampicillin/tetracycline resistant and which had the ability to use nitrate as terminal electron acceptor *via* functional nitrate reductase *i.e.* cells in which the *mogA* mutation had been successfully complemented by an appropriate cDNA clone. Since nitrate assimilation in *E. coli* is not efficient (Venables and Guest, 1968) it was necessary to create selective conditions in which nitrate reductase fulfils only a respiratory role. This was achieved by anaerobic growth on LN medium, which contains lactate as a carbon and energy source, ammonium and amino acids as sources of nitrogen and nitrate for use as terminal electron acceptor (Venables and Guest, 1968). 1 µl of the JL3705 cell sample transformed with the cDNA library was spread onto a LN/ampicillin plate and incubated aerobically overnight at 37 °C to provide an indication of the total number of cDNA clones screened (see Chapter 2, Materials and Methods).

After 3 days anaerobic incubation no colonies were observed on the test plates containing cells transformed with the cDNA library. The negative control plates also showed no growth, while 567 transformant colonies were counted between all 3 of the positive control plates. The aerobic test plate contained 2 ampicillin-resistant *E. coli* colonies suggesting that approximately 2000 colonies = clones had been screened per 100 µl JL3705 competent cells transformed with 0.5 µg excised

cDNA library. Using the equation detailed below it was deduced that for the *A. nidulans* cDNA library, with an average insert size of 2 kb, at least 30958 clones (colonies) would need to be screened in order to have a 99% probability of including any one desired clone. It was therefore decided to try to increase the transformation frequency in order to screen an increased number of cDNA clones.

$$N = \log (1 - P) / \ln (1 - f)$$

N = Number of recombinant clones (colonies or plaques/pfu.'s)

P = Desired probability, for 99% probability use 0.99

f = Average insert size (in bp) ÷ Genome size (31 x 10<sup>6</sup> bp for *A. nidulans*)

#### **4.3.2 Electro-transformation of *E. coli* Strain JL3705**

The technique of electro-transformation was used in order to gain an increase in transformation frequency. This method involves passing electrical pulses through *E. coli* cells to stimulate DNA uptake. Yields of up to 10<sup>10</sup> transformants/μg DNA have been obtained with *E. coli* strain LE392 or MC1061 and the plasmid pUC18 (Dower *et al.*, 1988). The efficiency of electro-transformation is in general 10 to 20 times higher than that obtained with maximally-competent cells prepared by chemical methods (Dower *et al.*, 1988).

*E. coli* strain JL3705 was electro-transformed with 0.5 μg of the excised *A. nidulans* cDNA library, which had been dialysed over distilled H<sub>2</sub>O for 1 h. Strain JL3705 was also electro-transformed with 0.5 μg dialysed pTPR3 to act as a positive control. A negative control was provided by electro-transformation of strain JL3705 without any DNA addition. The DNA was dialysed in order to increase it's purity which it

was thought might contribute towards enhanced transformation efficiency. Transformed cells were plated on LN selective medium as described previously, with an aerobic control again performed.

From the aerobic sample it was estimated that approximately 29000 clones had been screened per electro-transformation of 0.5  $\mu$ g library DNA. This represented an approximate 14-fold increase in transformation frequency when compared to the conventional transformation procedure and was very close to the target number of 30958 clones which needed to be screened to in order to have a 99% probability of finding a single gene copy. The 'no DNA' negative control plates were completely blank while the positive control plates were completely covered in colonies (uncountable number).

On the selection plates containing JL3705 cells transformed with the excised cDNA library only one *E. coli* colony was found able to grow after 3 days in an anaerobic environment. This single colony was removed and restreaked onto an LB/Ampicillin plate which was incubated aerobically overnight at 37 °C.

A 7.5 kb plasmid (designated pE1) was subsequently isolated from a single colony of this culture. When pE1 DNA was digested with *Kpn*I and *Sac*I (which cut at either end of the Bluescript vector polylinker), an insert of approximately 4.5 kb was identified in addition to the 3.0 kb Bluescript II KS (+/-) phagemid vector sequence.

#### **4.4      Retransformation of the *E. coli mogA* Mutant (Strain JL3705) with pE1**

To check the complementation of the *E. coli mogA* mutant by pE1, 0.5  $\mu$ g of this DNA was used to retransform strain JL3705. Strain JL3705 competent cells were transformed in the conventional manner as opposed to electro-transformation because a very high transformation

frequency was not necessary to check complementation with pE1. 0.5  $\mu\text{g}$  pTPR3 was transformed into JL3705 cells to act as a positive control and a cell sample transformed without DNA provided a negative control. Transformed cells from each treatment were split into 3 portions and spread onto 3 LN (ampicillin or tetracycline) selection plates.

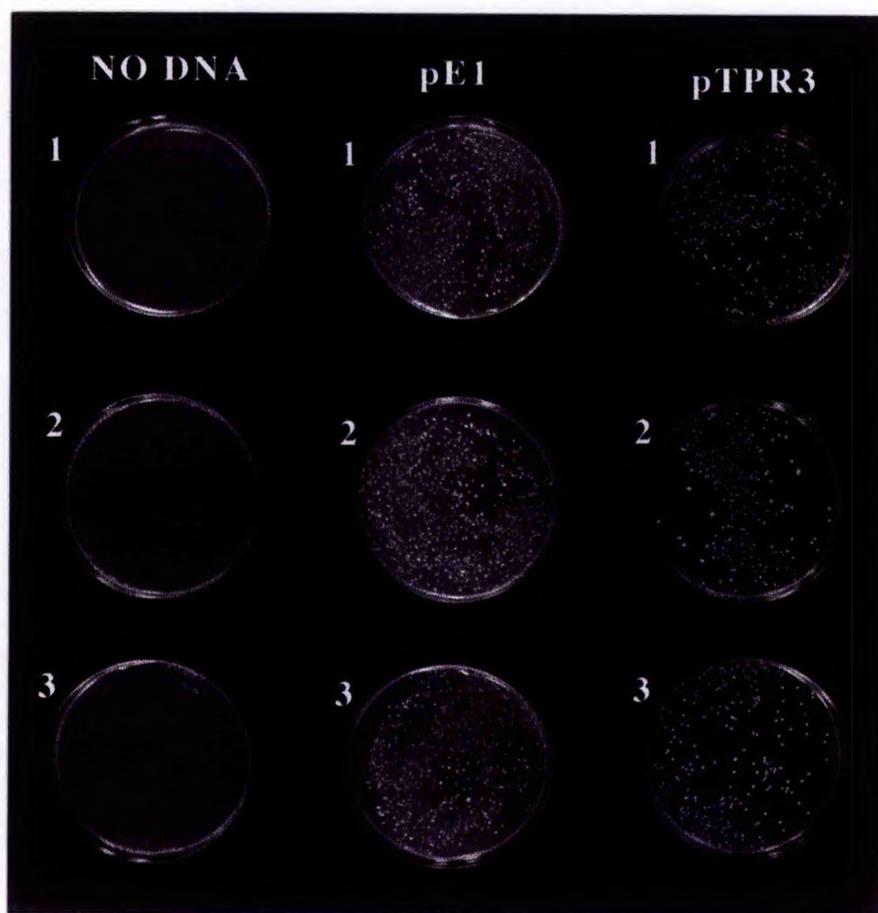
Figure 18 shows the appearance of the plates after incubation in an anaerobic environment for 3 days at room temperature. Approximately 2700 transformant colonies were produced per 0.5  $\mu\text{g}$  pE1. The positive control plasmid pTPR3 produced 492 transformant colonies per 0.5  $\mu\text{g}$  DNA. In the negative control no colonies were observed on the selection plates.

#### 4.5 Nitrate Reductase Activity in the pE1-Transformed *E. coli mogA* Mutant Strain

The *E. coli mogA* mutant strain JL3705 transformed with pE1 was examined for nitrate reductase activity using the nitrite overlay procedure (Glaser and DeMoss, 1971). The *mogA* mutant transformed with pTPR3 provided the positive control. The untransformed *mogA* mutant strain was used as the negative control. A single colony of each *E. coli* transformant to be tested was restreaked over the surface of a fresh LB plate. The resulting plates were then subjected to the nitrite overlay procedure.

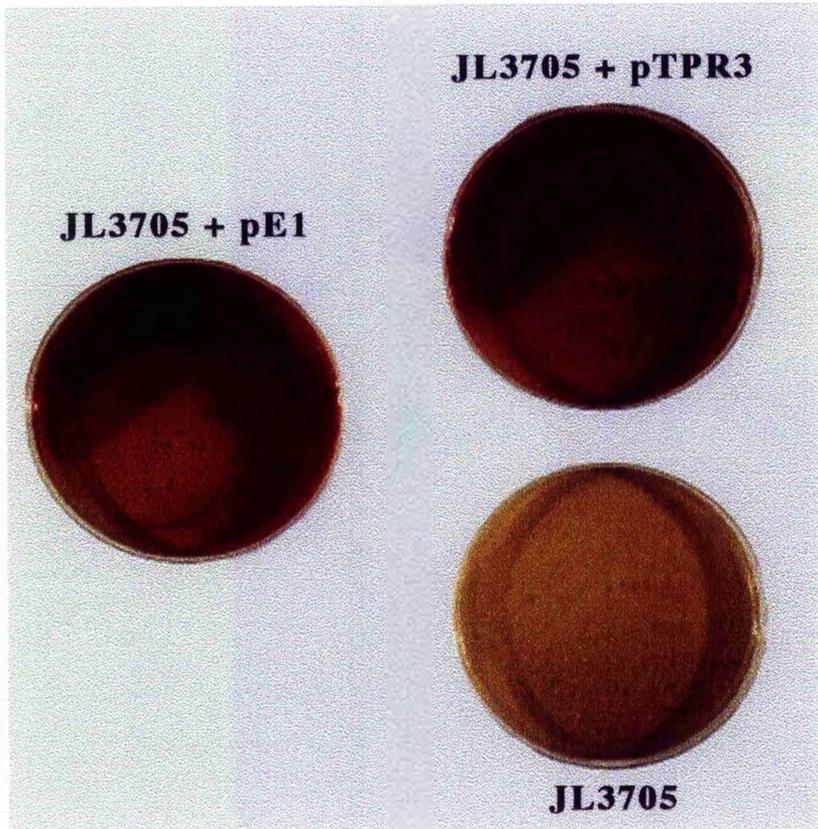
A pink colouration was observed in the agar surrounding the pE1- and pTPR3-transformed *E. coli* mutant growth indicating functional nitrate reductase activity (Figure 19). No nitrate reductase activity was observed in the untransformed mutant strain with the agar around the bacterial growth remaining colourless.

Figure 18: Complementation of *E. coli* strain JL3705 with the recombinant vector pE1



0.5  $\mu\text{g}$  of the recombinant vector pE1 was used to transform competent cells of strain JL3705. Transformation of strain JL3705 with 0.5  $\mu\text{g}$  pTPR3, and without any DNA addition, provided the positive and negative controls respectively. In each sample the transformed competent cells were divided into three equal portions and spread on triplicate LN/antibiotic plates (numbered 1, 2 and 3 above). Incubation of plates was carried out for 3 days, at room temperature, under anaerobic conditions.

Figure 19: Detection of nitrate reductase activity in pE1-transformed *E. coli* strain JL3705



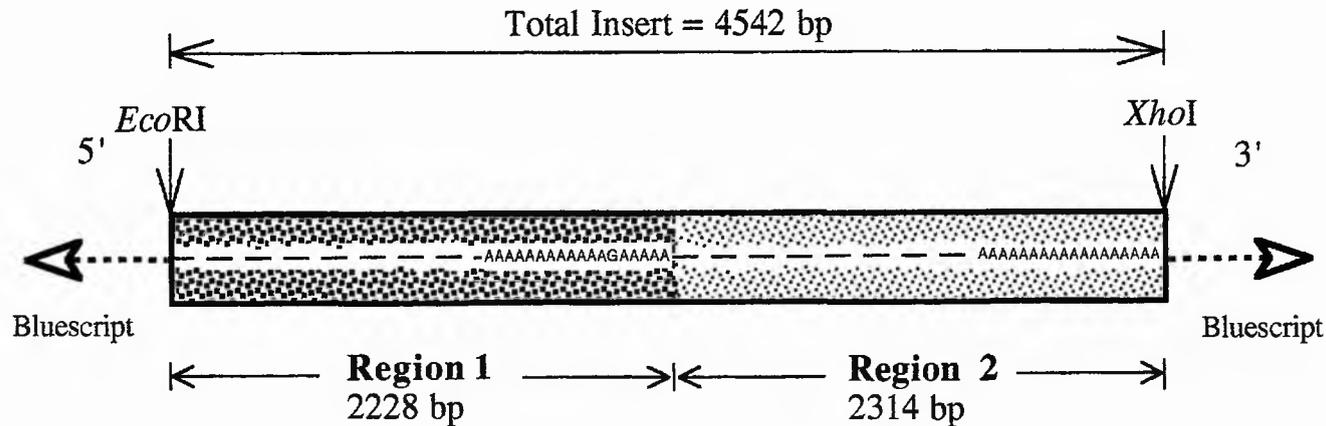
A single colony of pE1-transformed strain JL3705 was restreaked onto an LB plate and incubated overnight under anaerobic conditions at room temperature, before testing for the production of nitrate reductase using the nitrite overlay procedure (Glaser and DeMoss, 1971). A single colony of pTPR3-transformed strain JL3705 was treated in the same manner to provide a positive control. An isolated colony of untransformed strain JL3705 was also treated as described above, to act as a negative control. A colour change in the medium surrounding the *E. coli* growth, from colourless to pink, was indicative of nitrate reductase activity.

#### 4.6      Sequence Analysis of pE1

The cDNA clone pE1 was transformed into *E. coli* strain DH5 $\alpha$ , and DNA was prepared from this strain in order to achieve a high quality template for sequencing. The recombinant vector pE1 was completely sequenced in 1 strand, starting with the universal reverse and forward (-40) primers and moving along the sense and antisense DNA strands in both directions until the sequences overlapped. Figure 20 shows a schematic map of pE1. Two nucleotide regions, termed region 1 and region 2, were identified within the 4542 bp pE1 insert sequence. Region 1 consisted of 2228 bp within which an ORF was identified. This ORF continued right up to the 5' end of the inserted *A. nidulans* cDNA sequence therefore the translational start codon could not be predicted. The first of two poly(A) tails was located at the end of region 1, 92 bp downstream of the putative translational stop codon of the ORF. The region 1 poly(A) tail consisted of 17 adenine residues, interrupted by one guanine residue. A second DNA region was encountered after the first poly(A) tail which continued in the 3' direction. Region 2 consisted of 2314 bp which ended in a second poly(A) tail, composed of 18 adenine residues. No ORF was identified within region 2.

The presence of two DNA regions each with their own poly(A) tail was unexpected and suggested that perhaps 2 'joined' cDNA clones were present in pE1. It was considered necessary to isolate a genomic clone, equivalent to either or both of the DNA regions, which would determine whether the regions were contiguous within wild-type DNA or whether they represented two separate sections of DNA from different areas of the *A. nidulans* genome.

Figure 20: Schematic map of pE1



The *A. nidulans* cDNA clone insert consisting of 4542 bp was held within the *EcoRI/XhoI* restriction sites of the Bluescript II KS +/- phagemid vector polylinker. The thick dashed line (.....) represents the Bluescript vector and the arrows indicate the directions in which the vector continues. The thin dashed line (— —) within the shaded areas represents the cDNA insert sequence. The shaded areas indicate the 2 regions of the insert DNA. The position and sequence of the 2 poly(A) tails are shown at the end of each region.

#### 4.7 Isolation of Genomic Cosmid Clones Equivalent to Region 1 /Region 2 of pE1

The gene *cnxE* has been mapped to chromosome II by previous classical genetic studies therefore only clones contained within the chromosome II portion of the *A. nidulans* cosmid library were screened in order to isolate a genomic counterpart of both or either of the DNA regions identified in pE1. A fragment of DNA from within each region was amplified by PCR and used as probes in separate hybridisations against clones from the chromosome II section of the library (224 pWE15 and 186 LORIST2 cosmid clones). A 1.5 kb probe from within the region 1 ORF and a 1.5 kb probe from within region 2 were used in 2 separate hybridisation reactions. In the first of these reactions, the region 1 probe hybridised to 7 cosmid clones (see below) after 72 h incubation of the film with the filter at -70 °C.

<u>Clone Designation</u>	<u>Cosmid Number</u>
1	W15C02
2	W16B10
3	W19H04
4	L5G09
5	L8F01
6	L14A09
7	L23A09

L = LORIST2 clone

W = pWE15 clone

Cosmid clones 1, 2, 3, 4, 6, and 7 produced similar restriction patterns when digested with the same restriction enzyme (*Bam*HI)

suggesting a certain amount of overlap between the cosmids (Figure 21). Clone 5 produced only 2 bands when digested with *Bam*HI.

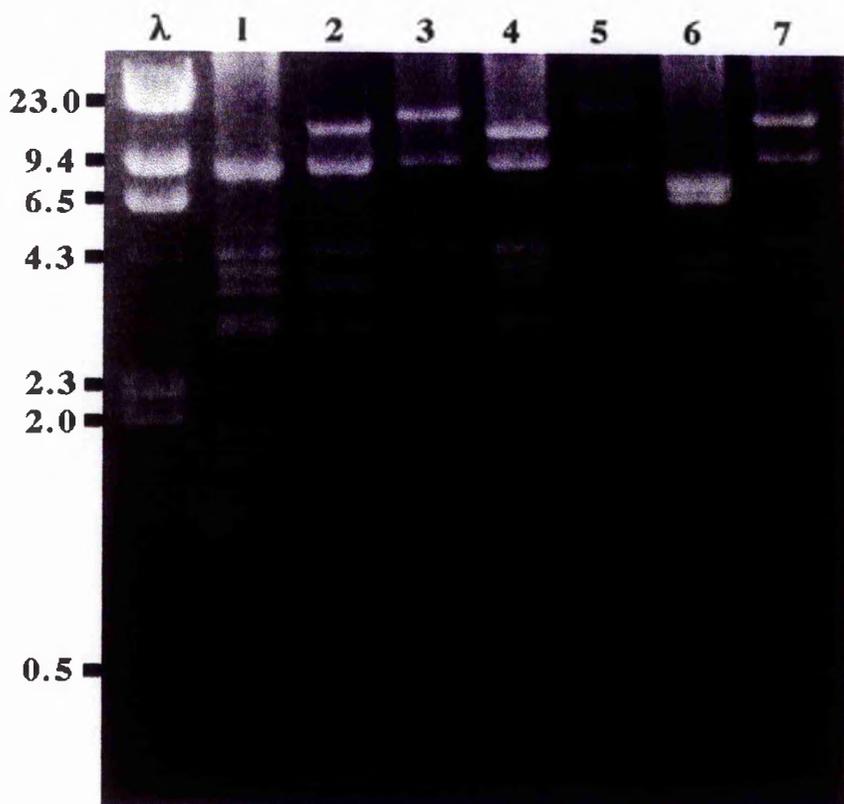
The blot was stripped of radioactive probe (section 2.15) and the second hybridisation performed. The 1.5 kb fragment from within region 2 of pE1 failed to hybridise to any of the cosmid clones screened. Sensitive X-OMAT AR fast film was incubated for 30 days at -70 °C in order to detect any weak hybridisation, however no signal was apparent after this time. To ensure that this negative result was not due to a failure in the hybridisation reaction, the experiment was repeated this time including a positive control, a separate filter containing pE1 DNA. The region 2 probe hybridised to the pE1 DNA but, as before, did not hybridise to any of the cosmid clones. Rehybridisation of the cosmid blot with the region 1 probe, again produced the original result (*i.e.* hybridisation to the same 7 cosmid clones) therefore the cosmid DNA present on the blot did not appear to have been altered by the stripping procedure between hybridisations.

The results suggest that DNA regions 1 and 2 are not contiguous within *A. nidulans* chromosomal DNA. Furthermore, a genomic sequence equivalent to region 2 does not exist within the chromosome II portion of the cosmid library suggesting that the sequence contained within region 2 is unlikely to encode *cnxE*.

#### **4.8** Further Investigation of the pE1 Region 2 Sequence

The pE1 region 2 sequence was assessed for similarities to protein sequences held within the Swiss-Prot and updated Swiss-Prot data bases using the BLASTX computer sequence comparison program. It was found that the sequence showed significant similarity to only one protein, a hypothetical 105.9 kDa protein termed Ybt6 from the yeast *Saccharomyces cerevisiae*, which is a putative integral membrane protein.

Figure 21: Restriction digest pattern of the cosmid clones which hybridised to a probe taken from within the ORF identified in region 1 of the cDNA clone pE1



Cosmid clones 1 (W15C02), 2 (W16B10), 3 (W19H04), 4 (L5G09), 5 (L8F01), 6 (L14A09) and 7 (L23A09) were individually digested with the restriction enzyme *Bam*HI and the resulting material analysed by electrophoresis in a 0.8 % agarose gel. The lane containing each cosmid digest is represented by its respective clone number. The symbol λ indicates the lane containing lambda (digested with *Hind*III) markers. The molecular sizes (kb) of the markers are shown.

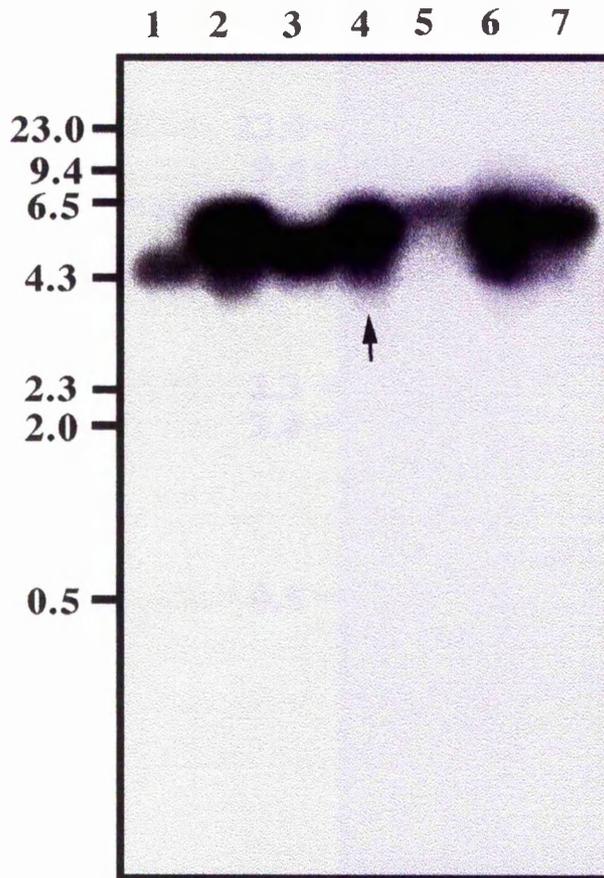
Several areas of similarity were identified, stretching from nucleotides 754 to 1863 along the length of the 2314 bp region. The similarity to Ybt6 indicates that an ORF may be present within region 2. Mistakes present within my sequence would result in reading frame shifts which could be responsible for the insertion of false stop codons, disguising the presence of any ORF. No further analysis of region 2 was undertaken.

#### **4.9 Construction of a Subclone Containing the Genomic Counterpart of the ORF Contained Within Region 1 of pE1**

It was necessary to subclone a cosmid fragment containing the entire gene corresponding to the ORF identified within region 1 of pE1 to allow sequencing. Sequencing directly from a cosmid clone would have been difficult, requiring large amounts of template cosmid DNA (approximately 10-15 µg cosmid DNA per sequencing reaction), and so it was considered necessary to obtain a smaller subclone. It was important to obtain a genomic sequence equivalent to region 1 of the cDNA clone in order to confirm that region 1 and region 2 were not contiguous within *A. nidulans* wild-type chromosomal DNA.

The nucleotide sequence obtained from pE1 suggested that the restriction enzyme *SalI* (amongst others) did not possess any sites within the region 1 ORF. Each of the 7 cosmid clones (identified in section 4.7) was digested with *SalI* and the resulting fragments subjected to Southern analysis using the same 1.5 kb region 1 probe which originally allowed isolation of the cosmids. The probe hybridised to 1 band within each of the cosmid digests (Figure 22). The hybridising fragments were of slightly different sizes suggesting that the gene was cut by the vector sequences in several (or all) of the cosmid clones. One of these highlighted *SalI* fragments, an approximately 5 kb fragment of cosmid 4

Figure 22: Identification of a cosmid fragment containing the genomic counterpart of the pE1 region 1 ORF suitable for subcloning



Cosmid clones 1-7 were each digested with the restriction enzyme *SalI* and electrophoresis carried out in a 0.8% agarose gel. The resulting gel was used to produce a Southern blot which was tested for hybridisation with a 1.5 kb PCR fragment taken from within the pE1, region 1 ORF. Autoradiography was carried out for 2 h at room temperature. The lane containing each cosmid digest is represented by its respective clone number. The molecular sizes (kb) of the lambda (digested with *HindIII*) markers are indicated. The 5 kb fragment of cosmid clone 4 which was picked for subcloning is highlighted by an arrow.

(L5G09, one of the LORIST2 clones), was chosen and ligated into the vector pUC19. The resulting genomic subclone was termed pK18.6.

#### **4.10      Structure of the *cnxE* Gene**

The nucleotide stretch of subclone pK18.6 equivalent to region 1 of pE1 was sequenced using primers deduced (in some cases) from the cDNA clone. Primers were chosen to allow sequence analysis of the opposite strand to that already obtained from pE1.

pK18.6 contained the full-length genomic counterpart of the ORF present within region 1 of pE1. Where the 5' end of the ORF sequence was terminated by the vector, the genomic clone sequence continued. The ORF continued for a further 122 bp in the genomic sequence until a stop codon was encountered. It became apparent that the ORF within region 1 of pE1 contained a slightly truncated version of the genomic gene coding region, missing 11 bp including and downstream of the putative translational start codon (which was identified in the genomic clone as the first methionine residue within the correct reading frame). 167 bp were sequenced upstream of the putative translational start codon on pK18.6. The 3' end of the genomic gene sequence deviated from that of pE1 at the position of the first base in the poly(A) tail of region 1. This confirmed that the region 1 ORF represented a distinct gene, separate from the rest of the sequence contained on pE1. Although no formal proof had yet been established it was likely that the gene previously termed the ORF within region 1 of pE1 represented *cnxE*, therefore from now on the ORF sequence (both on the cDNA clone pE1 and the genomic clone pK18.6) will be referred to as *cnxE* to ease description.

Figure 23 shows the *cnxE* genomic sequence obtained. Comparison of the genomic and cDNA clone sequences allowed the

Figure 23: Nucleotide sequence and inferred amino acid sequence of the *cnxE* gene

The genomic *cnxE* clone pK18.6 was sequenced in a single strand, opposing that already obtained from sequence analysis of the *cnxE* cDNA clone. Introns and non-coding regions present on the genomic clone were sequenced in both directions.

The numbers to the left of the sequence indicate nucleotides and the numbers to the right of the sequence indicate amino acids. Introns are presented in lower case letters. Italics have been used to denote the 5' and 3' splice sites and the sites of lariat formation within the introns (note that within the second intron there are 3 possible sites for lariat formation which have all been italicised). A possible candidate for a CAAT box is overlined. The pyrimidine-rich sequences are underlined. The translational stop codon is indicated by an asterisk. The single arrow head distinguishes the position of poly(A) tail addition and the position of the stop codon which defines the reading frame is indicated by three arrowheads. (Note that Figure 23 is contained on two sequential pages).

-167 TGGATGTTATTCGTGATGGTTGTCGAGGAAAGAACTTACAGAAATTTTGTGCCTGAGGTG  
 -107 TCAATGAGGGGTTCGGAGCGGAATAAATGCGACTCCGGACATCCCCCTCCTACTTTCTTTGGG  
 -47 ATCCTTCTCCAACCTCTTTTTACACTCATTTTCAAAGAGCGAGGGAAATGGCGGAGCAGC  
 M A E Q R 5  
 14 GATTGAAGGGCGCCATTCCTGGTGGTGTTCAGATACTGCATTCAGGATCCATCCTCGGATC  
 L K A A I L V V S D T A F Q D P S S D R 25  
 74 GAGTCTTCGGCACGCTGATCGATGTGTTCCCGGCCGACGGCCCTTGGGAGAAGCCGATCA  
 V F G T L I D V F P A D G P W E K P I T 45  
 134 CCAAGATAGTCCGGACGATGTACTTCAGATCCAGCGAGCTGTGTGCGACTGGACAGACA  
 K I V P D D V L Q I Q R A V C D W T D S 65  
 194 GTCCCAGTCCAGTGAATCTCGTTCTTATAAGTGGTGGAAACCGGCTTTGCGGTTAAGGATA  
 P S P V N L V L I S G G T G F A V K D N 85  
 254 ACACCCCCGAGGCTGTACGCCCTCTCCTTCACCGCCATGCACCTGGTCTCGTgtatgtac  
 T P E A V T P L L H R H A P G L V 102  
 314 attgatagattagttggagcgaactgaacacagGCATGGCATGATTGCCGCGTGGTT  
 H G M I A A S L 110  
 374 GAAGGTCACGCCTTgtgagttcatccgcatgagocaaatctcagcactaactaacgtatgc  
 K V T P F 115  
 434 agTTGCTATGATGGCTCGACCTGTGCTGGGACCCGACACAAGTCACTAGTTATCACTCT  
 A M M A R P V A G T R H K S L V I T L 135  
 494 GCCCGATCCCCAAAAGGTGCAAAGGAAAACCTTGAAGCTGTCATCAAACGTGTGCCTCA  
 P G S P K G A K E N L E A V I K L L P H 155  
 554 CGCATGTACCCAAGCGGCTGGTTCGGACTCGAGGACGCTCCACGCTGGAGGGATCAAGAA  
 A C T Q A A G A D S R T L H A G G I K K 175  
 614 GCTCGAAGCTGAAGCTGGGGTAAAGTCAGAGTCACAACATGATCATCATCACCATCATCA  
 L E A E A G V K S E S Q H D H H H H H H 195  
 674 TCACGAACATACTCACTCTCACTCTCATCTCATGGTTCATGGCCATGGACACGTCGTTCC  
 H E H T H S H S H S H G H G H G H V V P 215  
 734 AAGAGCACATACAACGCCTTCTGAACGGCGGTCAAATGACCCCGCAGCAGGGCAACCCG  
 R A H T T P S E R R S N D P A A G A T R 235  
 794 GCGATACCGGAGTTCGCCGTATCCCATGCTCTCCGTGGACGAAGCACTGAGGCAAGTTTC  
 R Y R E S P Y P M L S V D E A L R Q V S 255  
 854 AGCTCACACTCCGGAGCCGGAGGTTATCGAGGTTCCGGTGAATATGACCTCGTCCGATA  
 A H T P E P E V I E V P V N I D L V G Y 275  
 914 CGTTATCGCGGAAGATGTTTACGCCCGGAGGCGGTTCCAGCATACCTTGAAGTATCGT  
 V I A E D V Y A A E A V P A Y L A S I V 295  
 974 TGATGGATATGCTGTCAITGCTCCTGAATCTCCAGATGACGGACACAGTACAAAAGGGAAT  
 D G Y A V I A P E S P D D G H S T K G I 315  
 1034 CTTTCTGTGGCTTCTATTTACCCACGCCAATGAGGAGGGAGCGCTAGCACCGCTTGAGCC  
 F P V A S I T H A N E E G A L A P L E P 335  
 1094 AGGCACCATCGCCCGGATAACAACCGGCGCTCCTCTTCCGCCAATGCGAATCGGTTGGT  
 G T I A R I T T G A P L P P N A N A V V 354  
 1154 CATGGTGGAGGATACTTTACTAGGTCATCGACACCTGATGGCAAGGAAGAAGCAGTGT  
 M V E D T L L A S S T P D G K E E A T V 375

Figure 23 continued

1214 CGAGATTCTAACTGGAGAGATCAAGCCCAATGAGAACGTGCGGCAGCCGGGAAGCGATGT 395  
E I L T G E I K P N E N V R Q P G S D V  
1274 CGCTCTAGGATCCAGGATCCCTCAGCGGGGTGACCTAATCACGCTGTTCGGTGGTGAAT 415  
A L G S R I L Q R G D L I T P V G G E I  
1334 AGGCCTGCTGGCAGCTACCGGGACAAGGACCGTCAAGGTATTCAGAAGCCCGTCTGTTGGG 435  
G L L A A T G T R T V K V F K K P V V G  
1394 TGTTCCTGAGTACTGGCGACGAGCTCGTAGAGCACGACGATCCCCGGTCTCTTCAGGGAGG 455  
V L S T G D E L V E H D D P R S L Q G G  
1454 GCAGATTCGAGACTCAAATCGGCCCTCGATCCTATCTTGCCTGAAATCATGGGGGATTC 475  
Q I R D S N R P S I L S C L K S W G I P  
1514 CGCCGTTGATCTAGGTATTCGCCGCGATACGCCCTGCGGGCGAGCTGGAACAAAGCCTACG 495  
A V D L G I A R D T P A G E L E Q S L R  
1574 GGATGCTCTTCGTGGAGTCGGAAAAGTCCAACACCAGCGTCGATGTAATCATCACTACAGG 515  
D A L R G V G K S N T S V D V I I T T G  
1634 CGGTGTTTCTATGGGGAGCTCGATCTGCTAAAGCCCACCATTGAACGCTCTCTCGGGCGG 535  
G V S M G E L D L L K P T I E R S L G G  
1694 GACTATCCACTTCGGTTCGTGTCTCCATGAAACCAGGAAAGCCGACCACTTTTGGCCACCGT 555  
T I H F G R V S M K P G K P T T F A T V  
1754 TCCGTTCAAGCCAACGTCCTCTGCAGCAGGCCAACAAGAACGCAGCTCTCGACTGATCTT 575  
P F K P T S S A A G Q Q E R S S R L I F  
1814 CTCTCTCCCTGGCAACCCAGCCTCAGCACTGGTCACTTTGAATCTCTTTGTCTCCCTC 595  
S L P G N P A S A L V T L N L F V L P S  
1874 ATTACACAAGCTCATCGCCCTCGCCAGAAGCAGGCTGCATTAGGCATCGGCCCGGGCGT 615  
L H K L I G L G Q K Q A A L G I A P A L  
1934 TGGACTACCCCTTGTAGCAGTAACATTTGTCACATGCTTTCCCGCTTGACCCTAAACGCAC 635  
G L P L V A V T L S H A F P L D P K R T  
1994 GGAGTACCATCGGGCTATTTGTTACAGCGTCTCCCAAAGATGGACGGCTCTATGCCACAAG 655  
E Y H R A I V T A S P K D G R L Y A T S  
2054 TACAGGGGCAGAGGGCGTAGGACAGCGTAGTTCCCGGGTGGTAGTCTAGCCAGTGTCTAA 675  
T G A E G V G Q R S S R V G S L A S A N  
2114 TTCTCTCCTGGTGTCTCAGCCTGGGAAGGGCTCGATTGCGCAAGGGAGTCTGGTTCGAGGC 695  
S L L V L Q P G K G S I A Q G S L V E A  
2174 ACTGATGATGGGCCTATTTGTACGTGAAGTGTCTGCGGTTGCACTATGACAGGAGCGCTC 710  
L M M G P I V R E G A A V A L \*  
2234 AGCGTGGCACCAAAATAGAACACTACCTGCTCGTGAGTACGAATCCAGCCTTGAGGGGAAT  
2294 ATAGACTACTTGATTGCTGCCTAACCCCTCCAAGACTTATTCGCGTACTCCTTGTAAT  
2354 TTTTTTCCAGCAGTTACCAGCGGTCCAACCTGCGCCCCCAGGTGCTGGATAGTGTTCG  
2414 GACTCACAACTCGCGATGAATCTGACGGCATAGATATATGAGGAGT

identification of two introns: the first (intron 1) of 44 bp (starting at nucleotide position +306, where A in the translation initiation codon = +1) and the second (intron 2) of 48 bp (starting at nucleotide position +388). Both introns display good consensus sequences for 5' and 3' splice sites. Intron 1 contains a potential internal sequence for lariat formation *i.e.* TACTGAC (at nucleotide position +335), while within intron 2 there are 3 possible candidates for such a motif located at +411 (ATCTCAG), +418 (CACTAAC) and +422 (AACTAAC). The cDNA clone terminated 11 bp before the translational start codon in the *cnxE* coding region sequence, therefore although unlikely, another intron could theoretically be located within this area, or indeed within the *cnxE* 5' untranslated region. The *cnxE* coding region (minus the introns) was found to be 710 amino acids in length with a predicted protein weight of 78.1 kDa calculated using the equation indicated below.

$$\text{Number of amino acids} \times 110 = \text{Predicted protein molecular weight (in Daltons)}$$

Where 110 is the calculated average molecular weight of an amino acid residue (in Daltons).

As previously described, the translation initiation codon was identified as the first methionine within the open reading frame. The presence of a purine at nucleotide position -3 (*i.e.* GAAATG, where A in the translation initiation codon = +1) is in agreement with most (84%) fungal translational start sites (Unkles, 1992 and references therein).

Without formal identification of the transcriptional start site (*e.g.* by primer extension analysis), it is difficult to speculate on the location of promoter elements for the *cnxE* gene. Within the 167 bp 5'

upstream non-coding region which has been sequenced a putative CAAT motif is located at nucleotide -107 (TCAAT). No TATA motifs similar to the conserved sequence TATAAA can be found within the portion of the 5' non-coding region which has been sequenced. CT boxes or pyrimidine-rich sequences have been described in the promoters of yeast and filamentous fungal genes. These CT-rich motifs often occur immediately before the major transcription start point but can occur at any point in the 5' upstream region. Two possible CT boxes are present in the 5' non-coding region of *cnxE* which has been sequenced. These CT-rich sequences start at nucleotides -46 and -68 and are 19 and 18 residues in length respectively.

In the 3' non-coding region the position of the poly(A) tail was identified, 92 nucleotides downstream of the translational stop codon. A polyadenylation signal (AATAAA) could not be identified between the translational stop codon and the point of poly(A) tail addition however this is not unusual because many fungal genes lack this consensus sequence (Unkles, 1992).

#### **4.11 CnxE Protein Sequence Similarities**

The full length coding region sequence of *cnxE* was assessed for similarities to protein sequences held within the Swiss-Prot and updated Swiss-Prot data bases using the BLASTX computer sequence comparison program. *cnxE* was found to be similar to 3 eukaryotic proteins: 1) Gephyrin, a rat neuroprotein which functions to link the mammalian inhibitory glycine receptor to subsynaptic microtubules (Prior *et al.*, 1992; Kirsch and Betz, 1993; Kirsch *et al.*, 1993); 2) Cnx1, a protein from the plant *Arabidopsis thaliana* which is thought to be involved in MoCo biosynthesis although the specific function remains unknown (Stallmeyer *et al.*, 1995) and 3) Cinnamon, a protein from

*Drosophila melanogaster* which is again involved in MoCo biosynthesis (specific function unknown) (Kamdar *et al.*, 1994). The *cnxE* sequence also exhibits significant similarity to several prokaryotic proteins representing the *E. coli* MoCo biosynthesis proteins MoeA (Nohno *et al.*, 1988), MogA (Reiss *et al.*, 1987; Yura *et al.*, 1992; James *et al.*, 1993) and MoaB (Rivers *et al.*, 1993), and putative MoCo biosynthesis proteins from the bacterial species *Salmonella typhimurium* (K.K. Wong and H.S. Kwan, unpublished, sequence submitted to data bases) and *Methanococcus jannaschii* (Bult *et al.*, 1996) and the cyanobacterial species *Anabaena*, strain PCC 7120 (Ramaswamy *et al.*, 1996), *Synechococcus*, strain PCC 7942 (Rubio *et al.*, 1998) and *Synechocystis*, strain PCC 6803 (Kaneko *et al.*, 1995).

Table 5 displays the various protein similarities to the CnxE amino acid sequence in decreasing order of % identity (designated 1 to 13, where 1 represents the most similar sequence). The % identity and % similarity values were obtained by comparing each of the protein sequences with the deduced CnxE amino acid sequence using Bestfit alignments on the Seqnet computer program. It was not possible to align the full-length Cnx1 protein sequence with CnxE in this manner because Cnx1 exhibited 2 separate areas of similarity to CnxE (Figure 24). The N-terminal region of Cnx1 (from amino acid residues 2-448) was similar to the C-terminal region of CnxE (from amino acid residues 221-708), while the C-terminal of Cnx1 (from 462-670) was similar to the N-terminal of CnxE (from 1-207). Stallmeyer *et al.* (1995) have suggested that Cnx1, Gephyrin and Cinnamon contain 2 domains, termed the -E and -G domains, which are connected by a non-homologous 'linker' region. The E-domain is similar to the *E. coli* MoeA protein and the G-domain is similar to the *E. coli* MogA and MoaB protein molecules. It has been

Table 5: Protein similarities to *cnxE*

The protein sequences are listed in order of decreasing % identity. The % similarity values for each protein are also indicated.

<sup>a</sup>Number of inferred amino acid residues given for each protein.

For each of the eukaryotic proteins, Gephyrin, Cnx1 and Cinnamon, three %identity and %similarity values are given as total, E-domain and G-domain.

<sup>b</sup>The term 'total' refers to the %identity/%similarity of the protein molecule to the full-length CnxE protein sequence.

<sup>c</sup>The term 'E-domain' refers to the %identity/%similarity of the protein to amino acids 1-178 of CnxE *i.e.* the region of CnxE which is similar to *E. coli* MoeA.

<sup>d</sup>The term 'G-domain' refers to the %identity/%similarity of the protein to amino acids 237-704 of CnxE *i.e.* the region of CnxE which is similar to *E. coli* MogA and MoaB.

The full-length sequence of Cnx1 could not be directly compared to that of CnxE due to differences in domain organisation between the 2 proteins -N.A. indicates not applicable. The Cnx1 protein is ranked second in the table because on average the %identity to both the E- and G-domains of CnxE is lower than Gephyrin, but higher than Cinnamon (ranked 1 and 3 respectively).

Position	Organism	Protein	Known or Postulated Protein Function	% Identity	% Similarity	No. of Amino Acids <sup>a</sup>	Reference(s)
1	<i>Rattus norvegicus</i>	Gephyrin	Neuroprotein which anchors the inhibitory glycine receptor to subsynaptic microtubules	Total <sup>b</sup> : 41.8 E-Domain <sup>c</sup> : 45.1 G-Domain <sup>d</sup> : 39.1	Total <sup>b</sup> : 61.1 E-Domain <sup>c</sup> : 64.6 G-Domain <sup>d</sup> : 60.4	736	Prior <i>et al.</i> , (1992); Kirsch and Betz (1993); Kirsch <i>et al.</i> , (1993)
2	<i>Arabidopsis thaliana</i>	Cnx1	MoCo biosynthesis (specific function unknown)	Total <sup>b</sup> : N.A. E-Domain <sup>c</sup> : 40.4 G-Domain <sup>d</sup> : 40.2	Total <sup>b</sup> : N.A. E-Domain <sup>c</sup> : 61.7 G-Domain <sup>d</sup> : 56.3	670	Stallmeyer <i>et al.</i> , (1995)
3	<i>Drosophila melanogaster</i>	Cinnamon	MoCo biosynthesis (specific function unknown)	Total <sup>b</sup> : 34.4 E-Domain <sup>c</sup> : 37.1 G-Domain <sup>d</sup> : 33.3	Total <sup>b</sup> : 54.0 E-Domain <sup>c</sup> : 56.7 G-Domain <sup>d</sup> : 53.0	601	Kamdar <i>et al.</i> , (1994)
4	<i>Anabaena</i> sp. (strain PCC 7120)	MoeA	Putative MoCo biosynthesis protein (similarity to bacterial MoeA proteins)	33.9	51.9	463	Ramaswamy <i>et al.</i> , (1996)
5	<i>Escherichiacoli</i>	MoeA	MoCo biosynthesis (specific function unknown)	33.4	57.6	411	Nohno <i>et al.</i> , (1998)
6	<i>Escherichiacoli</i>	MogA	MoCo biosynthesis (specific function unknown)	32.8	52.6	195	Reiss <i>et al.</i> , (1987); Yura <i>et al.</i> , (1992); James <i>et al.</i> , (1993)
7	<i>Salmonella typhimurium</i>	MoeA	Putative MoCo biosynthesis protein (similarity to bacterial MoeA proteins)	32.7	56.8	412	Wong, K.K. and Kwan, H.S. (1996) -unpublished
8	<i>Synechococcus</i> sp. (strain PCC 7942)	MoeA	Putative MoCo biosynthesis protein (similarity to bacterial MoeA proteins)	30.7	56.4	403	Rubio <i>et al.</i> , (1998)
9	<i>Methanococcus jannaschii</i>	MJO886	Putative MoCo biosynthesis protein (similarity to bacterial MoeA proteins)	30.6	55.1	620	Bult <i>et al.</i> , (1996)
10	<i>Synechocystis</i> sp. (strain PCC 6803)	MoeA	Putative MoCo biosynthesis protein (similarity to bacterial MoeA proteins)	30.5	51.2	390	Kaneko <i>et al.</i> , (1995)
11	<i>Methanococcus jannaschii</i>	MJO666	Putative MoCo biosynthesis protein (similarity to bacterial MoeA proteins)	29.2	53.1	398	Bult <i>et al.</i> , (1996)
12	<i>Escherichiacoli</i>	MoaB	MoCo biosynthesis (specific function unknown)	28.8	47.5	170	Rivers <i>et al.</i> , (1993)
13	<i>Synechococcus</i> sp. (strain PCC 7942)	MoaCB	Putative MoCo biosynthesis protein (N-terminal similar to <i>E. coli</i> MoaC; C-terminal similar to <i>E. coli</i> MoaB)	22.7	45.4	319	Rubio <i>et al.</i> , (1998)

Figure 24: The two similar amino acid regions found between the *A. nidulans* CnxE and *A. thaliana* Cnx1 sequences

Identical amino acid residues are indicated by solid vertical lines. Dots imply related amino acids, with one dot indicating a greater evolutionary distance than two. Gaps (.) have been inserted to produce maximum alignment. Note that the N-terminal of CnxE is similar to the C-terminal of Cnx1, and *vice versa*.



shown that in Cnx1 the order of these 2 domains is inverted when compared to Gephyrin and Cinnamon (Stallmeyer *et al.*, 1995).

To ascertain how similar Cnx1 was to CnxE, in comparison to the other proteins mentioned in Table 5, the CnxE protein sequence was split into two domains (an E- and G-domain, following the nomenclature denoted by Stallmeyer *et al.*, (1995)) and each domain compared separately to the 3 eukaryotic protein homologues Gephyrin, Cnx1 and Cinnamon using the Bestfit function. The E-domain of CnxE was taken as the portion of the protein with similarity to *E. coli* MoeA and the G-domain was taken as the portion of the protein with similarity to *E. coli* MogA and MoaB, as denoted by the alignment detailed in Figure 25. The G-domain of CnxE was present at the N-terminal and stretched from amino acids 1-178, with the E-domain encompassing amino acid residues 237-704 at the C-terminal. In this respect CnxE resembles the eukaryotic proteins Gephyrin and Cinnamon which display the same order of domain organisation. A non-homologous linker region of 58 amino acids was identified between the two domains. Cnx1 was found to be 40.4% identical (61.7% similar) to the E-domain of CnxE and 40.2% identical (56.3%) similar to the G-domain of CnxE. When the average values of identity to both domains were taken into consideration Cnx1 was found to show more identity than Cinnamon, but less identity than Gephyrin to CnxE. The Cnx1 protein was therefore entered as being second most similar to CnxE in Table 5 between Gephyrin and Cinnamon, even although no values for the %similarity/%identity to the whole CnxE protein could be established.

Gephyrin was most similar to the full-length CnxE sequence, exhibiting 41.8% identity (61.1% similarity) (Table 5). As previously mentioned, Cnx1 was placed as being second most similar to CnxE due to the values obtained from comparisons with the CnxE E- and G-domains.

The other eukaryotic homologue Cinnamon was placed in third position with values of 34.4% identity (54.0% similarity) resulting from a comparison to full-length CnxE. The prokaryotic proteins with the highest similarity to CnxE were MoeA from *Anabaena* (strain PCC 7942) and MoeA from *E. coli* with 33.8% identity (51.8% similarity) and 33.4% identity (57.6% similarity) respectively. Putative MoeA protein sequences from several other prokaryotic species showed similarities to CnxE which ranged from 32.7% identity in *Salmonella typhimurium* to 29.2% identity in *Methanococcus jannaschii*. Interestingly, in the methanogenic bacterium *Methanococcus jannaschii* two separate proteins, MJ9886 and MJ0666, both claim similarity to MoeA sequences from other prokaryotes. The *E. coli* MoCo biosynthesis protein MogA showed 32.8% identity (52.6% similarity) to CnxE with *E. coli* MoaB the least similar of the three *E. coli* proteins exhibiting 28.8% identity (47.5% similarity) to CnxE. The protein MoaCB from *Synechococcus* (strain PCC 7942) was least similar to CnxE with 22.7% identity (45.4% similarity). MoaCB is similar in the N-terminal to *E. coli* MoaC and in the C-terminal to *E. coli* MoaB (Rubio *et al.*, 1998).

Figure 25 displays an alignment of CnxE with the eukaryotic proteins Gephyrin and Cinnamon, and the *E. coli* proteins MoeA, MogA and MoaB performed using Pileup and Prettybox functions on the Seqnet computer program. Due to the inverse arrangement of the Cnx1 domains it was not possible to align the Cnx1 molecule in this manner. The homology between the eukaryotic sequences is dispersed over the whole molecule. Figure 25 shows that the *E. coli* MogA and MoaB proteins are similar to the N-terminal of all three eukaryotic proteins and that *E. coli* MoeA is similar to the corresponding C-terminal regions.

Figure 25: Alignment of the deduced amino acid sequence of *A. nidulans* *cnxE* with Gephyrin from *R. norvegicus*, Cinnamon from *D. melanogaster* and MoeA, MogA and MoaB from *E. coli*.

Gaps (.) have been inserted to produce maximum alignment. The consensus sequence for the six proteins is shown. Note that the N-terminal regions of all three eukaryotic proteins (CnxE, Gephyrin and Cinnamon) are similar to MogA and MoaB from *E. coli*, while the corresponding C-terminal regions display similarity to *E. coli* MoeA. (Note that Figure 25 is contained on two sequential pages).

	1		50
CnxE <i>A. nidulans</i>	.....M	AEQRLKAAIL	VVSDTAFQDP SSDRVFGTLI DVFPADGPWE
Cinnamon <i>D. melanogaster</i>	.....	.MESITFGVL	TISDTCWQEP EKDTSGPILR QLIGET..FA
Gephyrin <i>R. norvegicus</i>	MATEGMILTN	HDHQIRVGV	TVSDSCFRNL AEDRSGINLK DLVQDPSLLG
MoeA <i>E. coli</i>	.....	.....	.....
MoaB <i>E. coli</i>	....MSQVST	EFIPTRIAL	TVSN..RRGE EDDTSGHYLR DSAQEA...G
MogA <i>E. coli</i>	.....	.MNTLRIGLV	SISDRASSGV YQDKGIPALE EWL TSA.LIT
<b>Consensus</b>	-----	-----RIG-L	TVSDT-F--P E-D-SGP-IR DL-QEA-L-G
	51		100
CnxE <i>A. nidulans</i>	KPIT..KIVP	DDVLQIQRAV	CDWTDSPSPV NLVLISGGTG FAVKDNTPEA
Cinnamon <i>D. melanogaster</i>	NTQVIGNIVP	DEKDIIQQEL	RKWID.REEL RVILTTGGTG FAPRDVTPEA
Gephyrin <i>R. norvegicus</i>	GTISAYKIVP	DEIEEIKETL	IDWCD.EKEL NLILTTGGTG FAPRDVTPEA
MoeA <i>E. coli</i>	.....	.....	.....
MoaB <i>E. coli</i>	HHVVDKAIVK	ENRYAIRAQV	SAWIAS.DDV QVVLITGGTG LTEGDOAPEA
MogA <i>E. coli</i>	PFELETRLIP	DEQAIIEQTL	CELVDEMS.C HLVLTTGGTG PARRDVTPEA
<b>Consensus</b>	-TIV--KIVP	DE--IIQQTL	CDWIDS--E- NLVLTGGTG FAPRDVTPEA
	101		150
CnxE <i>A. nidulans</i>	VTPLLHRHAP	GLVHGMIAAS	LKVTPFAMM. ARPVAGTRHK SLVITLPGSP
Cinnamon <i>D. melanogaster</i>	TRQLEKECP	QLSMYITLES	IKQTQYAAL. SRGLCGIAGN TLIINLPGSE
Gephyrin <i>R. norvegicus</i>	TKEVIEREAP	GMALAMLGMS	LNVTPLGML. SRPVCGIRGK TLIINLPGSK
MoeA <i>E. coli</i>	.....	.....	.....
MoaB <i>E. coli</i>	LLPLFDREVE	GFGEVFRMLS	FEEIGTSTLQ SRAVAGVANK TLIIFAMPGST
MogA <i>E. coli</i>	TLAVADREMP	GFGEQMRQIS	LHFVPTAIL. SRQVGVIRKQ ALIINLPGQP
<b>Consensus</b>	TLPLL-REAP	G-G--MRM-S	LKVTP-AML- SRPV-GIRGK TLIINLPGSP
	151		200
CnxE <i>A. nidulans</i>	KGAKENLEAV	IKLLPHACTQ	AAGADS..RT LHAG.....
Cinnamon <i>D. melanogaster</i>	KAVKECFQTI	SALLPHAVHL	IGDDVSLVRK THAEVQG...
Gephyrin <i>R. norvegicus</i>	KGSQECFQFI	LPALPHAIDL	LRDAIVKVKE VHDELEDLPS PPPPLSPPT
MoeA <i>E. coli</i>	.....	.....	.....
MoaB <i>E. coli</i>	KACRTAWENI	IAPQLDARTR	PCNFHPLKPK .....
MogA <i>E. coli</i>	KSIKETLEGV	KDAEGNVVH	GIFASVPYCI QLEGGPYVET APEVVAARFP
<b>Consensus</b>	K--KEC-E-I	IA-LPHAVTL	--DA---V-K -HAE----- -P-----
	201		250
CnxE <i>A. nidulans</i>	.GIKKLEAEA	GVKSESQHDH	HHHHHEHETH SHSHSHGHGH G.....
Cinnamon <i>D. melanogaster</i>	.....SAQK	SHICPHKTGT	GTDS.....
Gephyrin <i>R. norvegicus</i>	TSPHKQTEDK	GVQCEEEEE	KKDSGVASTE DSSSHITAA ALARKIPDSI
MoeA <i>E. coli</i>	.....	.....	.....
MoaB <i>E. coli</i>	.....	.....	.....
MogA <i>E. coli</i>	KSARRDVSE.	.....	.....
<b>Consensus</b>	-S--K--AEK	GV-CE-----	--DS----T- --S-SH----
	251		300
CnxE <i>A. nidulans</i>	.....HVVPR	AHTTSPERRS	NDPAAGAT..
Cinnamon <i>D. melanogaster</i>	.....	.....	.....
Gephyrin <i>R. norvegicus</i>	ISRGVQVLEP	DTASLSTTFS	ESPRAQATSR LSTASCPTPK VQSRCSSKEN
MoeA <i>E. coli</i>	.....	.....	.....
MoaB <i>E. coli</i>	.....	.....	.....
MogA <i>E. coli</i>	.....	.....	.....
<b>Consensus</b>	-----V-PR	-----S---S	--P-A-AT--
	301		350
CnxE <i>A. nidulans</i>	.....	.....RRYRE	SPYPMLSVDE ALRQVSAHTP EPEVIEVPVN
Cinnamon <i>D. melanogaster</i>	.....	.....RN	SPYPMLPVQE VLSLIF.... NTVQKT
Gephyrin <i>R. norvegicus</i>	ILRASHSAVD	ITKVARHRM	SPFPLTSMK AFITVLEMTPLGTEIINR
MoeA <i>E. coli</i>	.....	.....ME	FTTGLMSLDT ALNEMLSRVT PLTAQETLPL
MoaB <i>E. coli</i>	.....	.....	.....
MogA <i>E. coli</i>	.....	.....	.....
<b>Consensus</b>	-----	-----RR-RE	SPYP-LS-DE AL--VL--TP -L---EV---
	351		400
CnxE <i>A. nidulans</i>	IDLVGVIIE	DVYAAEAVPA	YLASIVDGYA VIAPESPDDG HSTKGIFFVA
Cinnamon <i>D. melanogaster</i>	ANL.NKIL.L	EMNAPVNIPP	FRASIKDGYA MKSTGFSGTK R.....VL
Gephyrin <i>R. norvegicus</i>	DGM.GRVLAQ	DVYAKDNLPP	FPASVKDGYA VRAADGPGDR F.....II
MoeA <i>E. coli</i>	VQCFGRILAS	DVVSPLDVP	FDNSAMDGYA VRLADIASGQ .....PLPVA
MoaB <i>E. coli</i>	.....	.....	.....
MogA <i>E. coli</i>	.....	.....	.....
<b>Consensus</b>	--L-GR-LA-	DVYAP-NVPP	F-ASIKDGYA VRAAD-PGD-----PVA

Figure 25 continued

401 450  
 CnxE *A. nidulans* SITHANEEGA LAPLEPTIA RITTGAPLPP NANAVVMVED TLLASSTPDG  
 Cinnamon *D. melanogaster* GCIAAGDSPN SLPLAEDECY KINTGAPLP. ....LQL. ....DKN  
 Gephyrin *R. norvegicus* GESQAGEQPT QIVMP.GQVM RVTTGAPIPC GADAVVQVED TELIRESDDG  
 MoeA *E. coli* GKSFAG.QPY HGEWPAGTCT RIMTGAPVPE GCEAVVMQEQ TEQMDNGVR.  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** G-S-AGEQP- --PLP-GT-- RITTGAP-P- GA-AVV-VED TEL----DDG

451 500  
 CnxE *A. nidulans* KEEATVEILT GEIKPNENVR QPGSDVALGS RILQRGDLIT PVGGEI.GLL  
 Cinnamon *D. melanogaster* QOESLVDIL. VEPQAGLDVR PVGYDLSTND RIFPA...LD PSPVVVKSL  
 Gephyrin *R. norvegicus* TEELFVRIL. VQARPGQDIR PIGHDIKRGE CVLAKGTHMG PSEI...GLL  
 MoeA *E. coli* .....FT AEVRSGQNR RRGEDISAGA VVFPAGTRLT TAEI...PVI  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** -EE--V-ILT VE-RPGQ--R P-G-DIS-G- R--PAGT-LT PSE----GLL

501 550  
 CnxE *A. nidulans* AATGTRTVKV FKKPVVGVLS TGDELVEHDD PRSLQGGQIR DSNRPSILSC  
 Cinnamon *D. melanogaster* ASVG.NRLIL SK.PKVAIVS TGSELCSRN Q..LTPGKIF DSNITMLTEL  
 Gephyrin *R. norvegicus* ATVGVTVEEV NKFPVAVMS TGNELLNPE D..LLPGKIR DSNRSTLLAT  
 MoeA *E. coli* ASLGIAEVPV IRKVRVALFS TGDELQLPGQ P..LGDGQIY DTNRLAVHLM  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** ASVG--EV-V -KKPVAV-S TGDEL--P-D P--L-PG-I- DSNR---L--

551 600  
 CnxE *A. nidulans* LKSWGIPAVD LGIARDTPAG ELEQSLRDAL RGVGKSNTSV DVIITGGVS  
 Cinnamon *D. melanogaster* LVYFGFNCMH TCVLSDTF.Q RTKESLLELF .....EVV DFVICSGGVS  
 Gephyrin *R. norvegicus* IQEHGYPTIN LGIVGDNP.D DLLNALNEGI .....SRA DVIITSGGVS  
 MoeA *E. coli* LEQLGCEVIN LGIIRDDP.H ALRAAFIEA. ....DSQA DVVISSGGVS  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** L--G-P-IN LGI-RDTP-- -L---L-EA- -----S-- DV-ITSGGVS

601 650  
 CnxE *A. nidulans* MGELDLLKPT IERSLGGTTH FGRVSMKPGK PTTFATVPFK PTSSAAGQOE  
 Cinnamon *D. melanogaster* MGDKDFVRSV LE.DLQFRIH CGRVNIKPGK PMTFASRK. ....  
 Gephyrin *R. norvegicus* MGEKDYLKQV LDIDLHAQIH FGRVFMKPGK PTTFATLIDID GV.....  
 MoeA *E. coli* VGEADYTKTI LEE..LGEIA FWKLAIKPGK PFAFGKL... ..  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** MGEKDYLK-V LE-DL-G-IH FGRV--KPGK PTTFATL--- -----

651 700  
 CnxE *A. nidulans* RSSRLIFSLP GNPASALVTL NLFVLP SLHK LIGLGQKQAA LGIAPALGLP  
 Cinnamon *D. melanogaster* ..DKYFFGLP GNPVSAFVTF HLFALPAIRF AAGWDRCKCS LSV.....  
 Gephyrin *R. norvegicus* ..RKIIFALP GNPVSAVVTC NLFVVPALRK MQGILDPRP. .TI.....  
 MoeA *E. coli* .SNSWFCGLP GNPVSATLTF YQLVQPLLAK LSGNTASGLP .....A  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** -S-K--FGLP GNPVSA-VTF NLFVLPALRK L-G----- L-I-----

701 750  
 CnxE *A. nidulans* LVAVTLSHAF PLDPKRTEYH RAIVTASPKD GRLYATSTGA EGVGQRSSRV  
 Cinnamon *D. melanogaster* .LNVKLLNDF SLDS.RPEFV RASVISKSGE .LYASVNG. ...NQISSRL  
 Gephyrin *R. norvegicus* .IKARLSCDV KLDP.RPEYH RCILTWHHQE PLPWAQSTG. ...NQSSRL  
 MoeA *E. coli* RQRVRTASRL KKTFRGLDFQ RGVLQRNADG ELEVTITG.. ..HQS  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** ---VRLS-DF KLDP-RPE-H RAI-T----E -LLYATSTG- ---NQ-SSRL

751 793  
 CnxE *A. nidulans* GSLASANSLI VLQPGKGSIA QGSLVEALMM GPIVREGAAV AL.  
 Cinnamon *D. melanogaster* QSIVGADVLI NLPARTSDRP LAKAGEIFPA SVL.RFDFIS KYE  
 Gephyrin *R. norvegicus* MSMRSANGLL MLPPKTEQYV ELHKGEVVDV MVIGRL....  
 MoeA *E. coli* SSFSLGNCFI VLERDRGNVE VGEWVEVEPF NALFGGL... ..  
 MoaB *E. coli* .....  
 MogA *E. coli* .....  
**Consensus** -S--SAN-L- VLPP-TG--- -G---EV-P- -V--R-----

#### **4.12 Sequence Analysis of the *cnxE* Gene from Mutant Strains**

Genomic DNA was extracted from various *cnxE* mutant strains and digested with the restriction enzyme *EcoRI* which was known to leave the *cnxE* coding region uncut. PCR was then used to amplify appropriate overlapping fragments of genomic DNA, containing the *cnxE* coding region to use as templates for sequencing. Table 6 summarises the location, nucleotide sequence and amino acid sequence of the *cnxE* mutations.

Where numbers specifying the positions of mutated nucleotide and amino acid residues are cited, they refer to the numbering system detailed for the wild-type *cnxE* sequence (Figure 23). In mutant strain *cnxE12*, a single base pair substitution was identified. The exchange of a guanine residue for an adenine residue at nucleotide position 1120 would be responsible for a change in the corresponding predicted amino acid sequence from glycine to aspartate at position 344. Mutant strain *cnxE849* was found to contain a single base pair substitution from guanine to thymine at nucleotide position 1206, which would ultimately change alanine 373 to serine. In the mutant *cnxE14* two separate single base pair substitutions were identified at nucleotide position 1376, where cytosine was changed for guanine and at nucleotide position 1380, where adenine was changed for guanine. A further mutation was identified within strain *cnxE14* with the deletion of an adenine residue at nucleotide position 1381. The cumulative effect of these mutations would be to cause a change in the predicted amino acid sequence from phenylalanine to leucine at position 429 and to shift the reading frame downstream of the deleted residue (affecting amino acid residues downstream of position 431) until early termination of translation at a stop codon reached at nucleotide position 1399 (where nucleotide 1399 represents the T of the TGA stop codon). In mutant strain *cnxE3*, the deletion of a single guanine

Table 6: Location, nucleotide sequence and amino acid sequence of *cnxE* mutations

Strain		Nucleotide Sequence <sup>abd</sup>		Amino Acid Sequence <sup>acd</sup>
Wild-type <i>cnxE12</i>	1130	ACA ACC GGC GCT CCT ACA ACC <u>GAC</u> GCT CCT	340	T T G A P T T <u>D</u> A P
Wild-type <i>cnxE849</i>	1200	GAA GAA GCG ACT GTC GAA GAA <u>TCG</u> ACT GTC	370	E E A T V E E <u>S</u> T V
Wild-type <i>cnxE14</i>	1371	GTA TTC AAG AAG CCC GTA <u>TTG</u> AAG <u>GXG</u> CCC	428	V F K K P V <u>L</u> K <u>G</u> P <sup>*(1399)</sup>
Wild-type <i>cnxE3</i>	1587	GGA GTC GGA AAG TCC GGA GTC GXA AAG TCC	500	G V G K S G V <u>E</u> <u>S</u> <u>P</u> <sup>*(1618)</sup>
Wild-type <i>cnxE16</i>	1650	GAG CTC GAT CTG CTA GAG CTC <u>GAA</u> CTG CTA	521	E L D L L E L <u>E</u> L L
Wild-type <i>cnxE13</i>	1710	CGT GTC TCC ATG AAA CGT GTX XCC ATG AAA	541	R V S M K R V <u>H</u> <u>E</u> <u>T</u> <sup>*(1979)</sup>

<sup>a</sup>Underlined letters indicate residues which are altered by the mutation.

<sup>b</sup>X represents the absence of the corresponding wild-type nucleotide in the mutant strain.

<sup>c</sup>\* indicates that the amino acid reading frame continues to be altered until early termination of the protein at a stop codon. The number in brackets represents the position of the first nucleotide in the stop codon.

<sup>d</sup>The numbers to the left of the DNA and amino acid sequences and the numbers in brackets after the \* symbol correspond to the numbering system detailed for the *A. nidulans* wild-type *cnxE* nucleotide and amino acid sequences (Figure 23).

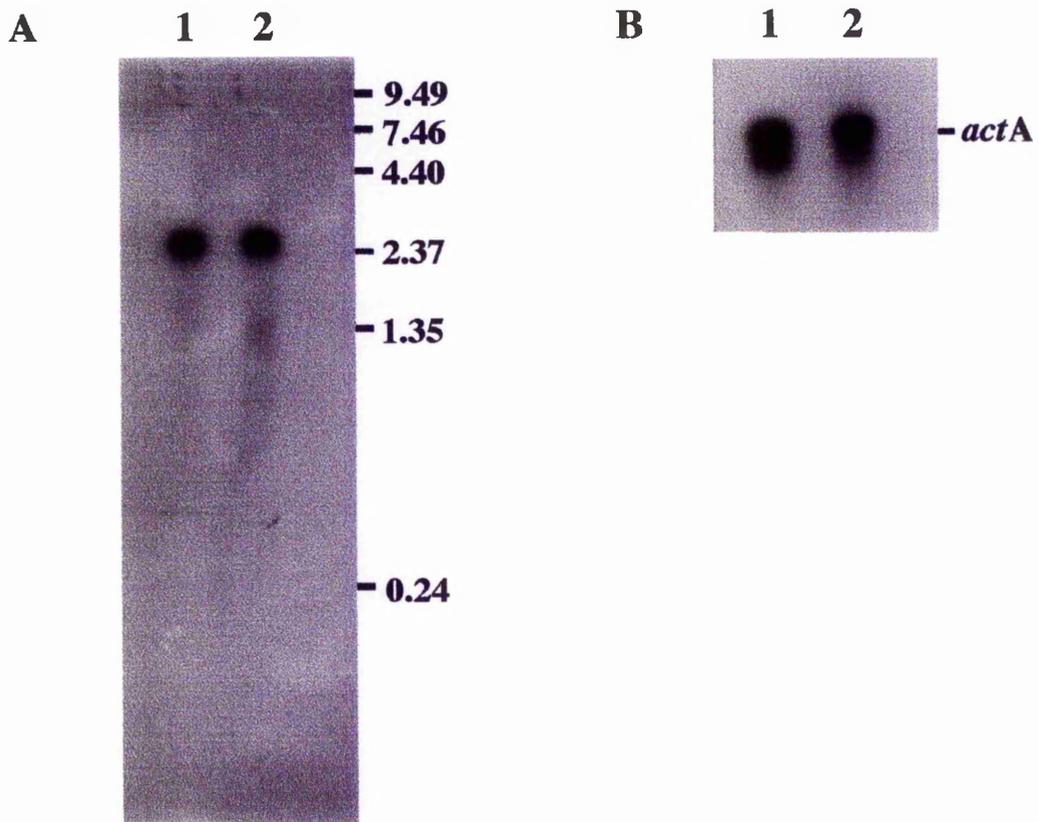
residue was identified at nucleotide position 1594. The removal of this base would result in a predicted reading frame shift from this point downstream until a premature stop codon is encountered at nucleotide 1618 (where nucleotide 1618 represents the T of the TAA stop codon). Translation of the *cnxE* gene within strain *cnxE3* would therefore be terminated at this point resulting in a truncated version of the CnxE molecule. In mutant strain *cnxE16*, a thymine residue was found to be substituted for an adenine residue at nucleotide position 1658, changing the inferred amino acid at position 523 from aspartate to glutamate. Sequence analysis of strain *cnxE13* revealed that 2 nucleotides, a cytosine and a thymine residue situated next to each other at positions 1715 and 1716 respectively had been deleted. The effect of this 2 bp deletion would be to cause a shift in the reading frame downstream of the mutation until early termination of translation at a stop codon, encountered at nucleotide 1979 (where nucleotide 1979 represents the T of the TGA stop codon).

Sequence analysis of the *cnxE* mutant strains was carried out by S.E. Unkles, Monash University, Melbourne.

#### **4.13 Northern Analysis of the *cnxE* Gene**

Northern analysis was carried out to determine the size of the *cnxE* transcript and to provide an indication of whether gene transcription was regulated by nitrate or ammonium as is the case with the nitrate reductase apoenzyme structural gene, *niaD*. mRNA was extracted from wild-type *A. nidulans* cells grown with either nitrate or ammonium as sole nitrogen source. The mRNA from both growth conditions was used in a Northern blot and probed with an approximately 1.5 kb PCR fragment taken from within the *cnxE* coding region of the cDNA clone pE1.

Figure 26: Northern analysis of the *cnxE* gene



Messenger RNA was extracted from wild-type *A. nidulans* cells grown in minimal medium with either 10 mM nitrate or 10 mM ammonium as sole nitrogen source. The mRNA was then used to produce a Northern blot which was hybridised with a 1.5 kb PCR fragment taken from within the *cnxE* gene coding region. Panel A shows the autorad after incubation with the hybridised Northern blot for 3 days at  $-70^{\circ}\text{C}$ . The molecular sizes (kb) of the RNA markers are indicated. The blot was stripped and rehybridised with a fragment specific for the constitutively expressed *A. nidulans actA* gene (Panel B). In both Panels, 1 = nitrate-grown cells and 2 = ammonium-grown cells.

The probe hybridised to a single 2.5 kb mRNA in both lanes (see Figure 26). The intensity of hybridisation appeared to be evenly matched between each of the mRNA samples. A 2.5 kb transcript indicates the existence of an approximately 0.4 kb untranslated leader sequence before the 2.1 kb *cnxE* gene coding region.

Rehybridisation with a fragment of the *A. nidulans* actin gene revealed that the distribution of mRNA between the 2 lanes of the Northern blot was reasonably equal. The expression of *cnxE* therefore does not appear to be controlled by nitrate or ammonium regulation.

#### 4.14 Testing for pE1-Complementation of the *moeA* Mutant Phenotype

An experiment was carried out in order to assess the ability of pE1 to functionally complement an *E. coli moeA* mutant strain. This was investigated because of the amino acid similarity observed between CnxE and the *E. coli* MoCo biosynthesis protein MoeA (see section 4.10). Competent cells were prepared from the *moeA* mutant strain NS9 (Rivers, 1991), and a 100 µl aliquot subsequently transformed with 0.5 µg pE1. To provide a positive control, competent cells were transformed with 0.5 µg pJR3, which contains the *moeAB* locus cloned into pUC8 (Reiss *et al.*, 1987). A negative control was provided by transforming NS9 competent cells without any DNA addition. Each sample of transformed cells was split into 3 equal portions, spread onto LN/ampicillin selection plates and incubated for 3 days under anaerobic conditions at room temperature, in an experiment analagous to that carried out in section 4.4, *i.e.* the retransformation of the *mogA* mutant strain with pE1.

After 3 days incubation under anaerobic conditions, the positive control plates containing the *moeA* mutant strain NS9 transformed with

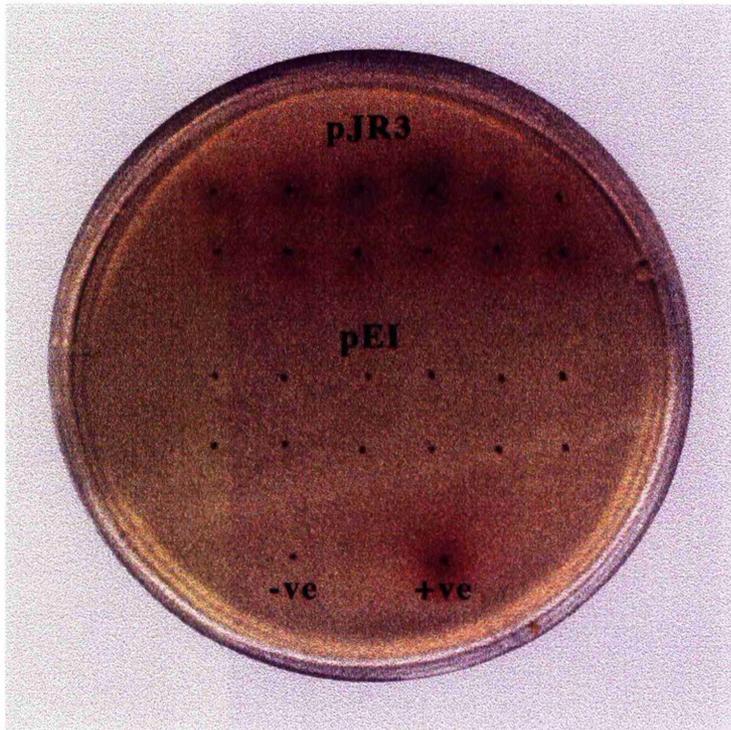
pJR3 possessed numerous (uncountable number), very small (approximately 0.5 mm in diameter), translucent bacterial colonies. The 'no DNA' negative control plates did not contain any colonies. The plates containing pE1-transformed NS9 cells also appeared to be devoid of any bacterial growth. Due to the strange appearance of the colonies on the positive control plates it was decided to firstly select for ampicillin resistance under aerobic conditions by appropriately pE1 or pJR3-complemented cells, and then to test such successfully transformed colonies in the nitrite overlay procedure to determine whether functional nitrate reductase activity had been restored to the *moeA* mutant.

As previously, 0.5 µg pE1 was used to transform *E. coli* strain NS9 competent cells. NS9 cells transformed with 0.5 µg pJR3 and without any DNA addition formed the positive and negative controls respectively. Each sample of transformed cells was split into three equal aliquots and spread on triplicate plates which were incubated overnight, aerobically, at 37 °C. The resulting plates were found to be covered in colonies which were each approximately 1 mm in diameter and beige in colouration. Twelve colonies were picked at random from the 3 plates containing cells transformed with pE1 and restreaked in small patches on a fresh LB plate alongside twelve colonies taken from the pJR3 plates which were included to act as a positive control. A single colony of untransformed strain NS9 was streaked out (in a single patch) on the same plate to provide a negative control, as was a colony of the wild-type *E. coli* strain MC4100 (Casadaban and Cohen, 1979) to act as a further positive control. The restreaked colonies were then tested in the nitrite overlay procedure (Figure 27).

The medium turned from colourless to pink around each of the twelve growth patches containing the *moeA* mutant transformed with pJR3 and around the wild-type growth patch indicating the presence of

NR activity in these samples. The twelve *E. coli* patches consisting of the *moeA* mutant transformed with pE1 did not change colour and appeared the same as the patch of growth containing the untransformed *moeA* mutant. This lack of colour change was indicative of the absence of nitrate reductase activity within these samples. The *A. nidulans* cDNA clone pE1 is therefore unable to functionally complement the *E. coli moeA* mutant strain NS9.

Figure 27: Testing for complementation of *E. coli* strain NS9 with pE1



Twelve single colonies containing pE1-transformed *E. coli* strain NS9 were picked at random and restreaked in small patches on a LB plate, alongside twelve colonies of pJR3-transformed strain NS9. A single colony of untransformed strain NS9 was streaked (in a single patch) on the same plate to provide a negative control (indicated by -ve on the figure), as was a colony of the wild-type *E. coli* strain MC4100 to provide a positive control (indicated by +ve on the figure). The restreaked colonies were subsequently incubated under anaerobic conditions, at room temperature, overnight and then tested using the nitrite overlay procedure (Glaser and DeMoss, 1971). A colour change in the medium surrounding the *E. coli* growth, from colourless to pink, was indicative of nitrate reductase activity.

## CHAPTER 5

### HPLC ANALYSIS OF *cnx* MUTANTS

#### 5.1 Brief Introduction

This chapter details the analysis of *A. nidulans cnx* mutant strains and the wild-type strain for cellular levels of precursor Z and molybdopterin, compounds which are formed as intermediates in the MoCo biosynthesis pathway (Johnson *et al.*, 1984; Wuebbens and Rajagopalan, 1993). As well as its presence as an intermediary compound molybdopterin is also formed by the dissociation of molybdenum from the molybdenum cofactor following denaturation of molybdoenzymes (Johnson *et al.*, 1980).

Although the oxygen sensitivity displayed by precursor Z and molybdopterin renders their direct measurement impossible, the presence and concentration of these molecules can be determined by detection and quantitation of their respective stable, inactive, oxidation products, namely compound Z and Form A. Precursor Z is converted to compound Z by either air or I<sub>2</sub>/KI oxidation (Johnson *et al.*, 1989; Wuebbens and Rajagopalan, 1993). Molybdopterin is converted to Form A by *in vitro* oxidation using I<sub>2</sub>/KI or another oxidising agent (Johnson *et al.*, 1984). Form A results from the loss of both S atoms and contains an acetylenic bond. A second derivative of molybdopterin, Form B, is produced when oxidation is carried out in air rather than with I<sub>2</sub> (Johnson *et al.*, 1984). During the conversion to Form B, one of the S atoms is lost while the other S atom forms a thiophene ring by linking to C-7 of the pterin ring. In this study, I<sub>2</sub>/KI oxidation was carried out in order to produce Form A. The Form A molecules were then dephosphorylated to produce Form A dephospho because removal of the phosphate group results in slower

elution and better separation of the compound during HPLC. Figure 28 details the conversion of precursor Z to compound Z and molybdopterin to Form A dephospho.

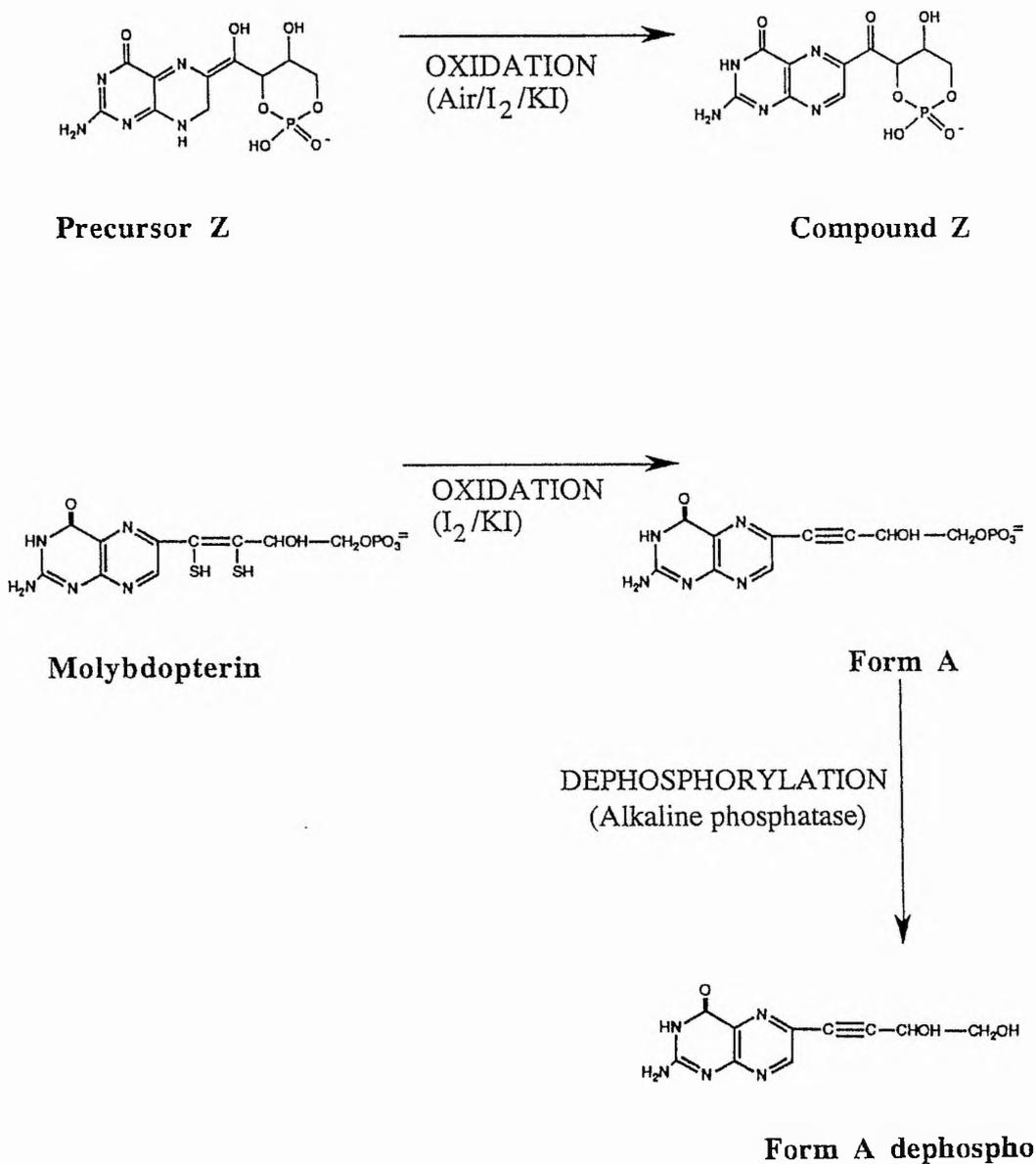
Both compound Z and Form A dephospho display fluorescence, characteristic of an oxidised pterin ring. The fluorescence exhibited by both compound Z and Form A dephospho permits the detection and identification of these compounds following separation by reversed phase HPLC.

The experiments detailed in this Chapter were carried out in collaboration with I. Heck, Dundee University.

## **5.2 HPLC Analysis of the *A. nidulans* Wild-type Strain Grown With Different Sole Nitrogen Sources**

The levels of precursor Z and molybdopterin *vis à vis* compound Z and Form A dephospho were determined in oxidised and dephosphorylated cell free extracts of wild-type cells grown with several different nitrogen sources. *A. nidulans* wild-type cells grown with nitrate, ammonium, adenine or nitrate plus proline as sole sources of nitrogen were analysed. Nitrate plus proline was tested to provide a control for comparison with the various mutant strains which were grown with proline to provide a usable source of nitrogen (*cnx* mutants cannot utilise nitrate as sole nitrogen source), but in the presence of nitrate to ensure induction of the molybdoenzyme nitrate reductase (sections 5.3, 5.4 and 5.5). The wild-type strain was tested with nitrate and adenine because these molecules represent substrates for the molybdoenzymes nitrate reductase and xanthine dehydrogenase respectively (Cove, 1979). Ammonium was tested because of its effect to repress the utilisation of alternative nitrogen sources *i.e.* to repress the synthesis of nitrate reductase and xanthine dehydrogenase (Cove, 1979).

Figure 28: Conversion of precursor Z to compound Z and molybdopterin to Form A dephospho



Precursor Z is converted to compound Z by either air or I<sub>2</sub>/KI oxidation. Molybdopterin is converted to Form A by oxidation using I<sub>2</sub>/KI (or an alternative oxidising agent). Form A is dephosphorylated with alkaline phosphatase to produce Form A dephospho.

### **5.2.1** Precursor Z Levels

The results presented in Figure 29 show the levels of precursor Z (measured as compound Z) present in the *A. nidulans* wild-type strain grown with different nitrogen sources. The levels of precursor Z were found to be approximately 3.5 times higher in nitrate grown cells than in ammonium grown cells and approximately 1.7 times higher in adenine grown cells than in ammonium grown cells. The wild-type cells grown with nitrate plus proline contained approximately 0.6 times less precursor Z than wild-type cells grown with nitrate alone, however still contained approximately 2.1 and 1.2 times more precursor Z than the ammonium and adenine grown cells respectively.

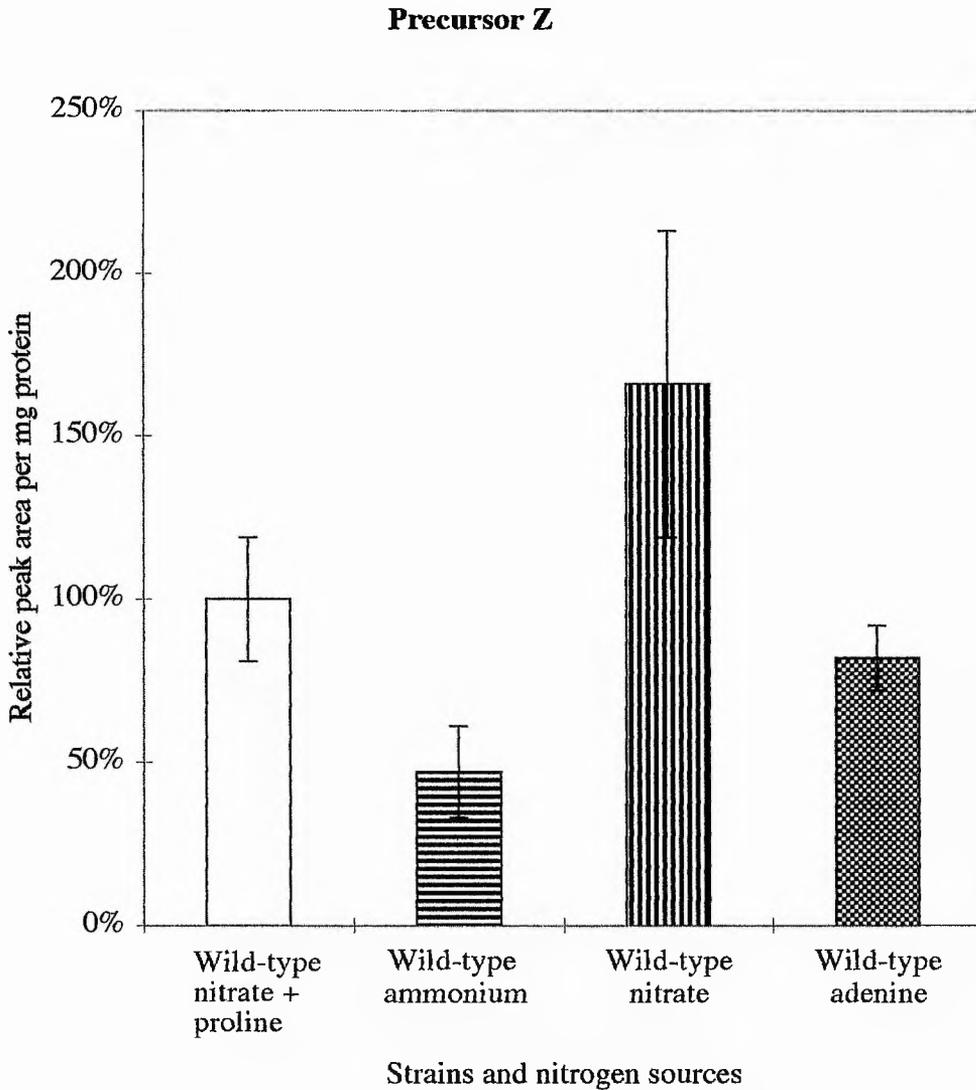
### **5.2.2** Molybdopterin Levels

The levels of molybdopterin (measured as Form A dephospho) present in the wild-type strain grown with different nitrogen sources is shown in Figure 30. Wild-type cells grown with nitrate plus proline contained the lowest amount of molybdopterin, with cells grown on ammonium, adenine and nitrate alone containing approximately 1.3, 1.4 and 1.7 times more molybdopterin respectively.

### **5.3** HPLC Analysis of *A. nidulans* *cnxA*, *cnxB* and *cnxC* Mutant Strains

Precursor Z and molybdopterin levels were estimated by measuring the cellular concentrations of compound Z and Form A dephospho in the *A. nidulans* *cnxA9*, *cnxB11* and *cnxC3* mutant strains, grown with nitrate plus proline as sole nitrogen source.

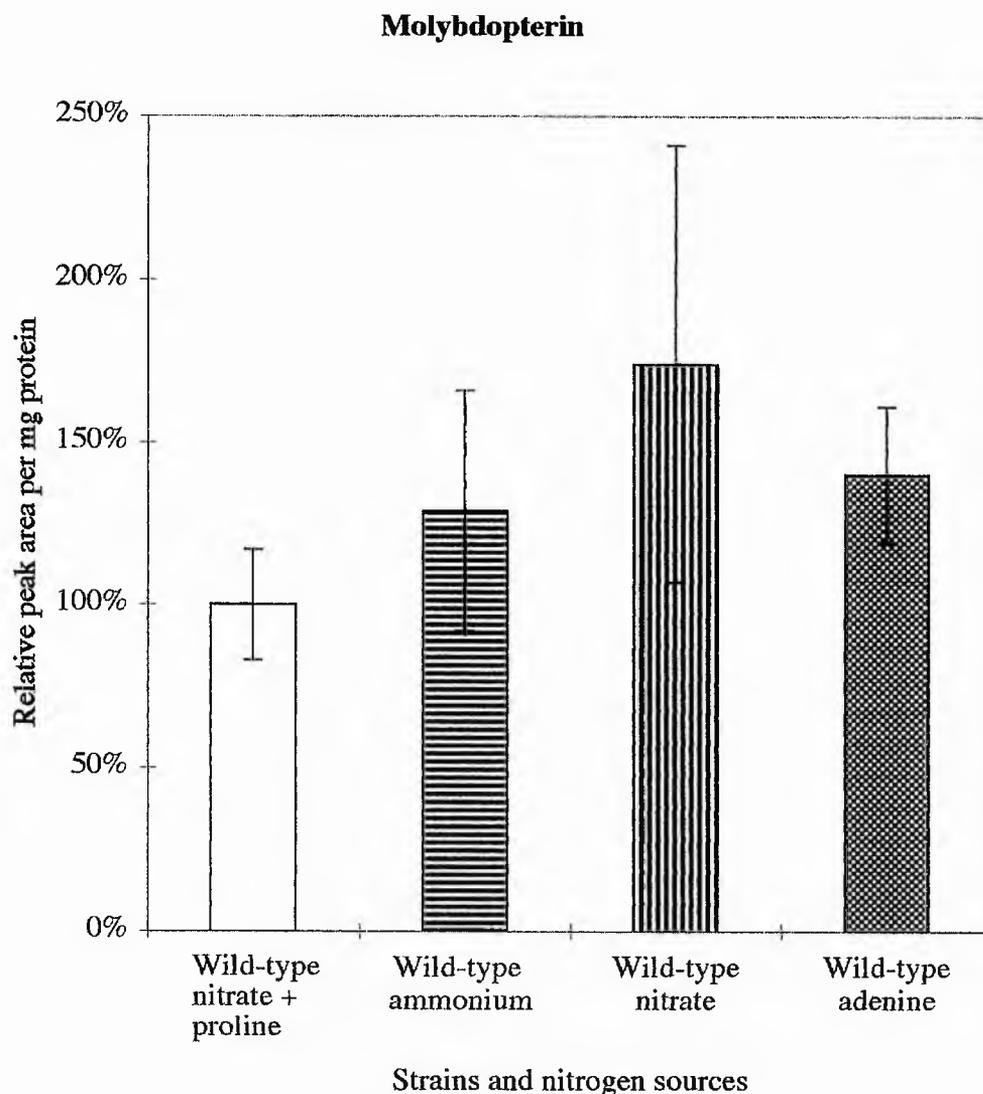
Figure 29: Levels of precursor Z detected in the *A. nidulans* wild-type strain grown with various different sole nitrogen sources



Precursor Z levels were measured as compound Z using reversed phase HPLC.

Fluorescent material co-chromatographing with purified compound Z was detected at 350/450 nm (emission/excitation). Three independent experiments were carried out for each nitrogen source. The levels of precursor Z were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain grown with nitrate plus proline. S.D. is indicated by the use of error bars. The equation used to calculate the S.D. values is detailed in the Appendix.

Figure 30: Levels of molybdopterin detected in the *A. nidulans* wild-type strain grown with various different sole nitrogen sources



Molybdopterin levels were measured as Form A dephospho using reversed phase HPLC. Fluorescent material co-chromatographing with purified Form A dephospho was detected at 370/450 nm (emission/excitation). Three independent experiments were carried out for each nitrogen source. The levels of molybdopterin were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain grown with nitrate plus proline. S.D. is indicated by the use of error bars.

### **5.3.1** Precursor Z Levels

The levels of precursor Z were found to be markedly decreased in each of the 3 mutant strains when compared to the wild-type strain grown under the same conditions (Figure 31). The *cnxA9*, *cnxB11* and *cnxC3* mutant strains yielded approximately 5.9, 5.3 and 3.2 times less precursor Z than the wild-type respectively.

### **5.3.2** Molybdopterin Levels

As can be seen from Figure 32, molybdopterin was not detected in strains *cnxA9*, *cnxB11* or *cnxC3*.

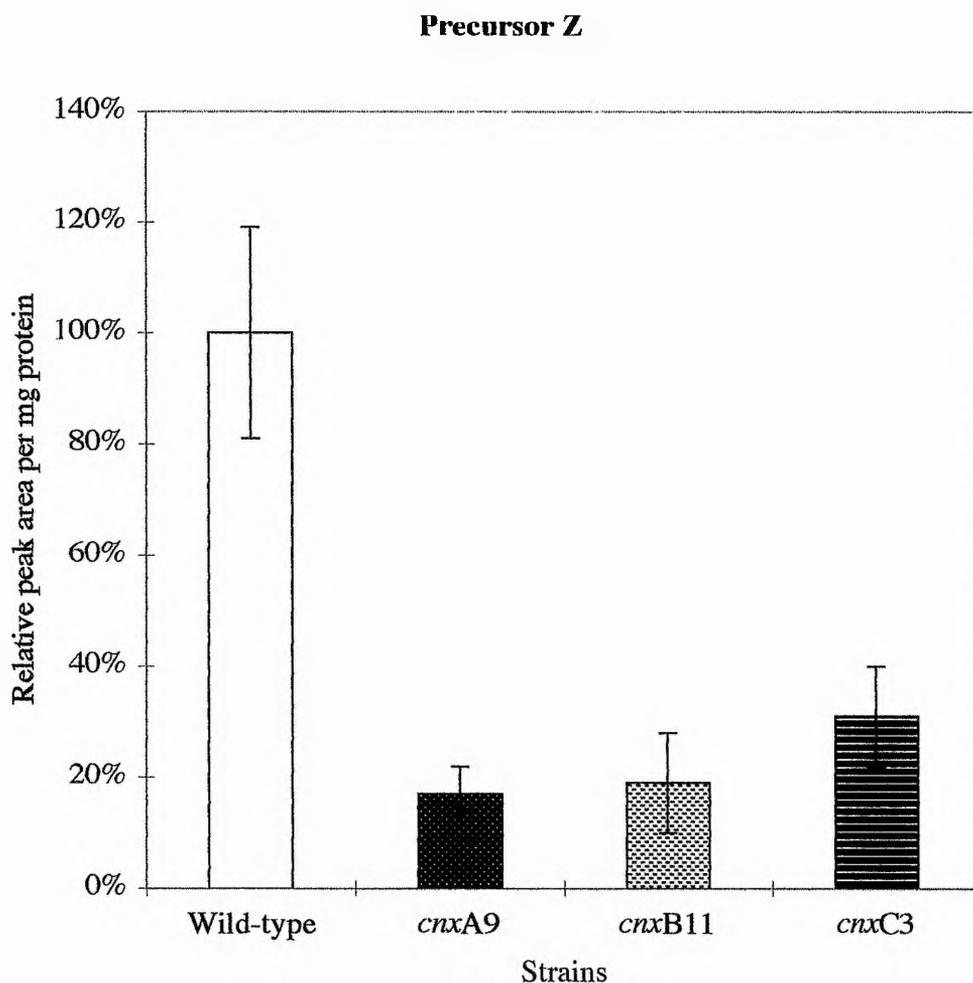
## **5.4** HPLC Analysis of *A. nidulans* *cnxG* Mutant Strains

Precursor Z and molybdopterin levels were determined by measuring the cellular concentrations of compound Z and Form A dephospho in 3 different *A. nidulans* *cnxG* mutant strains: *cnxG4*; *cnxG1222* and *cnxG2*, grown with nitrate plus proline as sole nitrogen source.

### **5.4.1** Precursor Z Levels

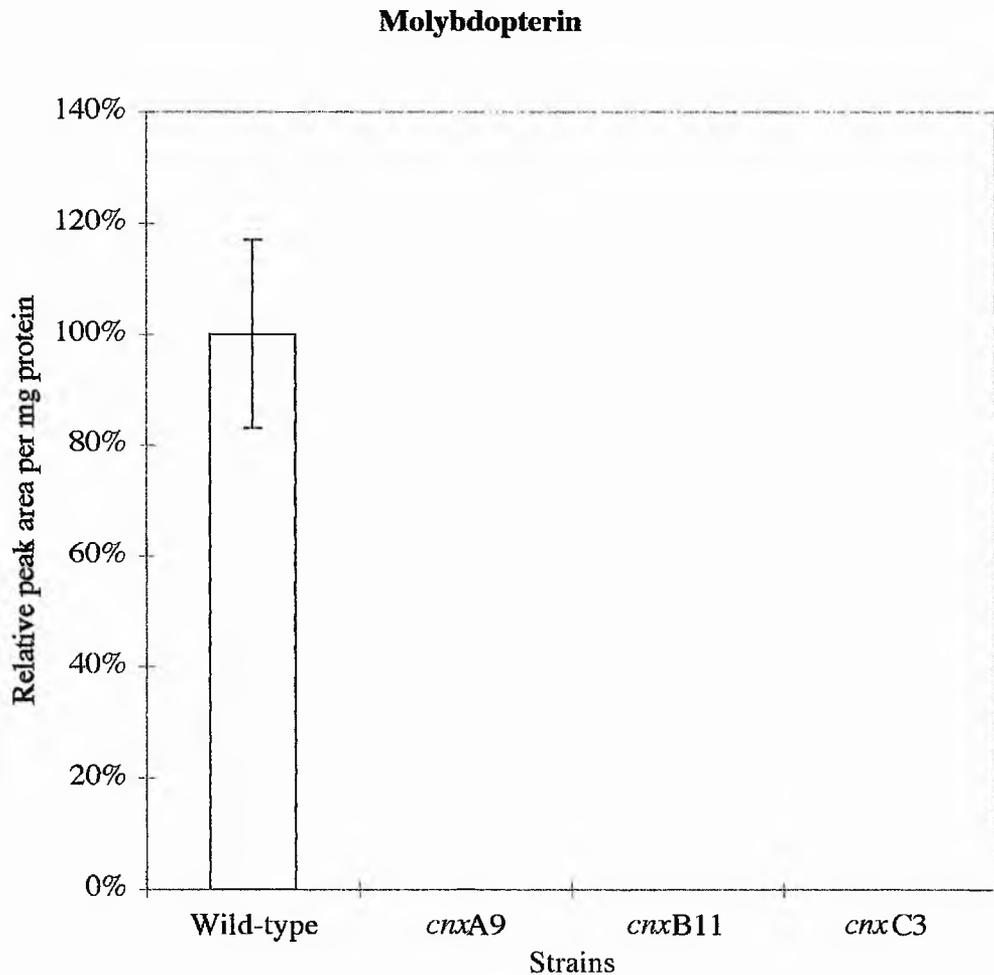
As depicted in Figure 33, precursor Z was found to be present at much higher levels in each of the *cnxG* mutant strains than in the wild-type strain grown under the same conditions. The *cnxG* mutant strains contained the following approximate increases in precursor Z concentrations when compared to wild-type cells: *cnxG4*, 25.3 times wild-type; *cnxG1222*, 17.2 times wild-type and *cnxG2*, 20.2 times wild-type.

Figure 31: Levels of precursor Z detected in the *A. nidulans* *cnxA9*, *cnxB11* and *cnxC3* mutant strains



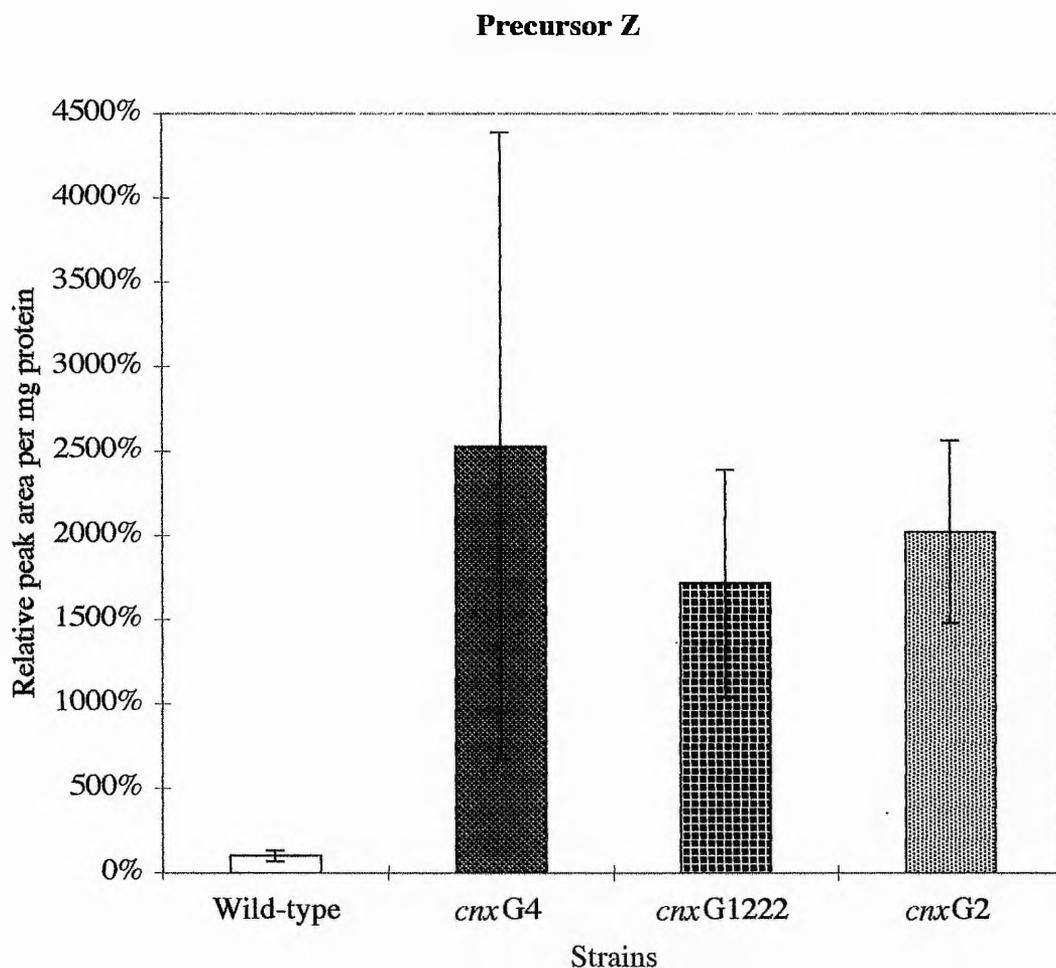
Precursor Z levels were measured as compound Z using reversed phase HPLC. Fluorescent material co-chromatographing with purified compound Z was detected at 350/450 nm (emission/excitation). Three independent experiments were carried out for each strain. The wild-type and all three mutant strains were grown with nitrate plus proline as sole nitrogen source. The levels of precursor Z were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain. S.D. is indicated by the use of error bars.

Figure 32: Levels of molybdopterin detected in the *A. nidulans* *cnxA9*, *cnxB11* and *cnxC3*



Molybdopterin levels were measured as Form A dephospho using reversed phase HPLC. Fluorescent material co-chromatographing with purified Form A dephospho was detected at 370/450 nm (emission/excitation). Three independent experiments were carried out for each strain. The wild-type and all three mutant strains were grown with nitrate plus proline as sole nitrogen source. The levels of molybdopterin were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain. S.D. is indicated by the use of error bars.

Figure 33: Levels of precursor Z detected in the *A. nidulans* *cnxG4*, *cnxG1222* and *cnxG2* mutant strains



Precursor Z levels were measured as compound Z using reversed phase HPLC. Fluorescent material co-chromatographing with purified compound Z was detected at 350/450 nm (emission/excitation). Three independent experiments were carried out for each strain. The wild-type and all three mutant strains were grown with nitrate plus proline as sole nitrogen source. The levels of precursor Z were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain. S.D. is indicated by the use of error bars.

#### **5.4.2 Molybdopterin Levels**

No molybdopterin was detected in the *cnxG4* and *cnxG1222* mutant strains (Figure 34). In the *cnxG2* mutant strain, 2% ( $\pm 1\%$ ) of the mean wild-type level of molybdopterin was recorded.

#### **5.5 HPLC Analysis of an *A. nidulans cnxE* Mutant Strain**

The levels of precursor Z and molybdopterin present in the *A. nidulans cnxE14* mutant strain were estimated by measuring the concentration of compound Z and Form A dephospho in cells grown with nitrate plus proline as sole nitrogen source.

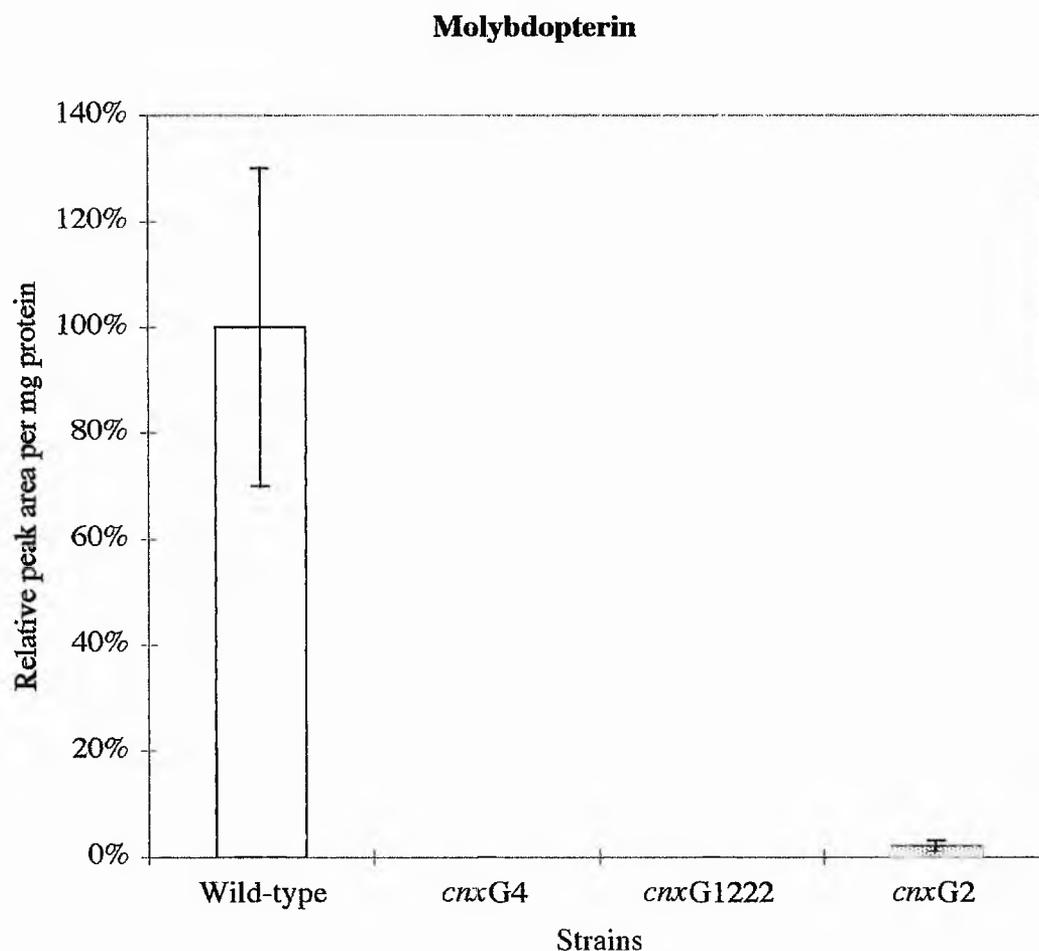
##### **5.5.1 Precursor Z Level**

HPLC analysis of the *cnxE14* mutant strain indicated that 1.6 times precursor Z was present in this strain when compared to the wild-type strain (Figure 35).

##### **5.5.2 Molybdopterin Level**

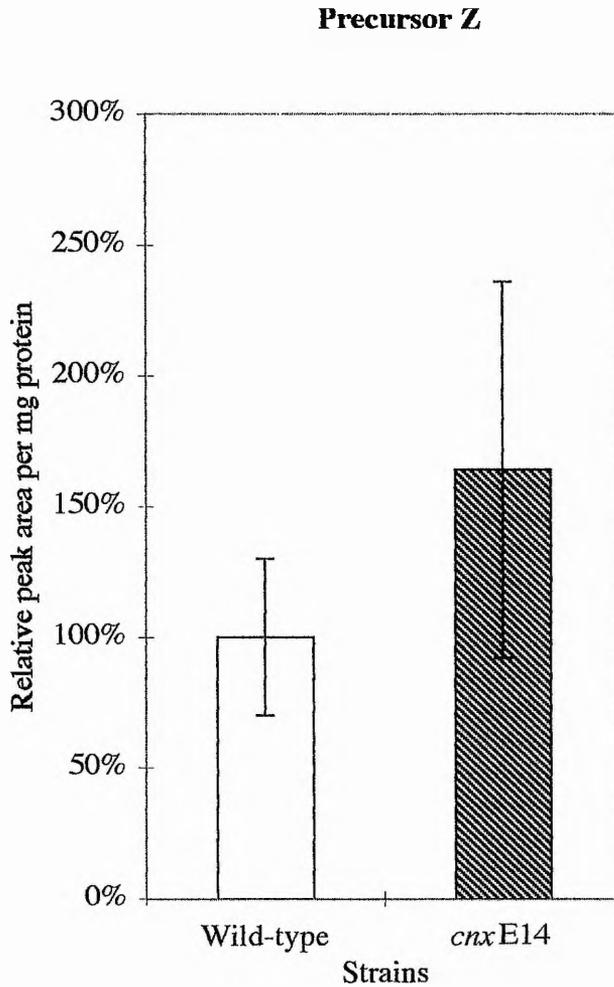
The *cnxE14* strain was found to contain approximately 2.3 times the amount of molybdopterin detected in the wild-type strain grown under the same conditions (Figure 36).

Figure 34: Levels of molybdopterin detected in the *A. nidulans* *cnxG4*, *cnxG1222* and *cnxG2* mutant strains



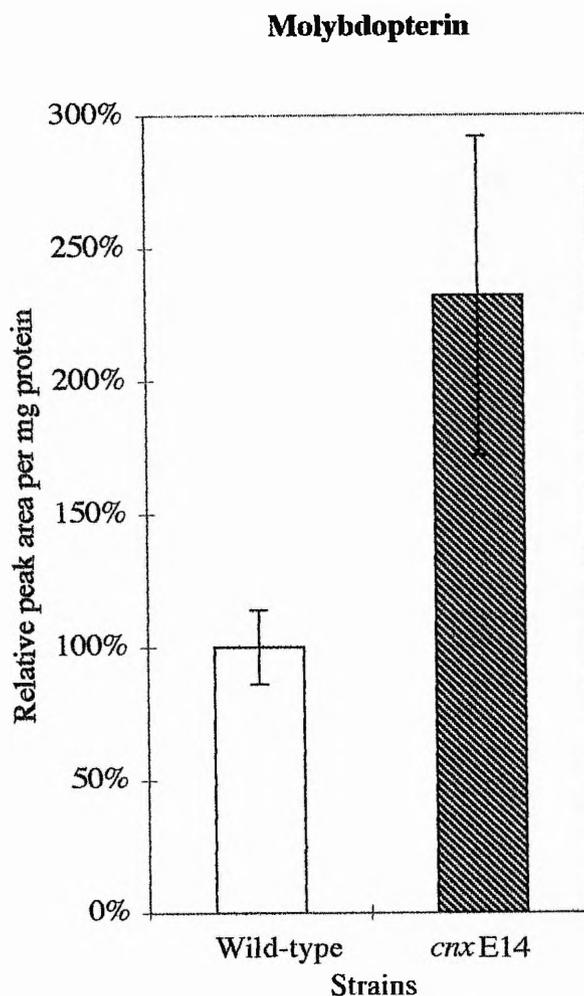
Molybdopterin levels were measured as Form A dephospho using reversed phase HPLC. Fluorescent material co-chromatographing with purified Form A dephospho was detected at 370/450 nm (emission/excitation). Three independent experiments were carried out for each strain. The wild-type and all three mutant strains were grown with nitrate plus proline as sole nitrogen source. The levels of molybdopterin were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain. S.D. is indicated by the use of error bars.

Figure 35: Levels of precursor Z detected in the *A. nidulans* *cnxE14* mutant strain



Precursor Z levels were measured as compound Z using reversed phase HPLC. Fluorescent material co-chromatographing with purified compound Z was detected at 350/450 nm (emission/excitation). Three independent experiments were carried out for both strains. The wild-type and the mutant strain were grown with nitrate plus proline as sole nitrogen source. The levels of precursor Z were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain. S.D. is indicated by the use of error bars.

Figure 36: Levels of molybdopterin detected in the *A. nidulans* *cnxE14* mutant strain



Molybdopterin levels were measured as form A dephospho using reversed phase HPLC. Fluorescent material co-chromatographing with purified form A dephospho was detected at 370/450 nm (emission/excitation). Three independent experiments were carried out for both strains. The wild-type and the mutant strain were grown with nitrate plus proline as sole nitrogen source. The levels of molybdopterin were calculated as relative peak areas of fluorescence per mg protein and expressed as percentages of the mean value for the wild-type strain. S.D. is indicated by the use of error bars.

## CHAPTER 6

### ATTEMPTS TO ISOLATE AN *E. coli moaD* GENE EQUIVALENT IN *A. nidulans*

#### 6.1 Brief Introduction

During the course of this study, the *A. nidulans cnxABC*, *cnxF* and *cnxH* genes were isolated and characterised (V.A. Appleyard and S.E. Unkles, personal communication). This means that all of the molybdenum cofactor biosynthesis genes originally identified by Cove and Pateman have been cloned and sequenced. From homology searches it was apparent that none of the isolated *A. nidulans cnx* genes were similar to the *E. coli moaD* locus. In *E. coli*, the *moaD* gene encodes a 10 kDa protein which forms the small subunit of MPT synthase (Pitterle and Rajagopalan, 1993). *cnxJ* was thought unlikely to encode the MPT synthase small subunit protein because the *cnxJ* gene product appears to be dispensible for cofactor synthesis with mutations within this locus only lowering, not abolishing, molybdoenzyme levels (Arst *et al.*, 1982). This chapter details the attempts made to isolate an *E. coli moaD* equivalent, from the *A. nidulans* genome.

#### 6.2 Cloning Strategy: Complementation of an *E. coli moaD* Mutant Strain with an Excised *A. nidulans* cDNA Library

Due to the earlier success in isolating the *A. nidulans cnxE* gene from an excised cDNA library by complementation of a *mogA* mutant strain (Chapter 4) it was thought that the same strategy might enable an *A. nidulans* gene (or genes) with the ability to functionally complement a *moaD* mutation to be detected. Electro-transformation was used to introduce the cDNA library into the *moaD* mutant cells because of the

high transformation frequencies achieved with this technique during the isolation of *cnxE* (Chapter 4, section 4.3.2).

*E. coli* strain MJ431 (Johnson and Rajagopalan, 1987b) was electro-transformed with 0.5 µg of the excised *A. nidulans* cDNA library (as used previously, Chapter 4), which had been dialysed over distilled H<sub>2</sub>O for 1 h. *E. coli* strain MJ431 was also electro-transformed with 0.5 µg dialysed pEM345 (which contains the *E. coli moaD* and *moaE* loci cloned into the expression vector pT7.5, refer to Rivers *et al.*, 1993) to act as a positive control. A negative control was provided by electro-transformation of strain MJ431 without any DNA addition. The DNA was dialysed in order to increase the purity which it was thought might contribute towards enhanced transformation efficiency.

Transformed cells were spread onto LN/ampicillin selection plates and incubated for 3 days under anaerobic conditions at room temperature. As described previously (Chapter 4, section 4.3.1) these conditions provided the selection of only those cells which were ampicillin resistant and which had the ability to use nitrate as terminal electron acceptor via functional nitrate reductase *i.e.* cells in which the *moaD* mutation had been successfully complemented by an appropriate cDNA clone. One µl of the MJ431 cell sample transformed with the cDNA library was spread onto an LN/ampicillin plate and incubated aerobically overnight at 37 °C to provide an indication of the total number of cDNA clones screened.

From the aerobic sample it was estimated that approximately 32000 clones had been screened during the electro-transformation using 0.5 µg library DNA. This satisfied the 30958 clones which needed to be screened in order to have a 99% probability of finding a single gene copy, as calculated in Chapter 4, section 4.3.1. After 3 days anaerobic incubation the negative control *i.e.* 'no DNA' plates showed no growth

while the positive control plates were covered in numerous colonies (uncountable number). The colonies growing on the positive control plates however appeared small (approximately 0.5 mm in diameter) and were translucent, as noted previously in Chapter 4, section 4.14 during the investigation of pJR3/pE1 transformation of the *moeA* mutant strain NS9.

On the selection plates containing MJ431 cells transformed with the excised cDNA library five small (approximately 0.5 mm in diameter), translucent *E. coli* colonies, similar to those present on the positive control plates, were visible after 3 days in an anaerobic environment. Each colony was removed and restreaked onto a fresh LB/ampicillin plate. The plates were then incubated aerobically overnight at 37 °C. Ten colonies from the positive control plates (*i.e.* the *moaD* mutant transformed with the *moaD* containing plasmid) were also restreaked and incubated in the same manner.

### **6.3 Testing the Putatively Complemented *moaD* Mutant Colonies for Nitrate Reductase Activity**

The five colonies resulting from electro-transformation of the MJ431 *moaD* mutant strain with the *A. nidulans* cDNA library were examined for nitrate reductase activity using the nitrite overlay procedure. The ten colonies restreaked from the plates containing the *moaD* mutant electro-transformed with pEM345, were also tested to act as a positive control. A further positive control was provided by the wild-type *E. coli* strain MC4100. The untransformed *moaD* mutant strain was used as the negative control. A single colony of each *E. coli* sample to be tested was restreaked as a small patch on a single LB plate which was then subjected to the nitrite overlay procedure.

A pink colouration was observed in the agar surrounding the wild-type MC4100 growth and around each of the ten pEM345-transformed MJ431 *moaD* mutant growth patches indicating the presence of functional nitrate reductase activity. The intensity of pink staining appeared similar between the wild-type and all ten of the pEM345-complemented mutant strain patches with any slight fluctuation in the pink colouration probably due to differences in the size of the patches and the amount of growth streaked. No nitrate reductase activity was observed in the untransformed mutant strain and in the five colonies resulting from electro-transformation of strain MJ431 using the excised *A. nidulans* cDNA library with the agar around each patch of bacterial growth remaining colourless. Further investigation of these five colonies and the plasmid DNA which they contained was not undertaken.

#### **6.4 Subsequent Attempts at Electro-transformation of Strain MJ431 with an Excised *A. nidulans* cDNA Library**

Two further attempts were made to complement the *E. coli moaD* mutant strain MJ431 with the excised *A. nidulans* cDNA library. In each attempt the experiment was carried out as detailed in section 6.2 using 0.5 µg of the excised cDNA library. In each case the aerobic control plates indicated that in excess of the 30958 colonies which needed to be screened in order to have a 99% probability of finding a single gene copy had been achieved. In the second of these attempts, the anaerobic incubation of plates was lengthened to 5 days and two activated catalysts were included within the anaerobic jar to ensure the establishment of anaerobic conditions.

In both electro-transformation attempts the negative *i.e.* 'no DNA' control plates were devoid of colonies and the positive control plates were completely covered in small (approximately 0.5 mm in

diameter), translucent colonies as observed in section 6.2. The longer incubation period of cells under anaerobic conditions using two catalysts did not seem to change the appearance or number of the *E. coli* colonies growing on any of the plates. In both attempts the test plates containing the *E. coli moaD* mutant strain MJ431 electro-transformed with the excised *A. nidulans* cDNA library did not contain any colonies.

Attempts to isolate an *E. coli moaD* gene equivalent in *A. nidulans* were therefore unsuccessful.

## CHAPTER 7

### DISCUSSION AND CONCLUSIONS

#### 7.1 The *A. nidulans* *cnxG* Locus

The *cnxG* gene has been isolated by functional complementation of the *A. nidulans* mutant strain *cnxG4*. Genomic DNA (in an *argB* bank) cotransformed with the autonomously replicating vector pHELP1 resulted in the formation of a replicating cointegrate plasmid pK which, after 'rescue' in *E. coli*, retained the ability to retransform the *cnxG* mutant strain. Subsequent Southern analysis of the recombinant vector pK showed that it contained sequences which hybridised to the *A. nidulans* *argB* gene sequence and the bacterial AMA1 and pUC19 sequences, confirming that recombination had taken place between the cotransformed DNA species. In *A. nidulans*, cointegrate plasmids are formed by homologous and non-homologous recombination as well as by end-to-end ligation of linear fragments (Aleksenko, 1994). As a result of these random recombination events it was possible that rearrangements had occurred so as to alter the *cnxG* gene sequence without affecting the gene function. The problem of gene rearrangement is a documented disadvantage of using this technique in gene isolation (reviewed in Kinghorn and Unkles, 1994) however, a fragment of pK which did not hybridise to any of the vector sequences used in the initial transformation allowed the eventual isolation of both genomic and cDNA clones of *cnxG* for further analysis.

Although several attempts were made to isolate *cnxG* by transformation experiments using either *Sau3AI* partially digested genomic DNA or the *argB* bank in conjunction with a variety of autonomously replicating vectors, and although many *E. coli*

transformants were produced as a result, only one such colony was able to repeatedly retransform the *cnxG4* mutant strain. The high number of *E. coli* transformants observed was probably due to the rescue of plasmids containing only autonomously replicating vector DNA, which may have lost (or rearranged) the formerly integrated *cnxG*-complementing sequence by additional recombination events. Transformation with cosmid clones containing *cnxG* proved difficult with only 1-11 nitrate-utilising transformants produced per 2  $\mu\text{g}$  DNA. This low frequency has been reported before (Johnstone, 1985), with frequencies of 10 transformants/ $\mu\text{g}$  DNA reported using a cosmid vector constructed by Yelton *et al.*, (1985).

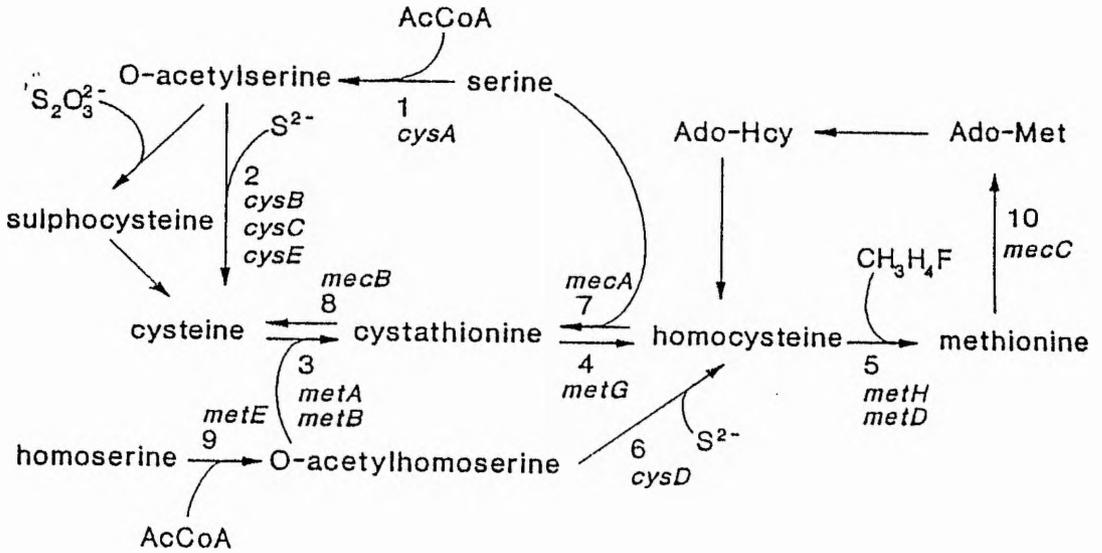
The *cnxG* locus consists of 1235 bp, encoding a protein of 396 amino acids with a predicted molecular weight of 43.56 kDa. The over-expression of *cnxG* resulted in the production of a protein with an apparent molecular weight of 53.5 kDa. The predicted molecular weight of the *cnxG* protein (396 amino acids) as inferred from the DNA sequence, with the addition of 19 amino acid residues coded for by pET21a vector sequences (including the His.Tag sequence), is 45.65 kDa (415 residues  $\times$  110, as calculated in section 3.10). This discrepancy in predicted and apparent protein weights may be due to the overall physico-chemical structure of CnxG which can effect protein mobility. Malelyated polypeptides, collagenases and other polypeptides with high proline contents (See and Jackowski, 1988), extremely basic proteins (Panyim and Chalkley, 1971) and very acidic proteins (Kaufmann *et al.*, 1984) give abnormally high molecular masses in SDS-PAGE and alterations of even a single amino acid within a protein can change the mobility in SDS-PAGE gels up to  $\pm 10\%$  (Seeburg *et al.*, 1984; Fasano *et al.*, 1985). Glycosylated proteins also exhibit artefactually high molecular mass estimates since SDS is only bound to the protein part of such molecules.

The solubility exhibited by the CnxG protein should remove the need for further protein solubilisation prior to future purification.

The CnxG protein is similar at the amino acid level to a variety of enzymes (from many different species) which are involved in catalysing transsulphuration reactions during the metabolism of cysteine, homocysteine and methionine. Such enzymes use a pyridoxal phosphate cofactor and all possess a conserved lysine residue found in the central section of the protein which forms the cofactor attachment site. The sequence around this residue is highly conserved and can be used as a signature pattern to detect this class of enzymes. The consensus pattern is [DQ]-[LIVM]-x(3)-[STAGC](2)-T-K-[FYW]-[LIVMF]-x-G-H-[SGN], where K is the pyridoxal phosphate attachment site (published as internet data only). The CnxG protein contains the pattern D-L-x(3)-S-G-S-K-Y-F-x-G-H-S, running from amino acid residues 198-212. The predicted protein length of 395 amino acids for the CnxG protein agrees fairly well with the typical size of an enzyme belonging to this family.

Transsulphuration pathways allow the interconversion of homocysteine and cysteine with the intermediary formation of cystathionine. Homocysteine can be used for the biosynthesis of methionine. An outline of sulphur amino acid metabolism in *A. nidulans* is presented in Figure 37. Biochemical diversity of transsulphuration reactions occurs among different species. In fungi, cysteine can be synthesised from O-acetyl-L-serine and sulphide *via* cysteine synthase (Paszewski and Grabski, 1974; 1975). This is the autotrophic pathway for sulphur utilisation in which inorganic sulphur is converted to organic sulphur, and shows similarity to the enteric bacterial and plant cysteine biosynthetic pathways (for reviews see Giovanelli (1987) and Soda (1987)). In fungi, plants and enteric bacteria cysteine sulphur is converted to homocysteine sulphur *via* transsulphuration consisting of cystathioine

Figure 37: Sulphur amino acid metabolism in *A. nidulans* (reproduced from Paszewski *et al.*, 1994)



1: serine transacetylase; 2: cysteine synthase; 3: cystathionine  $\gamma$ -synthase;  
 4: cystathionine  $\beta$ -lyase; 5: methionine synthase; 6: homocysteine  
 synthase; 7: cystathionine  $\beta$ -synthase; 8: cystathionine  $\gamma$ -lyase; 9:  
 homoserine transacetylase; 10: methionine adenosyltransferase.  
 AcCoA: acetylcoenzyme A;  $CH_3H_4F$ : methyltetrahydrofolate.

$\gamma$ -synthase and cystathionine  $\beta$ -lyase, and the homocysteine which is produced is used for the biosynthesis of methionine (Ono *et al.*, 1992). Cystathionine  $\gamma$ -synthase may alternatively catalyse the production of cystathionine from O-acetylhomoserine in fungi or from O-succinylhomoserine/O-phosphohomoserine in enteric bacteria and plants. Homocysteine may be produced directly from O-acetylhomoserine by homocysteine synthase or O-acetylhomoserine sulfhydrylase in fungal species. Fungal cysteine biosynthesis can also occur by the pathway of reverse transsulphuration. In this pathway cysteine is synthesised from methionine *via* S-adenosylmethioine (AdoMet), S-adenosylhomocysteine (AdoHcy), homocysteine and cystathionine. Mammals also utilise this pathway, but plants and enteric bacteria do not as far as we know (reviewed by Griffith, 1987).

The CnxG protein is similar in primary structure to cystathionine- $\gamma$ -synthase, methionine- $\gamma$ -lyase, cystathionine- $\gamma$ -lyase, O-acetylhomoserine sulfhydrylase, O-succinylhomoserine sulfhydrylase and cystathionine- $\beta$ -lyase from a number of different species. For the precise biochemical reactions catalysed by each individual enzyme refer to section 3.11. Transsulphuration enzymes are known to exhibit broad specificities: for example, deamination of L-serine may also be carried out by cystathionine  $\beta$ -lyase in *Escherichia coli* (Brown *et al.*, 1990). *In vivo* studies have also demonstrated the capacity of the yeast *Saccharomyces cerevisiae* homocysteine synthase to catalyse the conversion of ethionine to methioine (Cherest *et al.*, 1970). It could be suggested that CnxG also exhibits broad specificity and has evolved to catalyse a transsulphuration reaction during MoCo metabolism as well as catalysing similar reactions during cysteine/methionine biosynthesis, however if this were to be the case we might anticipate that the *cnxG* mutant phenotype would reflect the loss of all enzymatic functions carried

out by the *cnxG* gene product. *cnxG* mutants do not exhibit any growth requirement for sulphur or sulphur-containing amino acids which might be expected if an enzyme involved in cysteine/methionine biosynthesis was removed.

Alternatively, CnxG may be an enzyme which although originally involved in sulphur amino acid metabolism, has evolved to become dedicated to a sole function in MoCo biosynthesis. There are two major hypotheses which have been proposed to explain the evolution of metabolic diversity within multistep biosynthetic pathways: the first suggests that retrieval recruitment of new enzymes progressively built backwards from the final metabolite in the pathway (Horowitz, 1945); the second suggests that primitive enzymes possessed a very broad specificity which allowed later production of new enzymes following gene duplication (Jensen, 1976). *cnxG* does not exhibit similarity to any previously cloned MoCo biosynthesis genes. It seems likely that *cnxG* encodes a protein which has a similar function to the transsulphuration enzymes involved in the biosynthesis of sulphur containing amino acids yet carries out a distinct, perhaps unique, function in the MoCo biosynthesis pathway. Steps in pathways which appear to be unique or different in a particular species have been identified *e.g.* a yeast species designated PGR-13 has been shown to differ from all fungal/yeast species thus far tested in that it lacks the reverse transsulphuration pathway (Piotrowska and Paszewski, 1990). CnxG may represent a metabolic link between the pathways of sulphur amino acid biosynthesis and MoCo biosynthesis, in which cysteine/methionine derived sulphur is utilised during the anabolism of MoCo.

*cnxG* is similar to only one *A. nidulans* protein namely CysD (O-acetylhomoserine sulfhydrylase or homocysteine synthase) which catalyses the conversion of O-acetylhomoserine to homocysteine during

sulphur amino acid biosynthesis (Sienko, M., Topczewski, J. and Paszewski, A. (1995), unpublished, sequence submitted to database) as detailed in the following equation:



O-acetylhomoserine sulfhydrylase/homocysteine synthase, along with the enzymes cystathionine  $\beta$ -synthase and  $\gamma$ -cystathionase (cystathionine- $\gamma$ -lyase) form an 'alternative' pathway of cysteine synthesis in *A. nidulans* which operates when the main pathway of cysteine synthesis from O-acetylserine is impaired by mutation (Paszewski and Grabski 1974, 1975) (Figure 37). The CysD protein is 28.2% identical (50.1% similar) to *cnxG*. *cnxG* is not similar to any of the other enzymes involved in *A. nidulans* sulphur amino acid metabolism, however it would appear that *cysB* (encoding cysteine synthase) is the only gene to have been sequenced thus far (Paszewski *et al.*, 1994).

The initial donor of the reactive sulphur which is ultimately used to generate the dithiolene group unique to molybdopterin remains as yet unknown and has not been identified in any biological system. The hypothesis suggested here is that in *A. nidulans* CnxG is involved in the initial donation of sulphur to the MoCo biosynthesis pathway. In such a model, CnxG would donate sulphur to CnxF, which would in turn sulphurate the converting factor (for an overview of the *A. nidulans* MoCo biosynthesis pathway refer to section 7.5 and Figure 39).

This hypothesis is strengthened by the results obtained from HPLC analysis of three *cnxG* mutant strains, namely *cnxG4*, *cnxG1222* and *cnxG2*. Compound Z was accumulated to approximately 17.2-25.3 times the wild-type level in the *cnxG* mutants. Molybdopterin was absent from the *cnxG4* and *cnxG1222* mutants and was reduced to only 2%

( $\pm 1\%$ ) of the wild-type level in the *cnxG2* strain. The small amount of molybdopterin detected in the *cnxG2* mutant is very close to the detection limit of the HPLC and may actually represent an unrelated compound which is co-eluting with FormA dephospho. The accumulation of compound Z and abolition/reduction of molybdopterin suggests that the converting factor activity has been destroyed, possibly through the lack of sulphur donation to the *cnxF* gene product.

CnxG is more similar to Yhr2 than to any other protein sequence (showing 48.0% identity, 68.6% similarity) therefore the gene *yhr2* may represent the equivalent of *cnxG* in *S. cerevisiae*. Interestingly, no molybdenum cofactor biosynthesis genes have been identified in the yeast *S. cerevisiae*. Selected strains of *S. cerevisiae* are widely used in the manufacture of alcoholic beverages and fermented foods (Singleton and Sainsbury, 1993) therefore knowledge of the metabolic genes which function within this species is of great commercial significance. The recently cloned *A. nidulans* gene *cnxF* (which is thought to encode a sulphotransferase) is most identical to the protein encoded by the gene *yhr1* which lies immediately next to *yhr2* in *S. cerevisiae* (V.A. Appleyard, personal communication). The putative counterparts of *cnxG* and *cnxF* in *S. cerevisiae* are therefore located together within the genome. In *S. cerevisiae*, genes which carry out related or sequential functions in a pathway are sometimes found together in tightly linked clusters *e.g.* the gene clusters ARC and COR on chromosomes 5 and 10 of *S. cerevisiae* (Melnick and Sherman, 1993). Such gene clusters are also present in *A. nidulans*. Examples include the *crnA-niiA-niaD* gene cluster for nitrate assimilation (Tomsett and Cove, 1979; Greaves, 1989; Johnstone *et al.*, 1990; Kinghorn *et al.*, 1990; Unkles *et al.*, 1991), the four *prn* genes involved in proline utilisation (Arst and MacDonald, 1978; Hull, 1988; Sophianopoulou and Scazzocchio, 1989) and the three genes

representing the *npeA* locus required for penicillin biosynthesis (MacCabe *et al.*, 1990). In *A. thaliana*, the gene for cysteine synthase is located next to the 5' end of the genomic sequence for *cnx6* which encodes the large subunit of MPT synthase. The genes appear to be transcribed in opposite directions with only 1 kb separating the start codons of the two loci (Mendel, 1997). Perhaps in *A. thaliana*, the gene cysteine synthase can carry out a dual function in both MoCo biosynthesis and cysteine biosynthesis. Such a situation would differ from *A. nidulans* in which CnxG appears to be a distinct gene with a function dedicated to MoCo metabolism. Future studies will reveal if CnxG is indeed the initial donor of sulphur and whether this system is unique to *A. nidulans*.

The *cnxG4* mutant strain has been found to contain a single base pair substitution within the *cnxG* gene coding region. This single base pair change is predicted to result in the conversion of valine to phenylalanine at amino acid position 250 (using the numbering system detailed for the wild-type *cnxG* sequence (see Figure 12)). The mutated valine residue does not appear to be conserved throughout evolution as indicated by the alignment of CnxG with its five most similar sequences (Figure 13). The substitution does however exchange a small aliphatic, amino acid (valine) with a larger aromatic amino acid (phenylalanine). This may be responsible for a change in the structural folding, solubility or other behavior of the CnxG protein which renders it useless to function. An example of the drastic effects which can be caused by a single amino acid change is seen in the medical condition sickle-cell anaemia. In this molecular disease the substitution of valine for glutamate places a non-polar residue on the outside of haemoglobin S molecules which, when oxygenated, allow the formation of long fibrous precipitates that deform red blood cells giving them their sickle shape (Finch, 1978).

A mRNA transcript of 1.6 kb for *cnxG* was revealed in Northern analysis. This suggests a *cnxG* gene coding region of 1.2 kb following an approximately 0.4 kb untranslated leader sequence. The expression of *cnxG* does not appear to be regulated by the presence of nitrate or ammonium, with constitutive expression of the gene after cell growth on either nitrogen source. (For more discussion on the regulation of MoCo biosynthesis in *A. nidulans* refer to section 7.4). A second, faint RNA band of approximately 7.5 kb was also highlighted in Northern analysis however the lack of any similar bands in Southern analysis of digested total genomic DNA, led to the conclusion that this probably represented an artefact.

Another anomaly which emerged during this research was the isolation of cDNA clones which were longer than the genomic clone of *cnxG*, their sequences stretching much further in the 5' direction. Three of the cDNA clones of *cnxG* were approximately 1.3 kb, 2.7 kb and 3.0 kb longer than the *cnxG* gene coding region. One possible explanation for the long cDNA clones is that during cDNA bank construction, for an unknown reason, the reverse transcriptase did not function in the correct manner. It is known that the DNA synthesising activity of reverse transcriptase is error prone. The enzyme seems to switch templates during polymerisation by a 'copy choice' mechanism. In this 'copy choice' model recombination occurs during nucleic acid synthesis and involves the transfer of the polymerase from one template strand to another to form a new recombinant daughter strand (Singleton and Sainsbury, 1993). It is possible that recombination occurred in this way during cDNA bank construction to form the longer clones observed but it is unlikely that this would happen at such a high frequency and would surely not produce identical clones. This explanation is however supported by the

isolated *cnxE* cDNA clone pE1 in which two different cDNA clones appear to have become joined (Figure 20 shows a schematic map of pE1).

It could be proposed that the long cDNA clones contain an extra part of the *cnxG* gene which extends further in the 5' direction, however the plasmid pLJM18 which contains the 1.2 kb *cnxG* gene coding region plus only a further 572 bp upstream of this gene is able to complement the *cnxG4* mutant to high frequency, suggesting that the entire gene is contained on the cloned fragment. The sequenced 1kb portion of the three long cDNA clones which extend past the end of the genomic clone pLJM18 does not appear to contain any open reading frames, as revealed by DNA sequence analysis and Northern analysis of this region (section 3.8).

## 7.2 The *A. nidulans* *cnxE* Locus

The *cnxE* gene has been isolated by functional complementation of an *E. coli* *mogA* mutant strain. This method has previously been used to isolate other MoCo biosynthetic genes. The *A. thaliana* genes *cnx2* and *cnx3* were isolated by functional complementation of *E. coli* *moaA* and *moaC* mutant strains respectively (Hoff *et al.*, 1995). Another *A. thaliana* MoCo biosynthetic gene *cnx1* was isolated by functional complementation of an *E. coli* *mogA* mutant in the same manner (Stallmeyer *et al.*, 1995). Indeed, the first gene successfully cloned from *A. nidulans* was isolated using this technique (Kinghorn and Hawkins, 1982). The success of this method in cloning *cnxE* was facilitated by the use of electroporation to introduce DNA into the bacterial cells. Electroporation resulted in an approximately 14-fold increase in transformation efficiency when compared to that obtained using conventional bacterial transformation. This is in agreement with the data of Dower and colleagues (1988). These workers found that the efficiency of electro-transformation is in general

10 to 20 times higher than that obtained with maximally competent cells prepared by chemical methods.

Attempts to isolate *cnxE* by functional complementation of *A. nidulans* mutant strains *cnxE14* and *cnxE10* using either *Sau3AI* partially digested genomic DNA or the *argB* bank in conjunction with a variety of autonomously replicating vectors were unsuccessful. Many *E. coli* transformants were produced using total DNA isolated from each of the five initial *A. nidulans* transformants, however none of the plasmids harboured by these colonies possessed the ability to recomplement the *cnxE14* mutant strain. The high number of *E. coli* transformants observed was probably due to the rescue of plasmids containing only autonomously replicating vector DNA, which may have lost (or rearranged) the formerly integrated *cnxE*-complementing sequence by additional recombination events.

The cDNA clone which complemented the *E. coli mogA* mutation was composed of two DNA 'regions' each ending in their own poly(A) tail. Further analysis found that no sequence matching the second region of pE1 was present within the chromosome II portion of the cosmid library (where *cnxE* has been mapped). Moreover, genomic cosmid clones from the chromosome II portion of the library which were similar to the first region of pE1, contained a sequence which deviated from that of pE1 at the position of the first base in the poly(A) tail. From this data and from similarity searches it was concluded that the first region of pE1 contained the gene *cnxE*. The structure of pE1 supports the proposal suggested in section 7.1, that during cDNA library construction two different cDNA clones from different regions of the genome can become joined due to the reverse transcriptase enzyme switching templates during polymerisation by a 'copy choice' mechanism (Singleton and Sainsbury, 1993). The second region of pE1 exhibited homology to

only one protein, a hypothetical 105.9 kDa protein termed Ybt6 from *Saccharomyces cerevisiae* which is a putative integral membrane protein. The gene represented by the second region of pE1 does not appear to be located on chromosome II (as indicated by the failure of a region 2 DNA fragment to hybridise to any clone from the chromosome II portion of the cosmid library). The chromosomal location of this gene will probably be determined by future genetic studies, perhaps by the *A. nidulans* genome sequencing project which is currently underway in several laboratories.

The *cnxE* gene consists of 2130 bp, encoding a protein of 710 amino acids with a predicted weight of 78.1 kDa.

*cnxE* is similar to three *E. coli* proteins involved in MoCo biosynthesis (in order of decreasing similarity): 1) MoeA (Nohno *et al.*, 1988); 2) MogA (Reiss *et al.*, 1987; Yura *et al.*, 1992; James *et al.*, 1993) and 3) MoaB (Rivers *et al.*, 1993). *cnxE* is also similar to putative MoCo biosynthesis proteins from the bacterial species *Salmonella typhimurium* (K.K. Wong and H.S. Kwan, unpublished, sequence submitted to data bases) and *Methanococcus jannaschii* (Bult *et al.*, 1996) and the cyanobacterial species *Anabaena*, strain PCC 7120 (Ramaswamy *et al.*, 1996), *Synechococcus*, strain PCC 7942 (Rubio *et al.*, 1998) and *Synechocystis*, strain PCC 6803 (Kaneko *et al.*, 1995).

The *cnxE* gene is similar to three eukaryotic proteins (in order of decreasing similarity): 1) Gephyrin, a rat neuroprotein which functions to link the mammalian inhibitory glycine receptor to subsynaptic microtubules (Prior *et al.*, 1992; Kirsch and Betz, 1993; Kirsch *et al.*, 1993); 2) Cnx1, a protein from the plant *Arabidopsis thaliana* which is thought to be involved in MoCo biosynthesis although the specific function remains unknown (Stallmeyer *et al.*, 1995) and 3) Cinnamon, a protein from *Drosophila melanogaster* which is again involved in MoCo

biosynthesis (specific function unknown) (Kamdar *et al.*, 1994). The similarity exhibited by *cnxE* to each of the eukaryotic proteins stretches over the entire length of the molecule.

Stallmeyer and colleagues (1995) proposed that Cnx1, Gephyrin and Cinnamon contain two domains, termed the E- and G-domains, which are connected by a non-homologous 'linker' region. The E-domain is homologous to the *E. coli* MoeA protein and the G-domain is homologous to the *E. coli* MogA and MoaB protein molecules. In the *A. nidulans* CnxE protein two such domains are present (Figure 38). The G-domain is situated at the N-terminal of CnxE (stretching from amino acids 1-178, using the numbering system detailed for the wild-type *cnxE* sequence (Figure 23)) while the E-domain is situated at the C-terminal of the protein (encompassing amino acid residues 237-704) and in this respect CnxE resembles Gephyrin and Cinamon. In Cnx1 however the order of the two domains is inverted with the E-domain located at the N-terminal and the G-domain located at the C-terminal of the protein molecule (Stallmeyer *et al.*, 1995). It is not clear why the domain organisation should be reversed in *A. thaliana*.

A conserved amino acid signature pattern has been found within the Cinnamon, Gephyrin, MoaB and MogA proteins. The signature pattern is [LIVM](3)-x-T-G-G-T-G-x(4)-D (published as internet data only). In CnxE the pattern L-V-L-I-S-G-G-T-G-x(4)-D is present, running from amino acid residues 71-84. The signature pattern in CnxE is therefore slightly different to that already proposed with a serine (S) residue replacing the first threonine residue (T) (Table 7). A non-homologous linker region of 58 amino acids was identified between the two domains of CnxE. Within this region a histidine rich sequence is located running from amino acid residues 188 to 212 as detailed below: HDHHHHHHHEHTHSHSHSHGHGHGH.

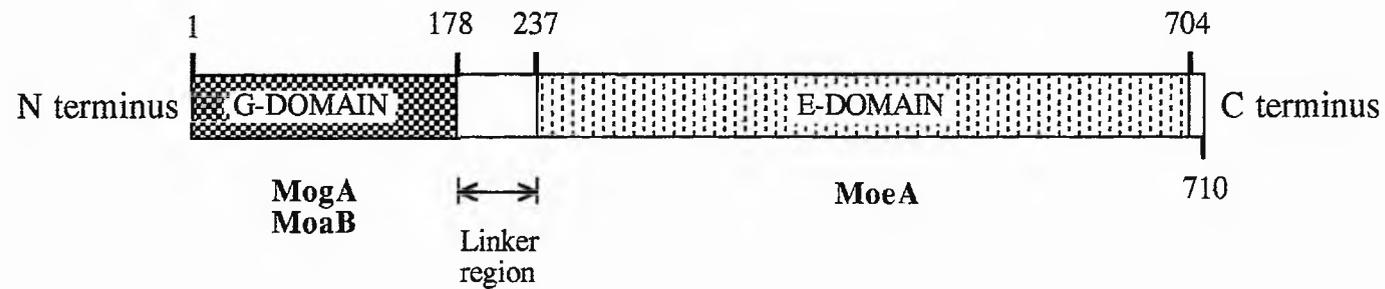
This sequence may have some functional significance, perhaps as a post-translational cleavage site for the production of two individual proteins corresponding to the E- and G-domains of the CnxE molecule.

During the course of evolution it appears that the genes encoding *E. coli* MogA, MoeA and MoaB proteins have become fused to produce the single *A. nidulans* *cnxE* gene encoding a two-domain multifunctional protein. In this regard *cnxE* resembles the *cnxABC* locus (see section 7.3). The joining together of these genes suggests that they carry out related or sequential functions during MoCo biosynthesis. The fusing of functionally related genes over evolution has been observed in other instances such as animal fatty acid synthesis (Amy *et al.*, 1992). Curiously in *E. coli*, where genes of related functions are often grouped together in operons, *moeA* and *moeB* are transcribed from a common promoter (Nohno *et al.*, 1988), *moaB* is found as part of the five-gene *moaA-moaE* operon (Rivers *et al.*, 1993) and *mogA* is held as a single gene locus ((Reiss *et al.*, 1987; Yura *et al.*, 1992; James *et al.*, 1993).

CnxE can replace the bacterial MogA protein function. Even although the cDNA clone pE1 lacked the first 11 bp before the translational start codon in the *cnxE* coding region sequence the ability to compensate for the *E. coli* *mogA* mutation was still conserved. Both the *cnxE* and *mogA* loci are molybdate repairable (Arst *et al.*, 1970; Stewart and MacGregor, 1982; Miller and Amy, 1983) and it seems likely that the function of *E. coli* MogA has been conserved in the N-terminal domain of *A. nidulans* CnxE. As previously discussed, the MogA protein is thought to have a role in intracellular molybdate processing or the insertion of molybdate into the cofactor, although the specific function remains unknown.

Although the N-terminal domain of CnxE is 33.4% identical (57.6% similar) to *E. coli* MoeA, the *A. nidulans* protein was found

Figure 38: Domain organisation in the *A. nidulans* CnxE protein



The approximate location of the G- and E- domains (shaded) in the CnxE protein molecule as determined by similarity to the individual *E. coli* proteins MogA, MoaB and MoeA. The numbers correspond to amino acid residues and open boxes represent regions of little or no similarity. The inter-domain linker region is highlighted. The same domain organisation is present in Gephyrin (*R. norvegicus*) and Cinnamon (*D. melanogaster*), however in the Cnx1 protein (*A. thaliana*) the order of the two domains is inverted.

Table 7: Signature pattern for proteins with similarity to *E. coli*  
MogA and MoaB

CnxE	<i>A. nidulans</i>	LVL	I	S	GGTG	EAVK	D
Cinnamon	<i>D. melanogaster</i>	VIL	T	T	GGTG	EAPR	D
Gephyrin	<i>R. norvegicus</i>	LIL	T	T	GGTG	EAPR	D
Cnx1	<i>A. thaliana</i>	LIL	T	L	GGTG	FTPR	D
MoaB	<i>E. coli</i>	VVL	I	T	GGTG	LTEG	D
MogA	<i>E. coli</i>	LVL	T	T	GGTG	EAPR	D
<hr/>							
Signature pattern		[LIVM](3)	X	X?*	GGTG	X(4)	D

\*Although the proposed signature pattern for this group of proteins (before inclusion of the CnxE and Cnx1 proteins, published as internet data only) indicated the presence of a conserved threonine (T) residue at this position, the CnxE and Cnx1 amino acid sequences display a serine (S) residue and a leucine (L) residue respectively.

unable to take over the activity of its bacterial counterpart. The same situation has been reported with Cnx1 from *A. thaliana*: a cDNA clone containing *cnx1* was found unable to functionally complement an *E. coli moeA* mutant (Stallmeyer *et al.*, 1995). In *E. coli* no specific function has yet been assigned to MoeA. However, a recent study using a unique class of *E. coli moeA* mutants, which possess molybdoenzyme activities only when grown with limited amounts of sulphur compounds, has led to the suggestion that MoeA is responsible for generating a form of 'activated' molybdenum produced by the combination of sulphide with molybdate (Hasona *et al.*, 1998). It has been speculated that this 'sulphide-activated' molybdenum is a type of thiomolybdenum which can be used in the formation of molybdenum cofactor from Mo-free molybdopterin. Another study by Menéndez and colleagues has revealed that the overexpressed and purified *A. nicotinovorans* MoeA protein associates into high-molecular-mass complexes and spontaneously forms gels at concentrations of 1 mg/ml or above (Menéndez *et al.*, 1997). It was also shown that MoeA has adenosinetriphosphatase activity, forms fibrillar structures and binds to neurotubulin and tubulin dimers (Menéndez *et al.*, 1997).

Studies using overexpressed and purified Cnx1 (the separate E- and G- domains and the entire protein) from *A. thaliana* have recently revealed a high affinity interaction between molybdopterin and the G-domain of Cnx1 (which is similar to *E. coli* MogA) (Schwarz *et al.*, 1997). The E-domain of Cnx1 (which is similar to *E. coli* MoeA) has also been shown to bind molybdopterin but with much lower affinity and in addition appears to possess a cytoskeleton-binding function (Schwarz *et al.*, 1997). Interestingly, the Cnx1 (and CnxE) homologue Gephyrin functions to fix neuroreceptors to the cytoskeleton in rat cells. Results showing a cytoskeleton-binding function for *A. nicotinovorans* MoeA

(Menéndez *et al.*, 1997) and the E-domain of the *A. thaliana* Cnx1 protein (Schwarz *et al.*, 1997) may suggest that the final step in MoCo biosynthesis takes place on an intracellular protein scaffold. Alternatively, an intracellular protein structure may be involved in the transport or storage of active molybdenum cofactor once molybdenum has been inserted into the MPT molecule. No binding of molybdenum to the G-domain of Cnx1 could be detected which was surprising given that *E. coli* *mogA* mutants are molybdate-repairable (Schwarz *et al.*, 1997). One suggested possibility for this finding is that sequential binding between MPT and molybdenum occurs where molybdenum binds only after the MPT-binding site is occupied. It is highly possible that the functions carried out by both domains of Cnx1 are conserved in CnxE however only future studies using the purified CnxE protein will determine whether this is indeed the case.

The ability of the CnxE cDNA clone pE1 to functionally complement the *E. coli* *moaB* locus could not be tested since no mutants defective in *moaB* have yet been described (Rivers *et al.*, 1993).

The DNA basis of 6 *cnxE* mutations (*cnxE12*, *cnxE16*, *cnxE849*, *cnxE14*, *cnxE3* and *cnxE13*) have been determined. In strain *cnxE12*, a single base pair substitution was identified which is predicted to result in a change from glycine to aspartate at position 344 (using the numbering system detailed for the wild-type *cnxE* sequence (see Figure 25)). This replaces an uncharged hydrophobic residue (glycine) with a negatively charged and polar substitute (aspartate). In strain *cnxE16*, a single base pair substitution was also identified, this time changing the inferred amino acid at position 523 from the small, polar and negatively charged residue aspartate to the larger, polar and negatively charged residue glutamate. Both glycine 344 and aspartate 523 are conserved among the Cinnamon, Gephyrin and MoeA proteins (Figure 25). The

disruption of CnxE activity by replacement of either residue combined with the conservation of these residues in similar proteins indicates that these amino acids play an important functional role. Highly conserved residues are often implicated in substrate binding, catalysis, dimer formation and other similar enzyme properties. For example in *A. nidulans* highly conserved cysteine residues within nitrate reductase have been suggested to function in the binding of both molybdenum cofactor and NADH, and in joining the enzyme's subunits (Hyde *et al.*, 1989). Strain *cnxE849* was found to contain a single base pair substitution ultimately converting alanine 373 to serine. Alanine 373 does not appear to be conserved among CnxE protein homologues (Figure 25) suggesting that this amino acid may not be essential for function, however the replacement of a hydrophobic moiety (alanine) with a polar substitute (serine) may result in alteration within the structural folding, solubility or other behavior of the protein so as to abolish CnxE activity.

The three further *cnxE* mutant strains analysed all contain mutations which are predicted to result in premature termination of translation: strain *cnxE14* has a change in the predicted amino acid sequence from phenylalanine to leucine at amino acid 429 and a shift in the reading frame downstream of amino acid 431 until early termination of translation at a stop codon reached at nucleotide position 1399 (where nucleotide 1399 represents the T of the TGA stop codon); in strain *cnxE3* the deletion of a single guanine residue was identified at nucleotide position 1594/amino acid position 502 resulting in a predicted reading frame shift from this point downstream until a premature stop codon is encountered at nucleotide 1618 (where nucleotide 1618 represents the T of the TAA stop codon) and sequence analysis of strain *cnxE13* revealed that 2 nucleotides, a cytosine and a thymine residue situated next to each other at positions 1715 and 1716 respectively had been deleted causing a

shift in the reading frame downstream of the mutation (from amino acid position 543 onwards) until early termination of translation at a stop codon, encountered at nucleotide 1979 (where nucleotide 1979 represents the T of the TGA stop codon).

Mutation in the *A. nidulans* *cnx* or *niaD* genes usually results in the constitutive synthesis of both nitrite reductase and of a mutant nitrate reductase (which can be detected because it still retains cytochrome *c* reductase activity), however some mutations in the *cnx* genes or the *niaD* gene result in 'normal' nitrate-inducible synthesis of both nitrite reductase and the mutant nitrate reductase (Cove, 1970; Cove and Pateman, 1969; Pateman *et al.*, 1964; Pateman *et al.*, 1967). Both the *cnxE12* and *cnxE16* mutants are inducible while the *cnxE3*, *cnxE13*, *cnxE14* and *cnxE849* mutants are constitutive (J.R. Kinghorn, unpublished). It has been suggested that the existence of these inducible and constitutive types of *cnx* mutants can be explained by the autoregulation exhibited by the nitrate reductase molecule. This autoregulation is thought to operate by a mechanism in which in the absence of nitrate, the functional nitrate reductase molecule is able to bind and inactivate the *nirA* gene product, thereby preventing stimulation of transcription of further nitrate reductase apoprotein. In the presence of nitrate, nitrate reductase would complex with it and no longer be able to inactivate the NirA protein (Cove, 1970). Thus, in addition to its catalytic function in nitrate assimilation, the nitrate reductase molecule has another role in the regulation of this pathway. It has been suggested that *niaD* and *cnx* mutants which drastically alter the conformation of nitrate reductase prevent *nirA* gene product inactivation and are therefore constitutive, while *niaD* and *cnx* mutants which affect only catalytic activity and not conformation remain normally regulated (are inducible). Both of the inducible mutants sequenced (*cnxE12* and *cnxE16*) have single base pair

changes in the *cnxE* gene altering single amino acid residues in the resulting protein. The constitutive mutant *cnxE849* also contains a single base pair change while the other constitutive mutants *cnxE3*, *cnxE13* and *cnxE14* possess mutations which result in early termination of CnxE protein translation. It could be suggested that the CnxE protein has a role in stabilising or influencing the conformation of the nitrate reductase molecule which is removed when the protein molecule is mutated in a certain way. Alternatively, some *cnx* mutations may still allow the production of a mutant MoCo molecule which is not functional but still allows the formation of a NR molecule with unaltered conformation.

All of the *cnxE* mutations analysed disrupt the E-domain of the predicted protein sequence between residues 344 and 710. Therefore in all of the six mutants tested the function of the G-domain remains viable. This result is perhaps unexpected given that the *E. coli* MogA protein function can be replaced by the *A. nidulans* *cnxE* gene product while the *E. coli* MoeA protein activity cannot. If each domain of CnxE has a different function it would seem likely that mutants defective for each activity should have been isolated during previous classical analysis in a manner analogous to the *cnxABC* locus. Early studies indicated that *cnxABC* comprised three tightly linked mutations (Cove and Pateman, 1963; Hartley, 1970). *cnxA* and *cnxC* mutants were found to complement each other, but not *cnxB* mutants, when grown as heterokaryons on nitrate (Cove and Pateman, 1963). Hartley (1970) described a 'complex' pattern of complementation within *cnxE* mutants, however other studies have not elaborated or corroborated this evidence. In the *A. thaliana* protein Cnx1 it has been suggested that the high affinity G-domain acts as the initial MPT-acceptor, with the low affinity E-domain taking over MPT binding when the concentration of MPT has reached a threshold level and the G-domain is becoming saturated. The fact that all the *cnxE*

mutant strains tested have mutations affecting only the E-domain of the CnxE protein may indicate that in *A. nidulans* the G-domain activity is dispensible, with its function being secondary to that of the E-domain. Site-directed mutagenesis of the DNA sequence encoding the G-domain might allow further investigation into the function of this part of the CnxE molecule during MoCo biosynthesis.

HPLC analysis of the *cnxE14* mutant strain showed that precursor Z was present at 1.6 times the wild-type level and molybdopterin was present at 2.3 times the wild-type level. Although there may appear to be a modest increase in precursor Z and molybdopterin levels within the mutant the increase is too slight to suggest an actual effect which could only be confirmed by analysing more samples of the *cnxE14* mutant and other *cnxE* mutant strains. Variations in the HPLC results may be produced by slight differences in cell physiology at the time of harvesting and by uneven sample homogenisation which is reflected in the margin of standard deviation. If CnxE does indeed function in the last step of MoCo biosynthesis it might be expected that precursor Z levels would remain unaffected and any significant accumulation in molybdopterin levels (due to disruption of the incorporation of molybdenum into the pterin molecules) may not be observed since in the wild-type strain both molybdopterin and molybdenum cofactor are converted to Form A dephospho.

The results for the *cnxE14* mutant strain contrast with that observed for the *E. coli mogA* mutant RK5206 in which Form A levels were significantly decreased in comparison to the wild-type (Joshi *et al.*, 1996). This decrease in MPT levels was explained by the suggestion that normally synthesised molybdopterin present in the *mogA* mutant may be degraded because further assembly of molybdenum cofactor and incorporation of MoCo into molybdoenzymes is interrupted, thereby only

leaving low levels of residual pterin in cells. It was also shown that the *mogA* mutation responsible for these lowered molybdopterin levels could not be overcome by growth on 1 mM molybdate in accord with the observed levels of phenotypic repair. This observation led to the suggestion that MogA functions as a molybdochelatase *i.e.* an enzyme which is responsible for the assembly of molybdenum cofactor from molybdopterin in the presence of low molybdate concentrations (Joshi *et al.*, 1996). It is unclear why the molybdopterin levels detected in the *E. coli mogA* and *A. nidulans cnxE* mutant strains should differ.

In conclusion it would appear that no significant alterations in either precursor Z or molybdopterin levels are present within the *cnxE14* mutant as compared to the wild-type strain, suggesting a putative role for CnxE in the final step of *A. nidulans* MoCo biosynthesis during the incorporation of molybdenum into MPT. This is derived from the fact that both molybdopterin and molybdenum cofactor are converted to Form A dephospho therefore if MPT is produced in the *cnxE14* mutant strain, instead of mature MoCo, it will be changed to an equivalent amount of Form A. It should however be noted that this result is inconclusive without further HPLC analysis of other *cnxE* mutant strains.

The hypothesis that CnxE is involved in such a 'final' MoCo biosynthetic step was suggested by early studies which looked at the sedimentation properties of nitrate reductase molecules found within *A. nidulans* wild-type and *cnx* mutant strains. In *A. nidulans* it was shown that wild-type nitrate reductase (which has NADPH: nitrate reductase activity) sedimented at 7.6 S, whereas certain *cnx* mutant nitrate reductase enzymes (which retain only cytochrome *c* reductase activity) sedimented at 4.5 S and other *cnx* mutant NR enzymes contained both the 7.6 S and the 4.5 S species (MacDonald *et al.*, 1974). *cnxE* mutants were unique in that they produced only the 7.6 S form of mutant NR. From these

observations it was postulated that the presence of MoCo allowed the subunits of the native enzyme to become aggregated. It was assumed that only 7.6 S cytochrome *c* reductase was detected in the *cnxE* mutants due to the production of molybdenum-free cofactor precursor ('empty' cofactor) in these strains which could still function to aggregate the NR monomers. This hypothesis did not explain why both 4.5 S and 7.6 S forms of cytochrome *c* reductase should be found in the other *cnx* strains. However a later study on xanthine dehydrogenase from *A. nidulans* (Lewis and Scazzocchio, 1977) suggested that the molybdeum cofactor serves to stabilise the dimeric enzyme structure but is not solely responsible for subunit aggregation therefore some 7.6 S dimerised NR could still be formed even in the absence of MoCo (or a Mo-free precursor) in the other *cnx* mutants. It is possible that cofactor binding results in slight conformational changes which in turn lead to strengthened subunit interactions, and that in *cnxE* mutants the molybdopterin which is formed still allows the same NR subunit stability. This theory is consistent with a role for CnxE in the incorporation of molybdenum into MPT molecules.

A single mRNA transcript of 2.5 kb for *cnxE* was revealed in Northern analysis. This suggests a *cnxE* gene coding region of 2.1 kb following an approximately 0.4 kb untranslated leader sequence. The expression of *cnxE* does not appear to be regulated by the presence of nitrate or ammonium, with constitutive expression of the gene after cell growth on either nitrogen source. (For more discussion on the regulation of MoCo biosynthesis in *A. nidulans* refer to section 7.4).

### 7.3 The *A. nidulans* *cnxABC* Locus

The *A. nidulans* *cnxABC* locus has recently been isolated (S.E. Unkles and J. Smith, personal communication) and found to consist of a single gene which is thought to encode a bifunctional protein with two catalytic domains. These two domains, termed CnxA and CnxC, are similar to *E. coli* MoaA and MoaC respectively. Results from this research show that strain *cnxA9* (which has a mutation in the CnxA domain), strain *cnxC3* (which has a mutation in the Cnx C domain) and strain *cnxB11* (which has a mutation affecting both domains) all contain low levels of precursor Z (approximately 3.2-5.9 times lower than the wild-type strain) and no detectable molybdopterin. The low precursor Z levels which were detected in the three mutant strains may actually represent another 'background' compound which is co-eluting with compound Z under the HPLC conditions employed. HPLC analysis has previously been used to show that an *E. coli* mutant strain containing the *moaA1* mutation is devoid of precursor Z (Johnson and Rajagopalan, 1987a; Johnson and Rajagopalan, 1987b).

The HPLC results, combined with the *E. coli* homologies, suggest that CnxABC is involved in the first stage of MoCo biosynthesis during the conversion of a phosphorylated guanosine residue to precursor Z. The exact mechanism of this conversion remains unclear (Wuebbens and Rajagopalan, 1995). The *cnxABC* locus encodes a single protein molecule which appears to carry out the combined functions of the *E. coli* MoaA and MoaC counterparts. An interesting future experiment might involve testing whether the *A. nidulans* *cnxABC* gene can functionally complement *E. coli* *moaA* and *moaC* mutations. In *A. thaliana*, the genes *cnx2* (which exhibits homology to MoaA) and *cnx3* (which exhibits homology to MoaC) were isolated by complementation of appropriate *E. coli* mutants (Hoff *et al.*, 1995). During the course of evolution the

*E. coli moaA* and *moaC* genes encoding closely related (and probably sequential) steps in the first stage of MoCo biosynthesis have become fused to produce a two-domain difunctional protein in *A. nidulans*. In this respect *cnxABC* appears to resemble the *cnxE* locus (section 7.2). Other examples of such multifunctional proteins have been found in *A. nidulans*, the AROM polypeptide, involved in the shikimate pathway, is an example of a pentafunctional enzyme (Lamb *et al.*, 1991) whose protein sequence shows similarity to five monofunctional shikimate pathway enzymes from *E. coli*. Genes carrying out related or sequential metabolic functions may become fused in order to channel substrates along the pathway, thus creating a higher flux than obtained by free diffusion of substrates.

#### 7.4 Regulation of MoCo Biosynthesis in *A. nidulans*

Little is known about the regulation of MoCo biosynthesis in eukaryotes. Expression of the *cnx* genes is required in three different metabolic contexts corresponding to growth on nitrate, purines or nicotinate as sole nitrogen source. The question remains whether *cnx* gene expression is induced through three independent pathway-specific systems or is constitutive. Early studies showed that levels of MoCo were not affected by mutations in the gene *nirA* (which is responsible for nitrate induction of the NiiA and NiaD enzymes), therefore suggesting that the *nirA* gene product is not involved in controlling the expression of the *cnx* genes (Garret and Cove, 1976). It was also demonstrated that cofactor levels were lower in the *niaD26* mutant strain grown with ammonium as nitrogen source, whether or not nitrate was also present indicating that the *cnx* genes are perhaps regulated by ammonium repression (Garret and Cove, 1979).

HPLC analysis of the wild-type *A. nidulans* strain revealed that precursor Z levels were approximately 3.5 times higher in nitrate grown

cells than in ammonium grown cells and approximately 1.7 times higher in adenine grown cells than in ammonium grown cells. The wild-type cells grown with nitrate plus proline contained approximately 0.6 times less precursor Z than wild-type cells grown with nitrate alone, but still contained approximately 2.1 and 1.2 times more precursor Z than the ammonium and adenine grown cells respectively. This suggests a slight increase in precursor Z formation in the presence of nitrate, however cells grown with nitrate plus proline as sole nitrogen source exhibit a much more modest increase in precursor Z levels. Growth on adenine also resulted in a slight increase in precursor Z levels but this was not as large as that observed for growth on nitrate. As nitrate and adenine serve to induce the molybdoenzymes nitrate reductase and xanthine dehydrogenase (purine hydroxylases I and II) respectively, growth on each of these nitrogen sources would be expected to result in an equivalent increase in precursor Z production.

The precursor Z levels do not correlate well with the observed cellular levels of molybdopterin grown with the various different nitrogen sources. Wild-type cells grown with nitrate plus proline contained the lowest amount of molybdopterin, with cells grown on ammonium, adenine and nitrate alone containing approximately 1.3, 1.4 and 1.7 times more molybdopterin respectively. Although growth on nitrate alone produced the largest increase in molybdopterin levels the increase was not as great as observed with precursor Z, and nitrate plus proline contained the lowest level of molybdopterin in contrast to that seen for precursor Z.

Molybdopterin levels within the wild-type strain do not appear to be significantly altered by growth on the different sole nitrogen sources. Although there may appear to be an increase in precursor Z levels with growth on nitrate (and to a lesser extent with growth on

adenine) it is difficult to tell if this is a significant effect. Variations in the HPLC results can be produced by slight differences in cell physiology at the time of harvesting and by uneven sample homogenisation which is reflected in the margin of standard deviation.

Northern analysis showed that the *cnxABC* transcript is expressed to a greater extent in wild-type cells grown with nitrate than with ammonium as sole nitrogen source (S.E. Unkles, personal communication). This is in contrast to the transcript abundance levels recorded for the *cnxE* and *cnxG* genes (and the *cnxF* and *cnxH* genes, (V.A. Appleyard, personal communication)) grown under the same conditions. It can be concluded that *cnx* gene expression occurs more or less constitutively in the presence of nitrate, adenine and ammonium but a modest increase on nitrate of molybdenum cofactor biosynthesis exists which is exerted only on the early steps of the pathway. This hypothesis correlates with the slight increase in precursor Z levels observed under growth on nitrate as the sole nitrogen source. It has been tentatively suggested that the *cnxJ* locus might be involved in the regulation of molybdenum cofactor expression in *A. nidulans* (Arst *et al.*, 1982). *cnxJ* mutations differ from other types of *cnx* mutation in that they effectively lower but do not abolish any or all of the three molybdoenzyme activities found in *A. nidulans*. Future cloning of the *cnxJ* locus will allow further characterisation of the role undertaken by its encoded protein product.

Analysis of RNA transcripts produced from the *A. thaliana* *cnx2* and *cnx3* genes (which are thought to catalyse steps equivalent to the *A. nidulans* *cnxABC* locus) has shown that there is no regulation of expression by nitrate (Hoff *et al.*, 1995). Contrastingly in *A. thaliana* the synthesis of Cnx1 (which is homologous to *cnxE*) was increased by growth on nitrate thus showing a co-regulation with nitrate reductase (Mendel, 1997). This data conflicts with that observed for *A. nidulans* in

which any nitrate-induction of MoCo biosynthesis seems to be exerted only on the first steps in the pathway, not during the later steps as would appear to be the case for *A. thaliana*.

In the prokaryotic system *E. coli*, expression of the *moa* locus appears to be induced to high levels under anaerobic conditions and is subject to repression by the molybdenum cofactor (Baker and Boxer, 1991). On the basis of this finding it has been suggested that the overall control of MoCo biosynthesis in *E. coli* is possibly mediated through feedback regulation of the *moa* operon by the end product of the pathway (active MoCo). In a separate study, the expression of the *moa* operon and the *moeA* gene appeared to be little affected by the nitrogen source present in the culture medium (Rubio *et al.*, 1998). The expression of the *E. coli* molybdate transport operon, *mod* is increased under conditions of molybdate limitation, however is not affected by oxygen or nitrate (Rech *et al.*, 1995) and the expression of the *mob* locus is apparently constitutive (Iobbi-Nivol *et al.*, 1995).

## **7.5**      **An Overview of Molybdenum Cofactor Biosynthesis in** ***A. nidulans***

### **7.5.1**    **The Proposed Function of Other *cnx* Genes**

As mentioned previously the *cnxABC* locus has been isolated and results from this research have helped to provide important information on the functional role of the encoded protein. During the course of this work, the *A. nidulans* *cnxF* and *cnxH* genes have also been cloned and characterised (V.A. Appleyard and S.E. Unkles, personal communication).

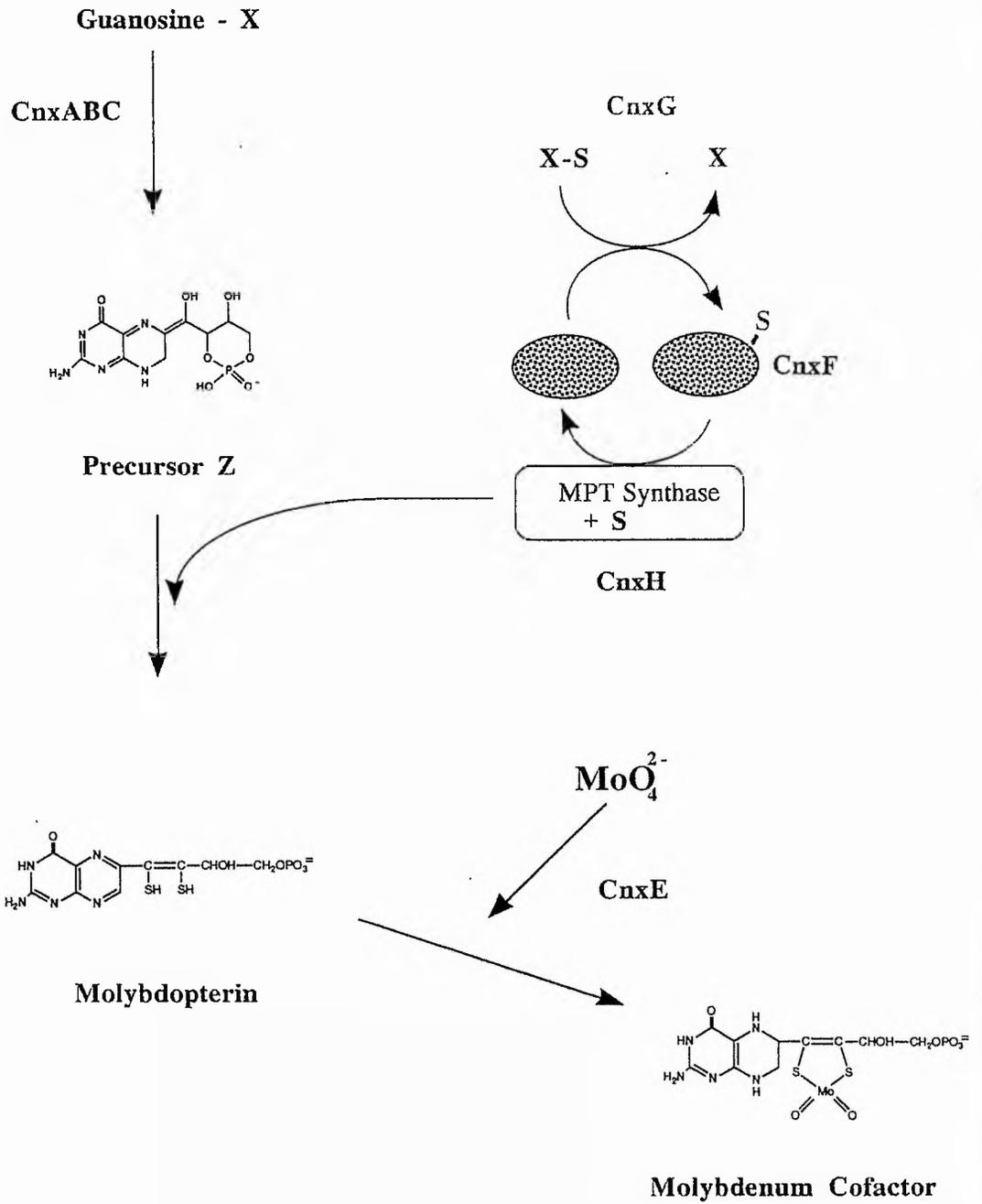
The *cnxF* gene exhibits homology to the *E. coli* molybdenum cofactor biosynthesis protein MoeB. In *E. coli*, MoeB is involved in the transfer of sulphur to the small subunit of MPT synthase (Pitterle and

Rajagopalan, 1993; Pitterle *et al.*, 1993). From this homology and HPLC analysis of *cnxF* mutants, which show significant accumulation of precursor Z and abolition of molybdopterin levels, it has been suggested that CnxF may function in a similar manner to *E. coli* MoeB *i.e.* in the activation of the small subunit of MPT synthase by the donation of sulphur to this molecule. *cnxH* is similar to the *E. coli* protein MoeA, which forms the large subunit of MPT Synthase (Pitterle and Rajagopalan, 1993). It therefore seems likely that in *A. nidulans*, CnxH may form the large subunit of the converting factor, with this hypothesis being strengthened by HPLC data on precursor Z and molybdopterin levels found within *cnxH* mutant strains.

An overview of MoCo biosynthesis in *A. nidulans* is presented in Figure 39. The model is based on that suggested by Stallmeyer *et al.*, (1995) to illustrate MoCo biosynthesis in *E. coli*. The bifunctional protein CnxABC is responsible for the conversion of a phosphorylated guanosine residue into precursor Z by an unknown mechanism. Precursor Z is then converted to molybdopterin by the transfer of sulphur from a converting factor which may possess one or two subunits. It would appear that at least one converting factor subunit is present in the form of CnxH, which is similar to the large subunit protein in *E. coli*. The question of whether *A. nidulans* contains a gene encoding a small subunit equivalent to *E. coli* MoeD is addressed in section 7.5.2. The CnxF protein functions to transfer sulphur, which is initially donated by CnxG, to the converting factor. The final step in *A. nidulans* MoCo biosynthesis, the incorporation of molybdenum into molybdopterin, is most likely carried out by CnxE. As mentioned previously, CnxE consists of 2 domains, E and G, which are similar to the *E. coli* MoeA and MogA proteins respectively. This may indicate a dual function for the CnxE protein however only future functional characterisation will allow confirmation of this suggestion.

Figure 39: Proposed overview of MoCo biosynthesis in *A. nidulans*

The putative bifunctional protein CnxABC is responsible for the conversion of a phosphorylated guanosine residue into precursor Z. Precursor Z is then converted to molybdopterin by the transfer of sulphur from the enzyme MPT synthase (also known as the converting factor), which may possess one or two subunits. In *A. nidulans*, it would appear that at least one MPT synthase subunit is present in the form of CnxH, which is similar to the large subunit protein in *E. coli*. CnxF functions to transfer sulphur, which is initially donated by CnxG, to the MPT synthase. The final step in *A. nidulans* MoCo biosynthesis, the incorporation of molybdenum into molybdopterin, is carried out by CnxE.



In *E. coli*, the *mod* operon is responsible for the uptake of molybdenum into the cell. No equivalent to *E. coli mod* has been isolated in *A. nidulans*. Very little is known about the way in which molybdate is transported into the *A. nidulans* cell but the existence of an uptake system which is common to the five structurally similar anions molybdate, sulphate, selenate, chromate and tungstate has been suggested (Arst, 1968).

A MoCo carrier protein may be responsible for distributing the molybdenum cofactor within the cell for incorporation into nitrate reductase and purine hydroxylases I and II. It is unknown whether such a carrier protein exists in *A. nidulans* but a putative MoCo carrier protein has been identified in several organisms including *Escherichia coli* (Amy and Rajagopalan, 1979) *Chlamydomonas reinhardtii* (Aguilar *et al.*, 1991) and *Vicia faba* (Kalakoutskii and Fernández, 1996).

Synthesised MoCo binds to the molybdenum cofactor domain of the nitrate reductase monomers. Several studies on NR defective mutants in *A. nidulans* have pinpointed single amino acid residues within the MoCo domain of the NR protein which are thought to have a critical role in MoCo domain functionality (Wilkinson and Crawford, 1993; Braaksma and Feenstra, 1982; LaBrie and Crawford, 1994; Meyer *et al.*, 1995; Garde *et al.*, 1995). The mutation of one residue, namely cysteine 150, has been found to abolish NR activity as well as MoCo domain partial activities (measured using the non-physiological electron donors MVH and BPB) and it has been suggested that this residue may be involved in the attachment of MoCo to the NR apoprotein or may be involved in subunit dimerisation (Garde *et al.*, 1995). The importance of this residue is highlighted by its conservation not only between nitrate reductase enzymes from different species but also between different molybdoenzymes.

### 7.5.2 Does a *moaD* Gene Equivalent Exist in *A. nidulans* ?

With the exception of *cnxJ*, all of the *A. nidulans* *cnx* loci which were originally identified by classical genetic studies have been cloned and sequenced. It has become apparent that none of the isolated *cnx* genes encodes a protein with similarity to the *E. coli* *moaD* locus. *cnxJ* is thought unlikely to encode the MPT synthase small subunit protein because the *cnxJ* gene product appears to be dispensible for cofactor synthesis with mutations within this locus only lowering, not abolishing, molybdoenzyme levels (Arst *et al.*, 1982). Although no *A. nidulans* gene with similarity to the small subunit of the converting factor in *E. coli* has been encountered, *cnxH* exhibits similarity to the large subunit of the converting factor which is encoded by *moaE* in *E. coli*. In *E. coli*, the small subunit of MPT synthase (encoded by *moaD*) must be activated by the addition of a single reactive sulphur before the whole converting factor can function (Pitterle *et al.*, 1993).

No other genes have been observed to exhibit the *cnx* gene phenotype. If a *moaD* equivalent does exist within the *A. nidulans* genome it may be possible that the gene is held as two (or more) copies meaning that both (or all) copies would need to be disrupted to reveal a mutant phenotype. Multiple copies of genes have been described in a number of organisms: multiple copies of the bone specific osteocalcin (OC) gene have been reported in mouse and rat (Rahman *et al.*, 1993) and in *Anabaena flos-aquae* seven copies of the gene *gvpA*, which encodes a protein component of cyanobacterial gas vesicles, were recorded (Hayes and Powell, 1995). Another possibility is that any mutation within such a gene is lethal to the cell, but this seems unlikely given that the other *cnx* mutants do not exhibit this phenomenon.

The experiments carried out in Chapter 6 of this study failed to isolate a gene capable of complementing an *E. coli* *moaD* mutant. This

may have been due to the absence of an appropriate clone from the portion of the *A. nidulans* cDNA library which was tested, or may be due to the inability of the *E. coli* mutant to functionally express an appropriate *A. nidulans* cDNA clone. The colonies resulting from complementation of the *E. coli moaD* mutant strain MJ431 with the positive control plasmid pEM345 were small (0.5 mm in diameter) and translucent in appearance. The same small, translucent colonies were observed during complementation of the *E. coli moeA* mutant strain NS9 with the positive control plasmid pJR3. The reason for this observation is unclear.

As mentioned before, the accumulation of precursor Z and the abolition of molybdopterin levels recorded in the *cnxG* mutants suggest that the *cnxG* gene product is involved in the transfer of sulphur at some stage in the MoCo biosynthetic pathway. It could be suggested that instead of a role in the initial donation of sulphur to the MPT synthase sulphotransferase (as presented in the overview detailed in Figure 39), *cnxG* carries out the equivalent function of *moaD* in *A. nidulans* (*i.e.* the donation of sulphur to precursor Z). To test this hypothesis any or all of the 6 *cnxG* cDNA clones which have been isolated in this research could be tested for their ability to complement an appropriate *E. coli moaD* mutant strain by selecting for the reconstitution of nitrate reductase activity during anaerobic growth on LN medium. In *Arabidopsis thaliana*, an *E. coli moaD* homologue has not yet been identified despite most of the other plant MoCo biosynthetic genes having already been cloned (Mendel, 1997). A human *E. coli moaD* equivalent has however recently been isolated (Mendel, 1997) suggesting conservation of the *moaD* gene through evolution to higher eukaryotes.

It will be interesting to see if the *A. nidulans* genome sequencing project which is currently being undertaken by a number of

laboratories will reveal any gene sequence(s) with homology to *E. coli moaD*.

## **7.6      Concluding Remarks**

In accordance with the aims of this study the *A. nidulans* molybdenum cofactor biosynthesis genes *cnxG* and *cnxE* have been isolated and characterised. The homologies exhibited by these genes, as well as DNA sequencing and HPLC analysis of *cnxG* and *cnxE* mutant strains, has allowed putative functional roles for the proteins encoded by these loci to be suggested. HPLC studies on *cnxA*, *cnxB* and *cnxC* mutant strains have also provided valuable information pertaining to the function of the CnxABC protein.

Future work could extend these findings to provide further details on the specific functions carried out by the proteins encoded by the *cnxG* and *cnxE* loci within the *A. nidulans* MoCo biosynthetic pathway. The *cnxG* protein overexpression plasmid pLIND-G constructed during this study should serve as a basis for such functionality studies. Overexpression and characterisation of the CnxE protein could also be a focus for future work. In the long term, X-ray crystallography may provide information on the structural features of the CnxG and CnxE proteins.

The isolation and characterisation of the *cnx* genes contributes to a general understanding of this important pathway within the model filamentous fungus *A. nidulans* and gives an indication of how this pathway resembles the mechanisms of MoCo biosynthesis in other species. The elucidation of the *A. nidulans* *cnxG* and *cnxE* sequences may allow the isolation of clones encoding protein counterparts in other species by a number of methods including the construction of degenerate primers or by heterologous hybridisation. The identification of human MoCo

biosynthesis genes may be particularly rewarding, perhaps helping us to understand the mechanisms underlying MoCo deficiency, and in the long term providing treatment for individuals affected by this devastating genetic condition. A recent review by Mendel (1997) suggests that several human MoCo biosynthesis genes have already been isolated (no data given).

The cosmid locations of the *cnxE* and *cnxG* genes derived from this study should help in the construction of the *A. nidulans* contig map which is currently underway using the chromosome-specific subcollections identified by Brody and colleagues (1991). The development of a such a contig map should greatly facilitate position-based cloning by chromosome walking from a previously cloned gene, with reference to the classical genetic map.

No other examples of fungal or yeast MoCo biosynthesis genes have so far been isolated. In fact, the cloning and preliminary functional characterisation of *cnxG* and *cnxE*, as well as the *cnxABC*, *cnxF* and *cnxH* loci means that *A. nidulans* has, to date, one of the most extensively investigated MoCo biosynthesis pathways, second only to *E. coli*. The genes which were first identified by classical studies on mutant strains more than 3 decades ago, and whose discovery led to the first proposal that a molybdenum cofactor common to several molybdoenzymes might exist, have now been isolated and characterised using recombinant DNA technology. Future work will build on these findings and functional studies may eventually lead to a greater understanding of the *A. nidulans* MoCo biosynthesis pathway.

## **BIBLIOGRAPHY**

- Aberg, B. (1947) On the mechanism of the toxic action of chlorates and some related substances upon young wheat plants. *Kungliga Lantbrukshogskolans Annaler* **15**, 37-107.
- Aguilar, M.R., Cárdenas, J. and Fernández, E. (1991) Regulation of molybdenum cofactor species in the green alga *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta* **1073**, 463-469.
- Aguilar, M.R., Kalakoutskii, K.L., Cárdenas, J. and Fernández, E. (1992) Direct transfer of molybdopterin cofactor to aponitrate reductase from a carrier protein in *Chlamydomonas reinhardtii*. *FEBS Letters* **307**, 162-163.
- Alderson, Y. and Scazzocchio, C. (1967) A system for the study of interlocus specifying for both forward and reverse mutation in at least eight gene loci in *Aspergillus nidulans*. *Mutation Research* **8**, 255-264.
- Aleksenko, A.Y. (1994) Cointegration of transforming DNAs in *Aspergillus nidulans*: a model using autonomously-replicating plasmids. *Current Genetics* **26**, 352-358.
- Aleksenko, A.Y. and Clutterbuck, A.J. (1997) Autonomous plasmid replication in *Aspergillus nidulans*: AMA1 and MATE elements. *Fungal Genetics and Biology* **21**, 373-387.

- Aleksenko, A.Y., Gems, D. and Clutterbuck, A.J. (1996) Multiple copies of MATE elements support autonomous plasmid replication in *Aspergillus nidulans*. *Molecular Microbiology* **20**, 427-434.
- Allen, R.M., Chatterjee, R., Madden, M.S., Ludden, P.W. and Shah, V.K. (1994) Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Critical Reviews Biotechnology* **14**, 225-249.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.
- Amy, N.K. and Rajagopalan, K.V. (1979) Characterisation of molybdenum cofactor from *Escherichia coli*. *Journal of Bacteriology* **140**, 114-124.
- Amy, C.M., Williams-Ahlf, B., Naggert, J. and Smith, S. (1992) Intron-exon organisation of the gene for the multifunctional animal fatty acid synthase. *Proceedings of the National Academy of Sciences (USA)* **89**, 1105-1108.
- Arnold, G.L., Greene, C.L., Stout, J.P. and Goodman, S.I. (1993) Molybdenum cofactor deficiency. *Journal of Pediatrics* **123**, 595-598.
- Arst, H.N. Jr. (1968) Genetic analysis of the first steps of sulphate metabolism in *Aspergillus nidulans*. *Nature* **219**, 268-270.

- Arst, H.N. Jr. and MacDonald, D.W. (1978) Reduced expression of a distal gene of the *prn* gene cluster in deletion mutants of *Aspergillus nidulans*: genetic evidence for a dicistronic messenger in a eukaryote. *Molecular and General Genetics* **163**, 7-22.
- Arst, H.N. Jr., MacDonald, D.W. and Cove, D.J. (1970) Molybdate metabolism in *Aspergillus nidulans*. I Mutations affecting nitrate reductase and xanthine dehydrogenase. *Molecular and General Genetics* **108**, 129-145.
- Arst, H.N. Jr., Tollervey, D.W. and Sealy-Lewis, H.M. (1982) A possible regulatory gene for the molybdenum-containing cofactor in *Aspergillus nidulans*. *Journal of General Microbiology* **128**, 1083-1093.
- Aviv, H. and Leder, P. (1972) Purification of biologically active globin messenger RNA by chromatography on oligo-thymidylic acid-cellulose. *Proceedings of the National Academy of Sciences (USA)* **69**, 1408-1412.
- Bachmann, B.J. (1983) Linkage map of *Escherichia coli* K-12, edition 7. *Microbiological Reviews* **47**, 180-230.
- Bagchi, S.N. and Singh, H.N. (1984) Genetic control of nitrate reduction in the cyanobacterium *Nostoc muscorum*. *Molecular and General Genetics* **193**, 82-84.

- Baker, K.P. and Boxer, D.H. (1991) Regulation of the *chlA* locus of *Escherichia coli* K12: involvement of molybdenum cofactor. *Molecular Microbiology* **5**, 901-907.
- Balance, D.J., Buxton, F.P. and Turner, G. (1983) Transformation of *Aspergillus nidulans* by the orotidine-5'-phosphate decarboxylase gene of *Neurospora crassa*. *Biochemical and Biophysical Research Communications* **112**, 284-289.
- Bamforth, F.J., Johnson, J.L., Davidson, A.G.F., Wong, L.T.K., Lockitch, G. and Applegarth, D.A. (1990) Biochemical investigation of a child with molybdenum cofactor deficiency. *Clinical Biochemistry* **23**, 537-542.
- Barton, A.B., Kaback, D.B., Clark, M.W., Keng, T., Ouellette, B.F.F., Storms, R.K., Zeng, B., Zhong, W.W., Fortin, N., Delaney, S. and Bussey, H. (1993) Physical localisation of yeast *CYS3*, a gene whose product resembles the rat gamma-cystathionase and *Escherichia coli* cystathionine gamma-synthase enzymes. *Yeast* **9**, 363-369.
- Bej, A.K. and Perlin, M.H. (1989) A high efficiency transformation system for the basidiomycete *Ustilago violaceae* employing hygromycin resistance and lithium-acetate treatment. *Gene* **80**, 171-176.
- Belfaiza, J., Parsot, C., Martel, A., de la Tour, C.B., Margarita, D., Cohen, G.N. and Saint-Girons, I. (1986) Evolution in biosynthetic pathways: two enzymes catalysing consecutive steps in methionine

biosynthesis originate from a common ancestor and possess a similar regulatory region. *Proceedings of the National Academy of Sciences (USA)* **83**, 867-871.

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

Bodie, E.A., Bower, B., Berka, R.M. and Dunn-Coleman, N.S. (1994) Economically important organic acid and enzyme products. In *Aspergillus: 50 Years On* (Edited by J.R. Kinghorn and S.D. Martinelli), pp. 561-602. Elsevier, Amsterdam.

Boles, R.G., Ment, L.R., Meyn, M.S., Horwich, A.L., Kratz, L.E. and Rinaldo, P. (1993) Short-term response to dietary therapy in molybdenum cofactor deficiency. *Annals of Neurology* **34**, 742-744.

Bonioli, E., DiStefano, A., Palmieri, A., Bertola, A., Bellini, C., Caruso, U., Fantasia, A.R., Minniti, G. and Dorche, C. (1996) Combined deficiency of xanthine oxidase and sulphite oxidase due to a deficiency of molybdenum cofactor. *Journal of Inherited Metabolic Disease* **19**, 700-701.

Börner, G., Karrasch, M. and Thauer, R.K. (1991) Molybdopterin adenine dinucleotide and molybdopterin hypoxanthine dinucleotide in formylmethanofuran dehydrogenase from *Methanobacterium thermoautotrophicum* (Marburg). *FEBS Letters* **290**, 31-34.

- Braaksma, F.J. and Feenstra, W.J. (1973) Isolation and characterization of chlorate resistant mutants of *Arabidopsis thaliana*. *Mutation Research* **19**, 175-185.
- Bright, S.W.J., Norbury, P.B., Franklin, J., Kirk, D.W. and Wray, J.L. (1983) A conditional-lethal *cnx*-type nitrate reductase-deficient barley mutant. *Molecular and General Genetics* **189**, 240-244.
- Brody, H. and Carbon, J. (1989) Electrophoretic karyotype of *Aspergillus nidulans*. *Proceedings of the National Academy of Sciences (USA)* **86**, 6260-6263.
- Brody, H., Griffith, J., Cuticchia, A.J., Arnold, J. and Timberlake, W.E. (1991) Chromosome-specific recombinant DNA libraries from the fungus *Aspergillus nidulans*. *Nucleic Acids Research* **19**, 3105-3109.
- Brown, E.A., D'Ari, R. and Newman, E.B. (1990) A relationship between L-serine degradation and methionine biosynthesis in *Escherichia coli* K12. *Journal of General Microbiology* **136**, 1017-1023.
- Brownlee, A.G. and Arst, H.N. Jr. (1983) Nitrate uptake in *Aspergillus nidulans* and the involvement of the third gene of the nitrate assimilatory gene cluster. *Journal of Bacteriology* **115**, 1138-1146.

Buchanan, R.J. and Wray, J.L. (1982) Isolation of molybdenum cofactor defective cell lines of *Nicotina tabacum*. *Molecular and General Genetics* **188**, 228-234.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., Fitzgerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J-F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodeck, A., Scott, J.L., Geoghagen, N.S.M., Weidman, J.F., Fuhrmann, J.L., Nguyen, D., Utterback, T.R., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Hurst, M.A., Kaine, B.P., Borodovsky, M., Klenk, H-P., Fraser, C.M., Smith, H.O., Woese, C.R. and Venter, J.C. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058-1073.

Campbell, W.H. and Kinghorn, J.R. (1990) Functional domains of assimilatory nitrate reductases and nitrite reductases. *Trends in Biochemical Sciences* **15**, 315-319.

Casadaban, M.J. and Cohen, S.N. (1979) Lactose genes fused to exogenous promoters in one step using a *Mu-lac* bacteriophage: *in vitro* probe for transcriptional control sequences. *Proceedings of the National Academy of Sciences (USA)* **76**, 4530-4533.

Cathala, G., Savouret, J-F., Mendez, B., West, B.L., Karin, M., Martial, J.A. and Baxter, J.D. (1983) A method for isolation of intact, translationally active ribonucleic acid. *DNA* **2**, 329-335.

- Cherest, H., Talbot, G. and de Robichon-Szulmajster, H. (1970) Role of homocysteine synthetase in an alternate route for methionine biosynthesis in *Saccharomyces cerevisiae*. *Journal of Bacteriology* **102**, 448-461.
- Cherest, H., Thomas, D. and Surdin-Kerjan, Y. (1993) Cysteine biosynthesis in *Saccharomyces cerevisiae* occurs through the transsulfuration pathway which has been built up by enzyme recruitment. *Journal of Bacteriology* **175**, 5366-5374.
- Clausen, T., Huber, R., Laber, B., Pohlenz, H-D. and Messerschmidt, A. (1996) Crystal structure of the pyridoxal-5'-phosphate dependent cystathionine beta-lyase from *Escherichia coli* at 1.83Å. *Journal of Molecular Biology* **262**, 202-224.
- Clutterbuck, A.J. (1994) Linkage map of *Aspergillus nidulans*. In *Aspergillus: 50 Years On* (Edited by S.D. Martinelli and J.R. Kinghorn), pp. 792-793. Elsevier, Amsterdam.
- Clutterbuck, A.J. (1997) The validity of the *Aspergillus nidulans* linkage map. *Fungal Genetics and Biology* **21**, 267-277.
- Coglievina, M., Klima, R., Bertani, I., Delneri, D., Zaccaria, P. and Bruschi, C.V. (1997) Sequencing of a 40.5 kb fragment located on the left arm of chromosome VII from *Saccharomyces cerevisiae*. *Yeast* **13**, 55-64.
- Cohen, S.H., Chang, A.C.Y. and Hsu, L. (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of

*Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Sciences (USA)* **69**, 2110-2114.

Contreras, R., Carrez, D., Kinghorn, J.R., van den Hondel, C.A.M.J.J. and Fiers, W. (1991) Efficient KEX2-like processing of a glucoamylase-interleukin-6 fusion protein by *Aspergillus nidulans* and secretion of mature interleukin-6. *Biotechnology* **9**, 378-381.

Cooley, R.N. and Tomsett, A.B. (1985) Determination of the subunit size of NADPH-nitrate reductase from *Aspergillus nidulans*. *Biochimica et Biophysica Acta* **831**, 89-93.

Cove, D.J. (1966) The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. *Biochimica et Biophysica Acta* **113**, 51-56.

Cove, D.J. (1970) Control of gene action in *Aspergillus nidulans*. *Proceedings of the Royal Society, London, Series B* **176**, 267-275.

Cove, D.J. (1976a) Chlorate toxicity in *Aspergillus nidulans*: the selection and characterisation of chlorate resistant mutants. *Heredity (Edinburgh)* **36**, 191-203.

Cove, D.J. (1976b) Chlorate toxicity in *Aspergillus nidulans*. Studies of mutants altered in nitrate assimilation. *Molecular and General Genetics* **146**, 147-159.

- Cove, D.J. (1979) Genetic studies of nitrate assimilation in *Aspergillus nidulans*. *Biological Reviews* **54**, 291-327.
- Cove, D.J. and Pateman, J.A. (1963) Independantly segregating loci concerned with nitrate reductase activity in *Aspergillus nidulans*. *Nature* **198**, 262-263.
- Cove, D.J. and Pateman, J.A. (1969) Autoregulation of the synthesis of nitrate reductase in *Aspergillus nidulans*. *Journal of Bacteriology* **97**, 1374-1378.
- Cramer, S.P., Wahl, R. and Rajagopalan, K.V. (1981) Molybdenum sites of sulfite oxidase and xanthine dehydrogenase. A comparison by EXAFS. *Journal of the American Chemical Society* **103**, 7721-7727.
- Crawford, N.M. (1995) Nitrate: nutrient and signal for plant growth. *The Plant Cell* **7**, 859-868.
- Crawford, N.M., Smith, M., Bellissimo, D. and Davis, R.W. (1988) Sequence and nitrate regulation of the *Arabidopsis thaliana* mRNA encoding nitrate reductase, a metalloflavoprotein with three functional domains. *Proceedings of the National Academy of Sciences (USA)* **85**, 5006-5010.
- Darlington, A.J. and Scazzocchio, C. (1967) Use of analogues and the substrate sensitivity of mutants in analysis of purine uptake and breakdown in *Aspergillus nidulans*. *Journal of Bacteriology* **93**, 937-940.

- Darlington, A.J. and Scazzocchio, C. (1968) Evidence for an alternative pathway of xanthine oxidation in *Aspergillus nidulans*. *Biochimica et Biophysica Acta* **166**, 569-571.
- Darlington, A.J., Scazzocchio, C. and Pateman, J.A. (1965) Biochemical and genetical studies of purine degradation in *Aspergillus*. *Nature* **206**, 599-600.
- Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) *Basic Methods in Molecular Biology*. Elsevier, New York.
- del Campillo-Campbell, A. and Campbell, A. (1982) Molybdenum cofactor requirement for biotin sulfoxide reduction in *Escherichia coli*. *Journal of Bacteriology* **149**, 469-478.
- Desjacques, P., Mousson, B., Vianey-Liaud, C., Boulieu, R., Bory, C., Baltassat, P., Divry, P., Zobot, H.T., Cotte, J. and Lagier, P. (1985) Combined deficiency of xanthine oxidase and sulphite oxidase: diagnosis of a new case followed by an antenatal diagnosis. *Journal of Inherited Metabolic Disease* **8**, 117-118.
- Dhawale, S.S., Paietta, J.V. and Marzluf, G.A. (1984) A new, rapid and efficient transformation procedure for *Neurospora*. *Current Genetics* **8**, 77-79.
- Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Research* **16**, 6127-6145.

- Downey, R.J. and Focht, W.J. (1974) Subunit character of the NADPH-nitrate reductase from *Aspergillus nidulans*. *Microbios* **11A**, 61-70.
- Downey, R.J. and Steiner, F.X. (1979) Further characterisation of the reduced NADP:nitrate oxidoreductase in *Aspergillus nidulans*. *Journal of Bacteriology* **137**, 105-114.
- Duchange, N., Zakin, M.M., Ferrara, P., Saint-Girons, I., Park, I., Tran, S.V., Py, M-C. and Cohen, G.N. (1983) Structure of the *metJBLF* cluster in *Escherichia coli* K12. Sequence of the *metB* structural gene and of the 5'- and 3'- flanking regions of the *metBL* operon. *Journal of Biological Chemistry* **258**, 14868-14871.
- Duesterhoeft, A., Erdmann, D. and Kroeger, M. (1991) Stepwise cloning and molecular characterisation of the *HgiDI* restriction-modification system from *Herpetosiphon giganteus* Hpa2. *Nucleic Acids Research* **19**, 1049-1056.
- Duran, M., Beemer, F., van der Heiden, C., Korteland, P.K., deBree, P.K., Brink, M. and Wadman, S.K. (1978) Combined deficiency of xanthine oxidase and sulphite oxidase: a defect of molybdenum metabolism or transport? *Journal of Inherited Metabolic Disease* **1**, 175-178.
- Eaves, D.J., Palmer, T. and Boxer, D.H. (1997) The product of the molybdenum cofactor gene *mobB* of *Escherichia coli* is a GTP-binding protein. *European Journal of Biochemistry* **246**, 690-697.

- Eki, T., Naitou, M., Hagiwara, H., Ozawa, M., Sasanuma, S-I., Sasanuma, M., Tsuchiya, Y., Shibata, T., Hanaoka, F. and Murakami, Y. (1996) Analysis of a 36.2 kb DNA sequence including the right telomere of chromosome VI from *Saccharomyces cerevisiae*. *Yeast* **12**, 149-167.
- Erickson, P.F., Maxwell, I.H., Su, L.J., Baumann, M. and Glode, L.M. (1990) Sequence of cDNA for rat cystathionine gamma-lyase and comparison of deduced amino acid sequence with related *Escherichia coli* enzymes. *Biochemistry Journal* **269**, 335-340.
- Fasano, O., Aldrich, T., Tamanoi, F., Taparowsky, E., Furth, M. and Wigler, M. (1984) Analysis of the transforming potential of the human H-ras gene by random mutagenesis. *Proceedings of the National Academy of Sciences (USA)* **81**, 4008-4012.
- Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6-13.
- Fidel, S., Doonan, J.H. and Morris, N.R. (1988) *Aspergillus nidulans* contains a single actin gene which has unique intron locations and encodes a  $\gamma$ -actin. *Gene* **70**, 283-293.
- Finch, J.T. (1978) Sickle-cell haemoglobin fibres. *Nature* **272**, 496-497.
- Fincham, J.R.S. (1989) Transformation in fungi. *Microbiology Reviews* **53**, 148-170.

- 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., McKenny, K., Sutton, G., Fitzhugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L-I., Glodek, A., Kelley, J.M., Weidmann, 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. and Venter, J.C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.
- Fogolino, M., Borne, F., Bally, M., Ball, G. and Patte, J.C. (1995) A direct sulfhydrylation pathway is used for methionine biosynthesis in *Pseudomonas aeruginosa*. *Microbiology* **141**, 431-439.
- Garde, J., Kinghorn, J.R. and Tomsett, A.B. (1995) Site-directed mutagenesis of nitrate reductase from *Aspergillus nidulans*. Identification of some essential and non-essential amino acids among conserved residues. *The Journal of Biological Chemistry* **270**, 6644-6650.
- Garrett, R.H. and Cove, D.J. (1976) Formation of NADPH-nitrate reductase activity *in vitro* from *Aspergillus nidulans* *niaD* and *cnx* mutants. *Molecular and General Genetics* **149**, 179-186.
- Gems, D., Johnstone, I.L. and Clutterbuck, J.A. (1991) An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene* **98**, 61-67.

- Gems, D., Aleksenko, A., Belenky, L., Robertson, S., Ramsden, M., Aleksenko, A., Belebky, L., Robertson, S., Ramsden, M., Vinetski, Y and Clutterbuck, A.J. (1993) An 'instant gene bank' method for gene cloning by mutant complementation. *Molecular and General Genetics* **242**, 467-471.
- Gentry-Weeks, C.R., Keith, J.M. and Thompson, J. (1993) Toxicity of *Bordetella avium* beta-cystathionine toward MC3T3-E1 osteogenic cells. *Journal of Biological Chemistry* **268**, 7298-7314.
- Gibson, T.J., Coulson, A.R., Sulston, J.E. and Little, P.F.R. (1987) Lorist 2, a cosmid with transcriptional terminators insulating vector genes from interference by promoters within the insert -effect on DNA yield and cloned insert frequency. *Gene* **53**, 275-281.
- Giovanelli, J. (1987) Sulfur amino acids of plants: an overview. *Methods in Enzymology* **143**, 419-426.
- Gish, W. and States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nature Genetics* **3**, 266-272.
- Glaser, J.H. and DeMoss, J.A. (1971) Phenotypic restoration by molybdate of nitrate reductase activity in *chlD* mutants of *Escherichia coli*. *Journal of Bacteriology* **108**, 854-860.
- Glatigny, A. and Scazzocchio, C. (1995) Cloning and molecular characterisation of *hxA*, the gene coding for the xanthine

dehydrogenase (purine hydroxylase I) of *Aspergillus nidulans*.  
*The Journal of Biological Chemistry* **270**, 3534-3550.

Greaves, P. (1989) Molecular genetical analysis of the nitrate utilisation gene cluster of *Aspergillus nidulans*. Ph.D. Thesis, University of St Andrews, UK.

Griffith, O.W. (1987) Mammalian sulfur amino acid metabolism: an overview. *Methods in Enzymology* **143**, 366-376.

Grunden, A.M. and Shanmugam, K.T. (1997) Molybdate transport and regulation in bacteria. *Archives of Microbiology* **168**, 345-354.

Grunden, A.M., Rosentel, J.K., Healy, F.G. and Shanmugam, K.T. (1996) Repression of the *Escherichia coli modABCD* (molybdate transport) operon by ModE. *Journal of Bacteriology* **178**, 735-744.

Guerrero, M.G., Vega, J.M. and Losada, M. (1981) The assimilatory nitrate-reducing system and its regulation. *Annual Review of Plant Physiology* **32**, 169-204.

Gwynne, D.I., Buxton, F.P., Williams, S.A., Garven, S. and Davies, R.W. (1987) Genetically engineered secretion of active interferon and a bacterial endoglucanase from *Aspergillus nidulans*. *Bio/Technology* **5**, 713-719.

Haddock, B.A. and Jones, C.W. (1977) Bacterial respiration. *Bacteriological Reviews* **41**, 47-99.

- Hartley, M.J. (1970) Contrasting complementation patterns in *Aspergillus nidulans*. *Genetical Research* **16**, 123-125.
- Hasona, A., Ray, R.M. and Shanmugam, K.T. (1998) Physiological and genetic analyses leading to identification of a biochemical role for the *moeA* (molybdate metabolism) gene product in *Escherichia coli*. *Journal of Bacteriology* **180**, 1466-1472.
- Hawker, K.L., Montague, P. and Kinghorn, J.R. (1992) Nitrate reductase and nitrite transcript levels in various mutants of *Aspergillus nidulans*: confirmation of autogenous regulation. *Molecular and General Genetics* **231**, 485-488.
- Hayes, P.K. and Powell, R.S. (1995) The *gvpA/C* cluster of *Anabaena flos-aquae* has multiple copies of a gene encoding GvpA. *Archives of Microbiology* **164**, 50-57.
- Hesseltine, C.W. (1965) A millenium of fungi, food and fermentation. *Mycologia* **57**, 149-157.
- Hinton, S.M. and Dean, D. (1990) Biogenesis of molybdenum cofactors. *Critical Reviews in Microbiology* **17**, 169-188.
- Hoff, T., Schnorr, K.M., Meyer, C. and Caboche, M. (1995) Isolation of two *Arabidopsis* cDNAs involved in early steps of molybdenum cofactor biosynthesis by functional complementation of *Escherichia coli* mutants. *The Journal of Biological Chemistry* **270**, 6100-6107.

- Hoover, T.R., Imperial, J., Ludden, P.W. and Shah, V.K. (1989) Homocitrate is a component of the iron-molybdenum cofactor of nitrogenase. *Biochemistry* **28**, 2768-2771.
- Horowitz, N.H. (1945) On the evolution of biochemical biosyntheses. *Proceedings of the National Academy of Sciences (USA)* **31**, 153-157.
- Hull, E.P. (1988) Molecular analysis of the proline catabolism gene cluster of *Aspergillus nidulans* and sequencing of the regulatory gene. Ph.D. Thesis, University of Essex, UK.
- Hyde, G.E., Wilberding, J.A., Meyer, A.L., Campbell, E.R. and Campbell, W.H. (1989) Monoclonal antibody-based immunoaffinity chromatography for purifying corn and squash NADH: nitrate reductases. Evidence for an interchain disulfide bond in nitrate reductase. *Plant Molecular Biology* **13**, 233-246.
- Inoue, H., Sugimoto, M., Inagaki, K., Esaki, N., Soda, K. and Tanaka, H. (1995) Structural analysis of the L-methionine gamma-lyase gene from *Pseudomonas putida*. *Journal of Biochemistry* **117**, 1120-1125.
- Ibbi-Nivol, C., Palmer, T., Whitty, P.W., McNairn, E. and Boxer, D.H. (1995) The *mob* locus of *Escherichia coli* K12 required for molybdenum cofactor biosynthesis is expressed at very low levels. *Microbiology* **141**, 1663-1671.

- James, R., Dean, D.O. and Debbage, J. (1993) Five open reading frames upstream of the *dnaK* gene of *E. coli*. *DNA Sequence* **3**, 327-332.
- Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. *Annual Review of Microbiology* **30**, 409-425.
- Johann, S. and Hinton, S.M. (1987) Cloning and nucleotide sequence of the *chlD* locus. *Journal of Bacteriology* **169**, 1911-1916.
- John, M.A. and Peberdy, J.F. (1984) Transformation of *Aspergillus nidulans* using the *argB* gene. *Enzyme and Microbial Technology* **6**, 386-389.
- Johnson, J.L. (1980) The molybdenum cofactor common to nitrate reductase, xanthine dehydrogenase and sulfite oxidase. In *Molybdenum and Molybdenum-Containing Enzymes* (Edited by M.P. Coughlan), pp. 347-383. Pergamon Press, Oxford.
- Johnson, J.L. and Rajagopalan, K.V. (1982) Structural and metabolic relationship between the molybdenum cofactor and urothione. *Proceedings of the National Academy of Sciences (USA)* **79**, 6856-6860.
- Johnson, M.E. and Rajagopalan, K.V. (1987a) *In vitro* system for molybdopterin biosynthesis. *Journal of Bacteriology* **169**, 110-116.
- Johnson, M.E. and Rajagopalan, K.V. (1987b) Involvement of *chlA*, E, M and N loci in *Escherichia coli* molybdopterin biosynthesis. *Journal of Bacteriology* **169**, 117-125.

- Johnson, J.L., Cohen, H.J. and Rajagopalan, K.V. (1974a) Molecular basis of the biological function of molybdenum. Molybdenum-free sulfite oxidase from livers of tungsten-treated rats. *Journal of Biological Chemistry* **249**, 5046-5055.
- Johnson, J.L., Rajagopalan, R.V. and Cohen, H.J. (1974b) Molecular basis of the biological function of molybdenum. Effect of tungsten on xanthine oxidase and sulfite oxidase in the rat. *Journal of Biological Chemistry* **249**, 859-866.
- Johnson, J.L., Hainline, B.E. and Rajagopalan, K.V. (1980a) Characterisation of the molybdenum cofactor of sulfite oxidase and nitrate reductase: identification of a pteridine as a structural component. *Journal of Biological Chemistry* **255**, 1783-1786.
- Johnson, J.L., Waud, W.R., Rajagopalan, K.V., Duran, M., Beemer, F.A. and Wadman, S.K. (1980b) Inborn errors of molybdenum metabolism: combined deficiencies of sulfite oxidase and xanthine dehydrogenase in a patient lacking the molybdenum cofactor. *Proceedings of the National Academy of Sciences (USA)* **77**, 3715-3719.
- Johnson, J.L., Hainline, B.E., Rajagopalan, K.V. and Arison, B.H. (1984) The pterin component of the molybdenum cofactor: structural characterisation of two fluorescent derivatives. *Journal of Biological Chemistry* **259**, 5414-5422.
- Johnson, J.L., Wuebbens, M.M., Mandell, R. and Shih, V.E. (1989) Molybdenum cofactor biosynthesis in humans. Identification of two

complementation groups of cofactor-deficient patients and preliminary characterisation of a diffusible molybdopterin precursor. *Journal of Clinical Investigation* **83**, 897-903.

Johnson, J.L., Bastian, N.R. and Rajagopalan, K.V. (1990a)

Molybdopterin guanine dinucleotide: a modified form of molybdopterin in the molybdenum cofactor of dimethyl sulfoxide reductase from *Rhodobacter sphaeroides forma specialis denitrificans*. *Proceedings of the National Academy of Sciences (USA)* **87**, 3190-3194.

Johnson, J.L., Rajagopalan, K.V. and Meyer, O. (1990b) Isolation and characterisation of a second molybdopterin dinucleotide:

molybdopterin cytosine dinucleotide. *Archives of Biochemistry and Biophysics* **283**, 542-545.

Johnson, J.L., Indermaur, L.W. and Rajagopalan, K.V. (1991a)

Molybdenum cofactor biosynthesis in *Escherichia coli*.

Requirement of the *chlB* gene product for the formation of molybdopterin guanine dinucleotide. *Journal of Biological Chemistry* **266**, 12140-12145.

Johnson, J.L., Chaudhury, M. and Rajagopalan, K.V. (1991b)

Identification of a molybdopterin-containing molybdenum cofactor in xanthine dehydrogenase from *Pseudomonas aeruginosa*.

*Biofactors* **3**, 103-107

Johnson, J.L., Rajagopalan, K.V., Mukund, S. and Adams, M.W.W.

(1993) Identification of molybdopterin as the organic component

of the tungsten cofactor in four enzymes from hyperthermophilic Archaea. *Journal of Biological Chemistry* **268**, 4848-4852.

Johnston, M., Andrews, S., Brinkman, R., Cooper, J., Ding, H., Dover, J., Du, Z., Favello, A., Fulton, L., Gattung, S., Geisel, C., Kirsten, J., Latreille, P., Louis, E.J., Macri, C., Mardis, E., Menezes, S., Mouser, L., Nhan, M., Rifkin, L., Riles, L., St.Peter, H., Trevaskis, E., Vaughan, K., Vignati, D., Wilcox, L., Wohldman, P., Waterston, R., Wilson, R. and Vaudin, M. (1994) Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VIII. *Science* **265**, 2077-2082.

Johnstone, I.L. (1985) Transformation of *Aspergillus nidulans*. *Microbiological Sciences* **2**, 307-311.

Johnstone, I.L., McCabe, P.C., Greaves, P., Gurr, S.J., Cole, G.E., Brow, M.A.D., Unkles, S.E., Clutterbuck, A.J., Kinghorn, J.R. and Innis, M.A. (1990) Isolation and characterisation of the *crnA-niiA-niaD* gene cluster for nitrate assimilation in *Aspergillus nidulans*. *Gene* **90**, 181-192.

Joshi, M.S., Johnson, J.L. and Rajagopalan, K.V. (1996) Molybdenum cofactor biosynthesis in *Escherichia coli mod* and *mog* mutants. *Journal of Bacteriology* **178**, 4310-4312.

Kalakoutskii, K. and Fernández, E. (1996) Different forms of molybdenum cofactor in *Vicia faba* seeds. The presence of molybdenum cofactor carrier protein and its purification. *Planta* **201**, 64-70.

- Kamdar, K.P., Shelton, M.E. and Finnerty, V. (1994) The *Drosophila* molybdenum cofactor gene *cinnamon* is homologous to three *Escherichia coli* cofactor proteins and to the rat protein Gephyrin. *Genetics* **137**, 791-801.
- Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N., Sugiura, M. and Tabata, S. (1995) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. *DNA Research* **2**, 153-166.
- Kaufmann, E., Geisler, N. and Weber, K. (1984) SDS-PAGE strongly overestimates the molecular masses of the neurofilament proteins. *FEBS Letters* **170**, 81-84.
- Kerjan, P., Cherest, H. and Surdin-Kerjan, Y. (1986) Nucleotide sequence of the *Saccharomyces cerevisiae* *MET25* gene. *Nucleic Acids Research* **14**, 7861-7871.
- Ketchum, P.A. and Downey, R.J. (1975) *In vitro* restoration of nitrate reductase: investigation of *Aspergillus nidulans* and *Neurospora crassa* nitrate reductase mutants. *Biochimica et Biophysica Acta* **385**, 354-361.
- Ketchum, P.A., Cambier, H.Y., Frazier, W.A. III., Madansky, C.H. and Nason, A. (1970) *In vitro* assembly of *Neurospora* assimilatory nitrate reductase from protein subunits of a *Neurospora* mutant and the xanthine oxidising or aldehyde oxidase systems of higher

- animals. *Proceedings of the National Academy of Sciences (USA)* **66**, 1016-1023.
- Kim, J. and Rees, D.C. (1992) Structural models for the metal centers in the nitrogenase molybdenum iron-protein. *Science* **257**, 1677-1682.
- Kinghorn, J.R. and Campbell, E.I. (1990) Amino acid sequence comparisons between bacterial, fungal and plant nitrate reductase and nitrite reductase proteins. In *Molecular and Genetic Aspects of Nitrate Assimilation* (Edited by J.L. Wray and J.R. Kinghorn), pp. 385-403. Oxford University Press, Oxford.
- Kinghorn, J.R. and Hawkins, A.R. (1982) Cloning and expression in *Escherichia coli* K-12 of the biosynthetic dehydroquinase function of the *arom* cluster gene from the eukaryote, *Aspergillus nidulans*. *Molecular and General Genetics* **186**, 145-152.
- Kinghorn, J.R. and Unkles, S.E. (1994) Molecular genetics and expression of foreign proteins in the genus *Aspergillus*. In *Aspergillus -Biotechnology Handbooks 7* (Edited by J.E. Smith), pp. 64-100. Plenum Press, New York.
- Kinghorn, J.R., Montague, P. and Unkles, S.E. (1990) Regulatory, structural and biotechnological aspects of the *Aspergillus nidulans* gene cluster for nitrate assimilation. In *Inorganic Nitrogen in Plants and Microorganisms; Uptake and Metabolism* (Edited by W.R. Ullrich, C. Rigano, A. Fuggi and P.J. Aparicio), pp. 287-295. Springer Verlag Press, Berlin.

- Kirsch, J. and Betz, H. (1993) Widespread expression of gephyrin, a putative glycine receptor-tubulin linker protein, in rat brain. *Brain Research* **621**, 301-310.
- Kirsch, J., Wolters, I., Triller, A. and Betz, H. (1993) Gephyrin antisense oligonucleotides prevent glycine receptor clustering in the spinal neurons. *Nature* **366**, 745-748.
- Kisker, C., Schindelin, H. and Rees, D.C. (1997) Molybdenum cofactor-containing enzymes: structure and mechanism. *Annual Review of Biochemistry* **66**, 233-267.
- Koschara, W. (1940) Urothione, a yellow pigment rich in sulphur from human urine. *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* **263**, 78-79.
- Kramer, S.P., Johnson, J.L., Ribeiro, A.A., Millington, D.S. and Rajagopalan, K.V. (1987) The structure of the molybdenum cofactor. Characterisation of di-(carboxamidomethyl) molybdopterin from sulfite oxidase and xanthine oxidase. *Journal of Biological Chemistry* **262**, 16357-16363.
- Krüger, B. and Meyer, O. (1986) The pterin (bactopterin) of carbon monoxide dehydrogenase from *Pseudomonas carboxydoflava*. *European Journal of Biochemistry* **157**, 121-128.
- Krüger, B. and Meyer, O. (1987) Structural elements of bactopterin from *Pseudomonas carboxydoflava* carbon monoxide dehydrogenase. *Biochimica et Biophysica Acta* **912**, 357-364.

- Kurlemann, G., Debus, O. and Schuierer, G. (1996) Dextromethorphan in molybdenum cofactor deficiency. *European Journal of Pediatrics* **155**, 422-423.
- Kuykendall, L.D. and Hunter, W.J. (1997) The sequence of a symbiotically essential *Bradyrhizobium japonicum* operon consisting of *trpD*, *trpC* and a *moaC*-like gene. *Biochimica et Biophysica Acta* **1350**, 277-281.
- Kwok, S. and Higuchi, R. (1989) Avoiding false positives with PCR. *Nature* **339**, 237-238.
- LaBrie, S.T. and Crawford, N.M. (1994) A glycine to aspartic acid change in the MoCo domain of nitrate reductase reduces both activity and phosphorylation levels in *Arabidopsis*. *The Journal of Biological Chemistry* **269**, 14497-14501.
- Lamb, H.K., Bagshaw, C.R. and Hawkins, A.R. (1991) *In vivo* overproduction of the pentafunctional *arom* polypeptide in *Aspergillus nidulans* affects metabolic flux in the quinate pathway. *Molecular and General Genetics* **227**, 187-196.
- Leach, J., Finkelstein, D.B. and Rambosek, J.A. (1986) Rapid miniprep of DNA from filamentous fungi. *Neurospora Newsletter* **33**, 32-33.
- Lee, K-Y., Erickson, R., Pan, S-S., Jones, G., May, F. and Nason, A. (1974) Effect of tungsten and vanadium on the *in vitro* assembly of assimilatory nitrate reductase utilising *Neurospora* mutant *nit-1*. *Journal of Biological Chemistry* **249**, 3953-3959.

- le Guen, L., Thomas, M. and Kreis, M. (1994) Gene density and organisation in a small region of the *Arabidopsis thaliana* genome. *Molecular and General Genetics* **245**, 390-396.
- Lewis, N.J. and Scazzocchio, C. (1977) The genetic control of molybdoflavoproteins in *Aspergillus nidulans*. A xanthine dehydrogenase I half-molecule in *cnx* mutant strains of *Aspergillus nidulans*. *European Journal of Biochemistry* **76**, 441-446.
- Lewis, N.J., Hurt, P., Sealy-Lewis, H.M. and Scazzocchio, C. (1978) The genetic control of molybdo-flavoproteins in *Aspergillus nidulans*. IV. A comparison between purine hydroxylase I and II. *European Journal of Biochemistry* **91**, 311-316.
- Li, Y. and Strohl, W.R. (1996) Cloning, purification, and properties of a phosphotyrosine protein phosphatase from *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* **178**, 136-142.
- Lorito, M., Hayes, C.K., Di Pietro, A. and Harman, G.E. (1993) Biolistic transformation of *Trichoderma harzianum* and *Gliocladium virens* using plasmid and genomic DNA. *Current Genetics* **24**, 349-356.
- Lu, Y., O'Dowd, B.F., Orrego, H. and Israel, Y. (1992) Cloning and nucleotide sequence of human liver cDNA encoding for cystathionine gamma-lyase. *Biochemical and Biophysical Research Communications* **189**, 749-758.
- Lueder, G.T. and Steiner, R.D. (1995) Ophthalmic abnormalities in molybdenum cofactor deficiency and isolated sulfite oxidase

deficiency. *Journal of Pediatric Ophthalmology and Strabismus* **32**, 334-337.

Luque, F., Mitchenall, L.A., Chapman, M., Christine, R. and Pau, R.N. (1993) Characterisation of genes involved in molybdenum transport in *Azotobacter vinelandii*. *Molecular Microbiology* **7**, 447-459.

MacCabe, A., Riach, M., Unkles, S.E. and Kinghorn, J.R. (1990) The *Aspergillus nidulans npeA* locus consists of three contiguous genes required for penicillin biosynthesis. *EMBO Journal* **9**, 279-287.

MacDonald, D.W. and Cove, D.J. (1974) Studies on temperature-sensitive mutants affecting the assimilatory nitrate reductase of *Aspergillus nidulans*. *European Journal of Biochemistry* **47**, 107-110.

MacDonald, D.W., Cove, D.J. and Coddington, A. (1974) Cytochrome *c* reductase from wild-type and mutant strains of *Aspergillus nidulans*. *Molecular and General Genetics* **128**, 187-199.

Martinelli, S.D. (1994) *Aspergillus nidulans* as an experimental organism. In *Aspergillus: 50 Years On* (Edited by S.D. Martinelli, and J.R. Kinghorn), pp. 33-59. Elsevier, Amsterdam.

Maupin-Furlow, J.A., Rosentel, J.K., Lee, J.H., Deppenmeier, U., Gunsalus, R.P. and Shanmugam, K.T. (1995) Genetic analysis of the *modABCD* (molybdate transport) operon of *Escherichia coli*. *Journal of Bacteriology* **177**, 4851-4856.

- McNicholas, P.M., Rech, S.A. and Gunsalus, R.P. (1997) Characterisation of the ModE DNA-binding sites in the control regions of *modABCD* and *moaABCDE* of *Escherichia coli*. *Molecular Microbiology* **23**, 515-524.
- Mehra, R.K. and Coughlan, M.P. (1984) Purification and properties of purine hydroxylase II from *Aspergillus nidulans*. *Archives of Biochemistry and Biophysics* **229**, 585-595.
- Mellor, R.B., Ronnenberg, J., Campbell, W.H. and Diekmann, S. (1992) Reduction of nitrate and nitrite in water by immobilised enzymes. *Nature* **355**, 717-719.
- Melnick, L. and Sherman, F. (1993) The gene clusters ARC and COR on chromosomes 5 and 10, respectively, of *Saccharomyces cerevisiae* share a common ancestry. *Journal of Molecular Biology* **233**, 372-388.
- Mendel, R.R. (1997) Molybdenum cofactor of higher plants: biosynthesis and molecular biology. *Planta* **203**, 399-405.
- Menéndez, C., Igloi, G., Henninger, H-P. and Brandsch, R. (1995) A pAO1-encoded molybdopterin gene (*moaA*) of *Arthrobacter nicotinovorans*: characterisation and site-directed mutagenesis of the encoded protein. *Archives of Microbiology* **164**, 142-151.
- Menéndez, C., Siebert, D. and Brandsch, R. (1996) MoaA of *Arthrobacter nicotinovorans* pAO1 involved in Mo-pterin cofactor synthesis is an Fe-S protein. *FEBS Letters* **391**, 101-103.

- Menéndez, C., Otto, A., Igloi, G., Nick, P., Brandsch, R., Schubach, B., Bottcher, B. and Brandsch, R. (1997) Molybdate-uptake genes and molybdopterin biosynthesis genes on a bacterial plasmid -characterisation of MoeA as a filament-forming protein with adenosinetriphosphatase activity. *European Journal of Biochemistry* **250**, 524-531.
- Meyer, C., Gonneau, M., Caboche, M. and Rouzé, P. (1995) Identification by mutational analysis of four critical residues in the molybdenum cofactor domain of eukaryotic nitrate reductase. *FEBS Letters* **370**, 197-202.
- Miller, J.B. and Amy, N.K. (1983) Molybdenum cofactor in chlorate-resistant and nitrate reductase-deficient insertion mutants of *Escherichia coli*. *Journal of Bacteriology* **155**, 793-801.
- Miller, J.B., Scott, D.B. and Amy, N.K. (1987) Molybdenum-sensitive transcriptional regulation of the *chlD* locus of *Escherichia coli*. *Journal of Bacteriology* **169**, 1853-1860.
- Minagawa, N. and Yoshimoto, A. (1982) Purification and characterisation of the assimilatory NADPH-nitrate reductase of *Aspergillus nidulans*. *Journal of Biochemistry* **91**, 761-774.
- Mize, C., Johnson, J.L. and Rajagopalan, K.V. (1995) Defective molybdopterin biosynthesis: clinical heterogeneity associated with molybdenum cofactor deficiency. *Journal of Inherited Metabolic Disease* **18**, 283-290.

- Mouncey, N.J., Mitchenall, L.A. and Pau, R.N. (1995) Mutational analysis of genes of the *mod* locus involved in molybdenum transport, homeostasis, and processing in *Azotobacter vinelandii*. *Journal of Bacteriology* **177**, 5294-5302.
- Murakami, Y., Naitou, M., Hagiwara, H., Shibata, T., Ozawa, M., Sasanuma, S-I., Sasanuma, M., Tsuchiya, Y., Soeda, E., Yokoyama, K., Yamazaki, M., Tashiro, H. and Eki, T. (1995) Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces cerevisiae*. *Nature Genetics* **10**, 261-268.
- Nason, A., Antoine, A.D., Ketchum, P.A., Frazier, W.A. III. and Lee, D.K. (1970) Formation of assimilatory nitrate reductase by *in vitro* interchromosomal complementation in *Neurospora crassa*. *Proceedings of the National Academy of Sciences (USA)* **65**, 137-144.
- Nason, A., Lee, K-Y., Pan, S-S., Ketchum, P.A., Lamberti, A. and Devries, J. (1971) *In vitro* formation of assimilatory reduced nicotinamide adenine dinucleotide phosphate: nitrate reductase from a *Neurospora* mutant and a component of molybdenum-enzymes. *Proceedings of the National Academy of Sciences (USA)* **68**, 3242-3246.
- Nishi, N., Tanabe, H., Oya, H., Urushihara, M., Miyanaka, H. and Wada, F. (1994) Identification of probasin-related antigen as cystathionine gamma-lyase by molecular cloning. *Journal of Biological Chemistry* **269**, 1015-1019.

- Nohno, T., Kasai, Y. and Saito, T. (1988) Cloning and complete nucleotide sequence of the *Escherichia coli* *chlEN* operon involved in molybdopterin biosynthesis. *Journal of Bacteriology* **170**, 4097-4102.
- Norbeck, J. and Blomberg, A. (1996) Protein expression during exponential growth in 0.7 M NaCl medium of *Saccharomyces cerevisiae*. *FEMS Microbiology Letters* **137**, 1-8.
- Ogier, J., Wadman, S., Johnson, J.L., Saudubray, J.M., Duran, M., Boue, J., Munnich, A. and Charpentier, C. (1983) Antenatal diagnosis of combined xanthine and sulphite oxidase deficiencies. *Lancet* **2**, 1363-1364.
- Okamoto, P.M. and Marzluf, G.A. (1993) Nitrate reductase of *Neurospora crassa*: the functional role of individual amino acids in the haem domain as examined by site-directed mutagenesis. *Molecular and General Genetics* **240**, 221-230.
- Ono, B-I., Tanaka, K., Naito, K., Heike, C., Shinoda, S., Yamamoto, S., Ohmori, S., Oshima, T. and Toh-E, A. (1992) Cloning and characterisation of the *CYS3* (*CYII*) gene of *Saccharomyces cerevisiae*. *Journal of Bacteriology* **174**, 3339-3347.
- Ono, B-I., Ishii, N., Naito, K., Miyoshi, S-I., Shinoda, S., Yamamoto, S. and Ohmori, S. (1993) Cystathionine gamma-lyase of *Saccharomyces cerevisiae*: structural gene and cystathionine gamma-synthase activity. *Yeast* **9**, 389-397.

- Ouellette, B.F.F., Clark, M.W., Keng, T., Storms, R.K., Zhong, W.W., Zeng, B., Fortin, N., Delaney, S., Barton, A.B., Kaback, D.B. and Bussey, H. (1993) Sequencing of chromosome I from *Saccharomyces cerevisiae*: analysis of a 32 kb region between the *LTE1* and *SPO7* genes. *Genome* **36**, 32-42.
- Palmer, T., Vasishta, A., Whitty, P.W. and Boxer, D.H. (1994) Isolation of protein FA: a product of the *mob* locus required for molybdenum cofactor biosynthesis in *Escherichia coli*. *European Journal of Biochemistry* **222**, 687-692.
- Palmer, T., Goodfellow, I.P., Sockett, R.E., McEwan, A.G. and Boxer, D.H. (1998) Characterisation of the *mob* locus from *Rhodobacter sphaeroides* required for molybdenum cofactor biosynthesis. *Biochimica et Biophysica Acta* **1395**, 135-140.
- Panyim, S. and Chalkley, R. (1969) High resolution acrylamide gel electrophoresis of histones. *Archives of Biochemistry and Biophysics* **130**, 337-346.
- Park, Y.M. and Stauffer, G.V. (1989) DNA sequence of the *metC* gene and its flanking regions from *Salmonella typhimurium* LT2 and homology with the corresponding sequence of *Escherichia coli*. *Molecular and General Genetics* **216**, 164-169.
- Paszewski, A. and Grabski, J. (1974) Regulation of S-amino acid biosynthesis in *Aspergillus nidulans*. *Molecular and General Genetics* **132**, 307-320.

- Paszewski, A. and Grabski, J. (1975) Enzymatic lesions in methionine mutants of *Aspergillus nidulans*: role and regulation of an alternative pathway for cysteine and methionine synthesis. *Journal of Bacteriology* **124**, 893-904.
- Paszewski, A., Brzywezy, J. and Natorff, R. (1994) Sulphur metabolism In *Aspergillus: 50 Years On* (Edited by S.D. Martinelli and J.R. Kinghorn), pp. 299-319. Elsevier, Amsterdam.
- Pateman, J.A., Cove, D.J., Rever, B.M. and Roberts, D.B. (1964) A common cofactor for nitrate reductase and xanthine dehydrogenase which also regulates the synthesis of nitrate reductase. *Nature* **201**, 58-60.
- Pateman, J.A., Rever, B.M. and Cove, D.J. (1967) Genetical and biochemical studies of nitrate assimilation in *Aspergillus nidulans*. *Biochemical Journal* **104**, 103-111.
- Pienkos, P.T., Shah, V.K. and Brill, W.J. (1977) Molybdenum cofactors from molybdoenzymes and *in vitro* reconstitution of nitrogenase and nitrate reductase. *Proceedings of the National Academy of Sciences (USA)* **74**, 5468-5471.
- Piotrowska, M. and Paszewski, A. (1990) A yeast with unusual sulphur amino acid metabolism. *Journal of General Microbiology* **136**, 2283-2286.
- Pitterle, D.M. and Rajagopalan, K.V. (1993) The biosynthesis of molybdopterin in *Escherichia coli*. Purification and characterisation

of the converting factor. *The Journal of Biological Chemistry* **268**, 13499-13505.

Pitterle, D.M., Johnson, J.L. and Rajagopalan, K.V. (1993) *In vitro* synthesis of molybdopterin from precursor Z using purified converting factor. Role of protein-bound sulfur in formation of the dithiolene. *The Journal of Biological Chemistry* **268**, 13506-13509.

Plunkett, G. III., Burland, V., Daniels, D.L. and Blattner, F.R. (1993) Analysis of the *Escherichia coli* genome. III. DNA sequence of the region from 87.2 to 89.2 minutes. *Nucleic Acids Research* **21**, 3391-3398.

Pontecorvo, G., Roper, J.A., Hemmons, L.H., MacDonald, K.D. and Bufton, A.W.J. (1953) The genetics of *Aspergillus nidulans*. *Advances in Genetics* **5**, 141-238.

Prior, P., Schmitt, B., Grenningloh, G., Pribilla, I., Multhaup, G., Beyreuther, K., Maulet, Y., Werner, P., Langosch, D. and Kirsch, J. (1992) Primary structure and alternative splice variants of gephyrin, a putative glycine receptor-tubulin linker protein. *Neuron* **8**, 1161-1170.

Punt, P.J., Oliver, R.P., Dingemans, M.A., Pouwels, P.H. and van den Hondel, C.A.M.J.J. (1987) Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. *Gene* **56**, 117-124.

- Rahman, S., Oberdorf, A., Montecino, M., Tanhauser, S.M., Lian, J.B., Stein, G.S., Laipis, P.J. and Stein, J.L. (1993) Multiple copies of the bone-specific osteocalcin gene in mouse and rat. *Endocrinology* **133**, 3050-3053.
- Rajagopalan, K.V. and Johnson, J.L. (1992) The pterin molybdenum cofactors. *The Journal of Biological Chemistry* **267**, 10199-10202.
- Ramaswamy, K.S., Endley, S. and Golden, J.W. (1996) Nitrate reductase activity and heterocyst suppression on nitrate in *Anabaena* sp. strain PCC7120 require *moeA*. *Journal of Bacteriology* **178**, 3893-3898.
- Ravanel, S., Ruffet, M.L. and Douce, R. (1995) Cloning of an *Arabidopsis thaliana* cDNA encoding cystathionine beta-lyase by functional complementation in *Escherichia coli*. *Plant Molecular Biology* **29**, 875-882.
- Rech, S., Deppenmeier, U. and Gunsalus, R.P. (1995) Regulation of the molybdate operon, *modABCD*, of *Escherichia coli* in response to molybdate availability. *Journal of Bacteriology* **177**, 1023-1029.
- Reiss, J., Kleinhofs, A. and Klingmuller, W. (1987) Cloning of seven differently complementing DNA fragments with *chl* functions from *Escherichia coli* K12. *Molecular and General Genetics* **206**, 352-355.
- Riach, M.B.R. and Kinghorn, J.R. (1996) Genetic transformation and vector developments in filamentous fungi. In *Fungal Genetics*,

*Principles and Practice* (Edited by K. Bos), pp. 209-233.  
Wiley Press, New York.

Richey, M.G., Marek, E.T., Schardl, C.L. and Smith, D.A. (1989)  
Transformation of filamentous fungi with plasmid DNA by  
electroporation. *Phytopathology* **79**, 844-853.

Rivers, S.L. (1991) A physical and genetic analysis of the *chlA* locus of  
*Escherichia coli*. Ph.D. Thesis, University of Dundee, UK.

Rivers, S.L., McNairn, E., Blasco, F., Giordano, G. and Boxer, D.H.  
(1993) Molecular genetic analysis of the *moa* operon of *Escherichia*  
*coli* K12 required for molybdenum cofactor biosynthesis.  
*Molecular Microbiology* **8**, 1071-1081.

Roberts, C.F. (1969) Isolation of multiple aromatic mutants in  
*Aspergillus nidulans*. *Aspergillus Newsletter* **10**, 19-20.

Rouzé, P. and Caboche, M. (1992) Nitrate reduction in higher plants:  
molecular approaches to function and regulation. In *Inducible Plant*  
*Proteins* (Edited by J.L. Wray), pp. 45-77. Cambridge University  
Press, Cambridge.

Rubio, L.M., Flores, E. and Herrero, A. (1998) The *narA* locus of  
*Synechococcus* sp. strain PCC 7942 consists of a cluster of  
molybdopterin biosynthesis genes. *Journal of Bacteriology* **180**,  
1200-1206.

- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual. Second Edition, Cold Spring Harbor Laboratory Press, New York.
- Sanderson, K.E. and Hartman, P.E. (1978) Linkage map of *Salmonella typhimurium*, edition V. *Microbiological Reviews* **42**, 471-519.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences (USA)* **74**, 5463-5467.
- Santini, C.L., Iobbi-Nivol, C., Romane, C., Boxer, D.H. and Giordano, G. (1992) Molybdoenzyme biosynthesis in *Escherichia coli*: *in vitro* activation of purified nitrate reductase from a *chlB* mutant. *Journal of Bacteriology* **174**, 7934-7940.
- Scazzocchio, C. (1973) The genetic control of molybdoflavoproteins in *Aspergillus nidulans*. Allopurinol-resistant mutants constitutive for xanthine dehydrogenase. *European Journal of Biochemistry* **36**, 428-445.
- Scazzocchio, C. (1974) The genetic determination of molybdoflavoenzymes in *Aspergillus nidulans*. *Journal of the Less-Common Metals* **36**, 461-464.

- Scazzocchio, C. (1980) The genetics of the molybdenum-containing enzymes. In *Molybdenum and Molybdenum-Containing Enzymes* (Edited by M.P. Coughlan), pp. 487-515. Pergamon Press, Oxford.
- Scazzocchio, C. (1994) The purine degradation pathway, genetics, biochemistry and regulation. In *Aspergillus: 50 Years On* (Edited by S.D. Martinelli and J.R. Kinghorn), pp. 221-257. Elsevier, Amsterdam.
- Scazzocchio, C. and Arst, H.N. Jr. (1989) Regulation of nitrate assimilation in *Aspergillus nidulans*. In *Molecular and Genetic Aspects of Nitrate Assimilation* (Edited by J.L. Wray and J.R. Kinghorn), pp. 314-363. Oxford University Press, Oxford.
- Scazzocchio, C., Holl, F.B. and Foguelman, A.I. (1973) The genetic control of molybdoflavoproteins in *Aspergillus nidulans*. Allopurinol resistant mutants constitutive for xanthine dehydrogenase. *European Journal of Biochemistry* **36**, 428-439.
- Schwarz, G., Boxer, D.H. and Mendel, R.R. (1997) Molybdenum cofactor biosynthesis. The plant protein Cnx1 binds molybdopterin with high affinity. *Journal of Biological Chemistry* **272**, 26811-26814.
- Sealy-Lewis, H.M., Scazzocchio, C. and Lee, S. (1978) A mutation defective in the xanthine alternative pathway of *Aspergillus nidulans*: Its use to investigate the specificity of *uaY* mediated induction. *Molecular and General Genetics* **164**, 303-308.

- See, Y.P. and Jackowski, G. (1988) *Protein Structure: A Practical Approach* (Edited by T.E. Creighton). University of Oxford Press, Oxford.
- Seeburg, P.M., Colby, W.W., Capon, D.J., Goedel, D.V. and Levinson, A.D. (1984) Biological properties of human c-Ha-ras1 genes mutated at codon 12. *Nature* **312**, 71-75.
- Shah, V.K. and Brill, W.J. (1977) Isolation of an iron-molybdenum cofactor from nitrogenase. *Proceedings of the National Academy of Sciences (USA)* **74**, 3249-3253.
- Shanmugam, K.T., Stewart, V., Gunsalus, R.P., Boxer, D.H., Cole, J.A., Chippaux, M., DeMoss, J.A., Giordano, G., Lin, E.C.C. and Rajagopalan, K.V. (1992) Proposed nomenclature for the genes involved in molybdenum metabolism in *Escherichia coli* and *Salmonella typhimurium*. *Molecular Microbiology* **6**, 3451-3454.
- Sharp, P.A., Sugden, B. and Sambrook, J. (1973) Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose ethidium bromide electrophoresis. *Biochemistry* **12**, 3055-3063.
- Singh, H.N. and Sonie, K.C. (1977) Isolation and characterisation of chlorate-resistant mutants of the blue-green alga *Nostoc muscorum*. *Mutation Research* **43**, 205-212.

- Singleton, P. and Sainsbury, D. (1993) In *Dictionary of Microbiology and Molecular Biology* (Edited by P. Singleton and D. Sainsbury), pp. 741-742. John Wiley and Sons, Chichester.
- Slot, H.M., Overweg-Plandsoen, W.C., Bakker, H.D., Abeling, N.G., Tamminga, P., Barth, P.G. and Van Gennip, A.H. (1993) Molybdenum-cofactor deficiency: an easily missed cause of neonatal convulsions. *Neuropediatrics* **24**, 139-142.
- Soda, K. (1987) Microbiological sulfur amino acids: an overview. *Methods in Enzymology* **143**, 453-459.
- Solomonson, L.P. and Barber, M.J. (1990) Assimilatory nitrate reductase: Functional properties and regulation. *Annual Review of Plant Physiology and Plant Molecular Biology* **41**, 225-253.
- Solomonson, L.P., Barber, M.J., Robbins, A.P. and Oaks, A. (1986) Functional domains of assimilatory NADH: nitrate reductase from *Chlorella*. *Journal of Biological Chemistry* **261**, 11290-11294.
- Sophianopoulou, V. and Scazzocchio, C. (1989) The proline transport protein of *Aspergillus nidulans* is very similar to amino acid transporters of *Saccharomyces cerevisiae*. *Molecular Microbiology* **3**, 705-714.
- Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503-517.

- Stallmeyer, B., Nerlich, A., Schiemann, J., Brinkmann, H. and Mendel, R.R. (1995) Molybdenum cofactor 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. *The Plant Journal* **8**, 101-112.
- Steiner, F.X. and Downey, R.J. (1982) Isoelectric focusing and two dimensional analysis of purified nitrate reductase from *Aspergillus nidulans*. *Biochimica et Biophysica Acta* **706**, 203-211.
- Stewart, V. and MacGregor, C.H. (1982) Nitrate reductase in *Escherichia coli* K12: involvement of *chlC*, *chlE* and *chlG* loci. *Journal of Bacteriology* **151**, 788-799.
- Stouthamer, A.H. (1969) A genetical and biochemical study of chlorate-resistant mutants of *Salmonella typhimurium*. *Antonie van Leeuwenhoek* **35**, 505-521.
- Studier, F.W. and Moffat, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113-130.
- Tilburn, J., Scazzocchio, C., Taylor, G.G., Zabicky-Zissman, J.H., Lockington, R.A. and Davies, R.W. (1983) Transformation by integration in *Aspergillus nidulans*. *Gene* **26**, 205-221.
- Timberlake, W.E. and Marshall, M.A. (1989) Genetic engineering of filamentous fungi. *Science* **244**, 1313-1317.

Tomb, J-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H-P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C. (1997) The complete genome sequence of *Helicobacter pylori*. *Nature* **388**, 539-547.

Tomsett, A.B. (1989) The genetics and biochemistry of nitrate assimilation in ascomycete fungi. In *Nitrogen, Phosphorus and Sulphur Utilisation by Fungi* (Edited by L. Boddy, R. Merchant and D.J. Read), pp. 33-57. Cambridge University Press, Cambridge.

Tomsett, A.B. and Cove, D.J. (1979) Deletion mapping of the *niiA-niaD* gene region of *Aspergillus nidulans*. *Genetical Research* **34**, 19-32.

Turnbull, I.F., Smith, D.R.J., Sharp, P.J., Cobon, G.S. and Hynes, M.J. (1990) Expression and secretion in *Aspergillus nidulans* and *Aspergillus niger* of a cell surface glycoprotein from the cattle tick, *Boophilus microplus*, by using the fungal *amdS* promoter system. *Applied and Environmental Microbiology* **56**, 2847-2852.

Ullmann, A., Jacob, F. and Monod, J. (1967) Characterisation by *in vitro* complementation of a peptide corresponding to an operator-

proximal segment of the  $\beta$ -galactosidase structural gene of *Escherichia coli*. *Journal of Molecular Biology* **24**, 339-343.

Unkles, S.E. (1992) Gene organisation in industrial filamentous fungi. In *Applied Molecular Genetics of Filamentous Fungi* (Edited by J.R. Kinghorn and G.Turner), pp. 28-53. Blackie Academic and Professional Press, Glasgow.

Unkles, S.E., Hawker, K.L., Grieve, C., Campbell, E.I., Montague, P. and Kinghorn, J.R. (1991) *crnA* encodes a nitrate transporter in *Aspergillus nidulans*. *Proceedings of the National Academy of Sciences (USA)* **88**, 204-208.

van den Hondel, C.A.M.J.J., Punt, P.J. and van Gorcom, R.F.M. (1992) Heterologous gene expression in filamentous fungi. In *More Gene Manipulations in Fungi* (Edited by J.W. Bennett and L.L. Lasure), pp. 396-428. Academic Press, San Diego.

van Gennip, A.H., Abeling, N.G., Stroomer, A.E., Overmars, H. and Bakker, H.D. (1994) The detection of molybdenum cofactor deficiency: clinical symptomatology and urinary metabolite profile. *Journal of Inherited Metabolic Disease* **17**, 142-145.

Venables, W.A. and Guest, J.R. (1968) Transduction of nitrate reductase loci of *Escherichia coli* by phages P1 and  $\lambda$ . *Molecular and General Genetics* **103**, 127-140.

- Vogels, G.H. and van der Drift, C. (1976) Degradation of purines and pyrimidines by microorganisms. *Bacteriological Reviews* **40**, 403-468.
- Vogelstein, B. and Gillespie, D. (1979) Preparative and analytical purification of DNA from agarose. *Proceedings of the National Academy of Sciences (USA)* **76**, 615-619.
- Wadman, K.S., Duran, M., Beemer, F.A., Cats, B.P., Johnson, J.L., Rajagopalan, K.V., Saudubray, J.M., Ogier, H., Charpentier, C., Berger, R., Smit, G.P.A., Wilson, J. and Krywawych, S. (1983) Absence of hepatic molybdenum cofactor: an inborn error of metabolism leading to a combined deficiency of sulphite oxidase and xanthine dehydrogenase. *Journal of Inherited Metabolic Disease* **6**, 78-83.
- Wahl, R.C., Warner, C.K., Finnerty, V. and Rajagopalan, K.V. (1982) *Drosophila melanogaster* *ma-1* mutants are defective in the sulfuration of desulfo Mo hydroxylases. *The Journal of Biological Chemistry* **257**, 3958-3962.
- Wahl, G.M., Lewis, K.A., Ruiz, J.C., Rothenberg, B., Zhao, J. and Evans, G.A. (1987) Cosmid vectors for rapid genomic walking, restriction mapping and gene transfer. *Proceedings of the National Academy of Sciences (USA)* **84**, 2160-2164.
- Walkenhorst, H.M., Hemschemeier, S.K. and Eichenlaub, R. (1995) Molecular analysis of the molybdate uptake operon, *modABCD*, of

*Escherichia coli* and *modR*, a regulatory gene. *Microbiological Research* **150**, 347-361.

Wang, G., Angermuller, S. and Klipp, W. (1993) Characterisation of *Rhodobacter capsulatus* genes encoding a molybdenum transport system and putative molybdenum-pterin-binding proteins. *Journal of Bacteriology* **175**, 3031-3042.

Warner, C.K. and Finnerty, V. (1981) Molybdenum hydroxylases in *Drosophila*. II Molybdenum cofactor in xanthine dehydrogenase, aldehyde oxidase and pyridoxal oxidase. *Molecular and General Genetics* **184**, 92-96.

Wernars, K., Goosen, T., Wennekes, B.M.J., Swart, K., van den Hondel, C.A.M.J.J. and van den Broek, H.W.J. (1987) Cotransformation of *Aspergillus nidulans*: a tool for replacing fungal genes. *Molecular and General Genetics* **209**, 71-77.

Wilkinson, J.Q. and Crawford, N.M. (1993) Identification and characterisation of a chlorate-resistant mutant of *Arabidopsis thaliana* with mutations in both nitrate reductase structural genes *NIA1* and *NIA2*. *Molecular and General Genetics* **239**, 289-297.

Williams, L.G., Bernhardt, S.A. and Davis, R.H. (1970) Copurification of pyrimidine-specific carbamyl phosphate synthetase and aspartate transcarbamylase of *Neurospora crassa*. *Biochemistry* **9**, 4329-4335.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latreille, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J. and Wohldman, P. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**, 32-38.

Wuebbens, M.M and Rajagopalan, K.V. (1993) Structural characterisation of a molybdopterin precursor. *The Journal of Biological Chemistry* **268**, 13493-13498.

Wuebbens, M.M. and Rajagopalan, K.V. (1995) Investigation of the early steps of molybdopterin biosynthesis in *Escherichia coli* through the use of *in vivo* labelling studies *The Journal of Biological Chemistry* **270**, 1082-1087.

Yamamoto, I., Saki, T., Liu, S-M. and Ljungdahl, L.G. (1983) Purification and properties of NADP-dependant formate dehydrogenase from *Clostridium thermoaceticum*, a tungsten-selenium-iron protein. *Journal of Biological Chemistry* **258**, 1826-1832.

- Yamagata, S., D'Andrea, R.J., Fujisaki, S., Isaji, M. and Nakamura, K. (1993) Cloning and bacterial expression of the *CYS3* gene encoding cystathionine gamma-lyase of *Saccharomyces cerevisiae* and the physicochemical and enzymatic properties of the protein. *Journal of Bacteriology* **175**, 4800-4808.
- Yelton, M.M., Hamer, J.E. and Timberlake, W.E. (1984) Transformation of *Aspergillus nidulans* by using a *trpC* plasmid. *Proceedings of the National Academy of Sciences (USA)* **76**, 5259-5263.
- Yelton, M.M., Timberlake, W.E. and van den Hondel, C.A.M.J.J. (1985) A cosmid for selecting genes by complementation in *Aspergillus nidulans*: selection of the developmentally regulated *yA* locus. *Proceedings of the National Academy of Sciences (USA)* **82**, 834-838.
- Yura, T., Mori, H., Nagai, H., Nagata, T., Ishihama, A., Fujita, N., Isono, K. and Mizobuchi, K. (1992) Systematic sequencing of the *Escherichia coli* genome: analysis of the 0-2.4 min region. *Nucleic Acids Research* **20**, 3305-3308.
- Yurdakök, M. and Coskun, T. (1997) Hypouricemia and molybdenum-cofactor deficiency. *Journal of Pediatrics* **130**, 162.

## APPENDIX

### Amino acids and their symbols:

<u>Amino Acid</u>	<u>Single Letter Code</u>	<u>Three Letter Abbreviation</u>
Alanine	A	Ala
Cysteine	C	Cys
Aspartic Acid	D	Asp
Glutamic Acid	E	Glu
Phenylalanine	F	Phe
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Lysine	K	Lys
Leucine	L	Leu
Methionine	M	Met
Asparagine	N	Asn
Proline	P	Pro
Glutamine	Q	Gln
Arginine	R	Arg
Serine	S	Ser
Threonine	T	Thr
Valine	V	Val
Tryptophan	W	Trp
Tyrosine	Y	Tyr

Equation used to calculate the standard deviation:

$$SD = \sqrt{s^2}$$

where SD = the standard deviation

$s^2$  = the variance

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

where  $s^2$  = variance of the data

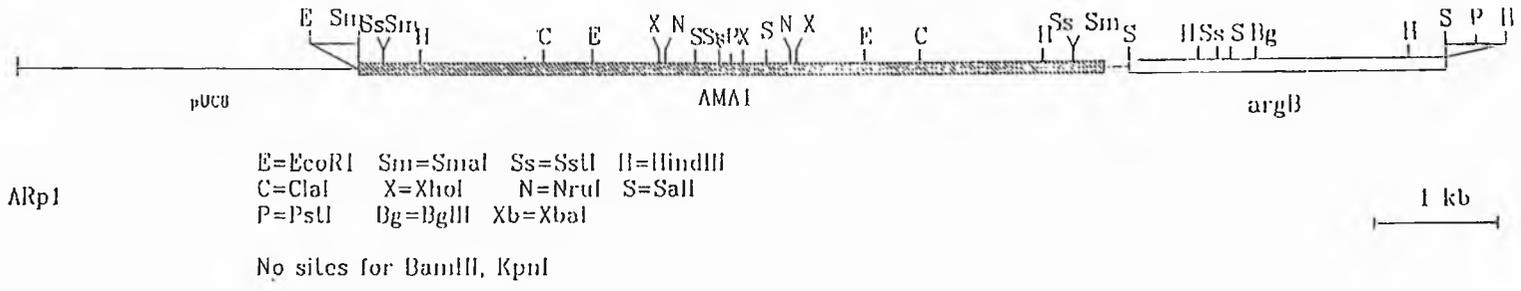
$\sum$  = the sum of

$x$  = the numerical value of a data point

$\bar{x}$  = the mean of the data

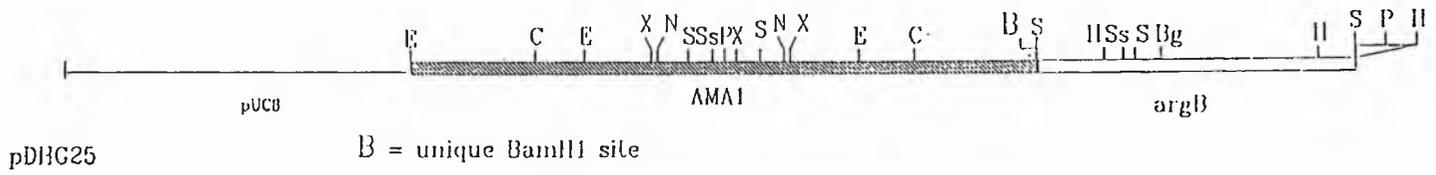
$n$  = the number of data points (sample size)

**Structures of Arp1, pDHG25 and pHELP1:**

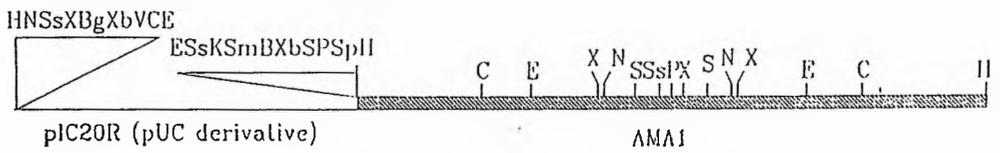


ARp1

E=EcoRI Sm=SmaI Ss=SstI H=HindIII  
 C=ClaI X=XhoI N=NruI S=Sall  
 P=PstI Bg=BglII Xb=XbaI



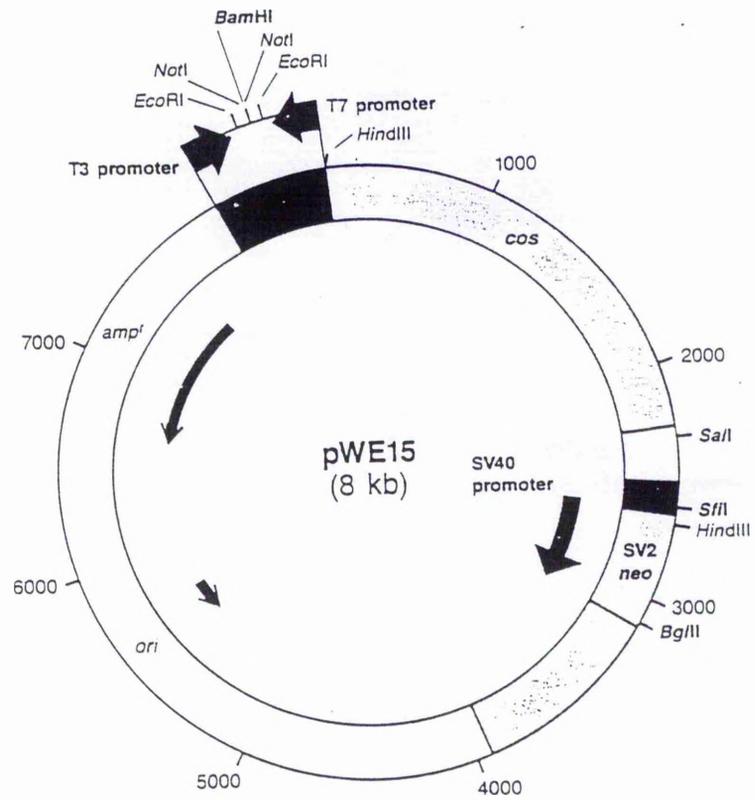
pDHG25



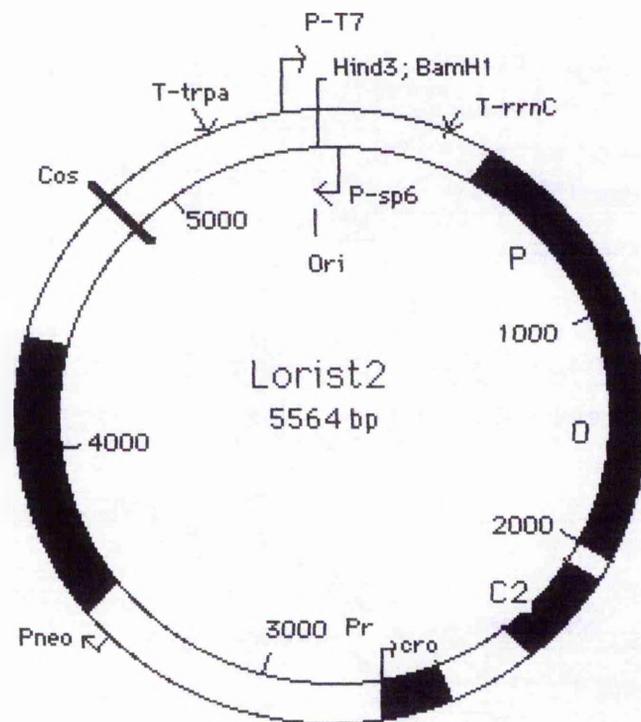
pHELP1

K=KpnI B=BamHI - unique sites  
 V=EcoV

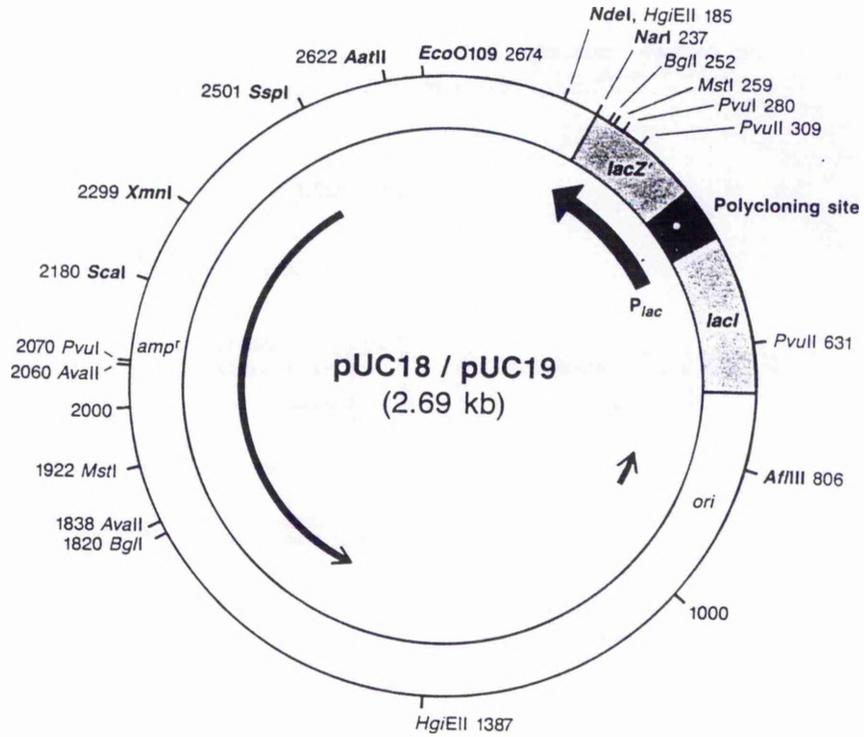
Structure of cosmid pWE15:



Structure of cosmid LORIST2:



## Structure of pUC19:

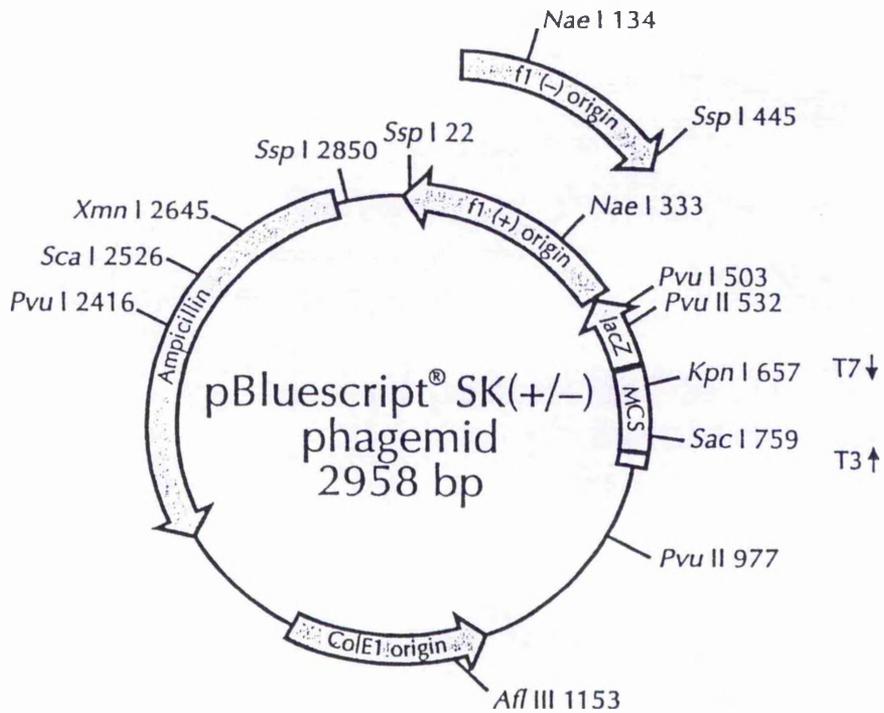
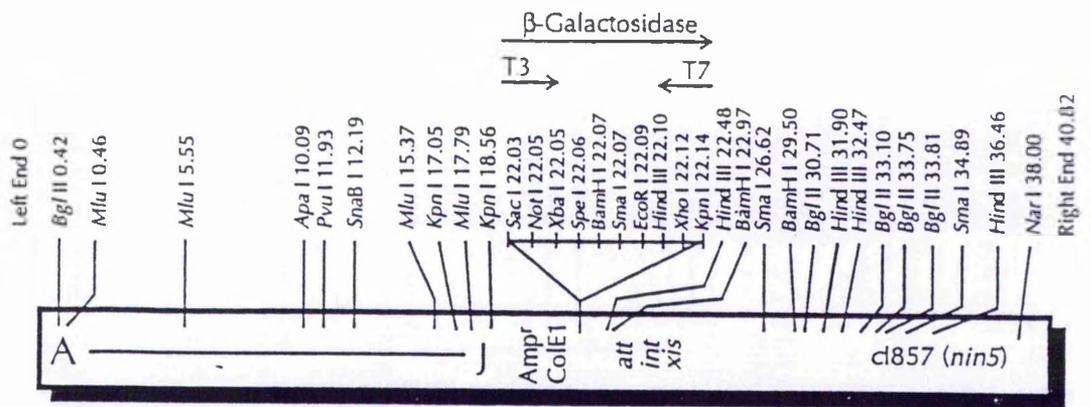


### Polycloning Site

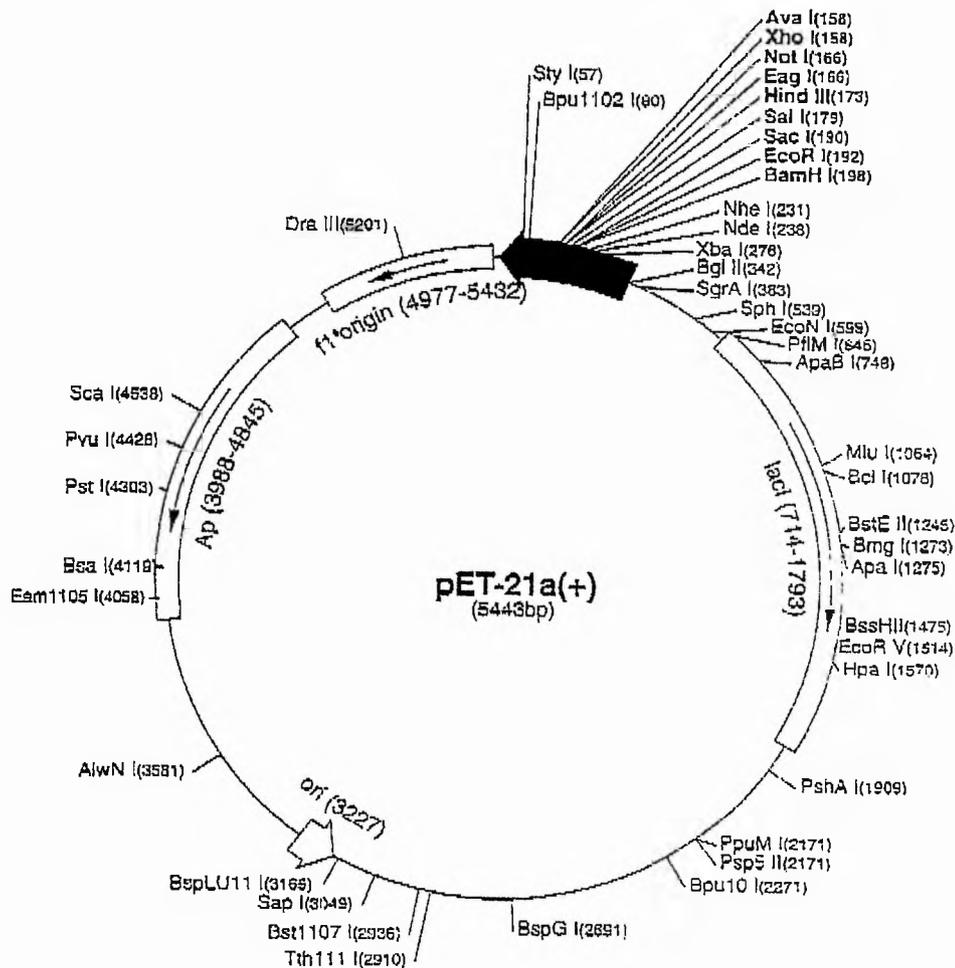
pUC19

1	2	3	4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	5	6	7	8	
Thr	Met	Ile	Thr	Pro	Ser	Leu	His	Ala	Cys	Arg	Ser	Thr	Leu	Glu	Asp	Pro	Arg	Val	Pro	Ser	Ser	Asn	Ser	Leu	Ala	
ATG	ACC	ATG	ATT	ACG	CCA	AGC	TTG	CAT	GCC	TGC	AGG	TCG	ACT	CTA	GAG	GAT	CCC	CGG	GTA	CCG	AGC	TCG	AAT	TCA	CTG	GCC
				<i>HindIII</i>		<i>SphI</i>		<i>PstI</i>		<i>SalI</i> <i>AccI</i> <i>HincII</i>		<i>XbaI</i>		<i>BamHI</i>		<i>SmaI</i> <i>XmaI</i>		<i>KpnI</i>		<i>SacI</i>		<i>EcoRI</i>				

Structure of the Uni-ZAP XR lambda-phage vector and the pBluescript KS+ phagemid:



# Structure of pET21a:



## pET-21a-d(+) Cloning/Expression Region

