

NUCLEAR GENES ENCODING THE ADENINE NUCLEOTIDE TRANSLOCATOR OF  
MAIZE MITOCHONDRIA.

BY

ALISON BAKER

DOCTOR OF PHILOSOPHY,  
UNIVERSITY OF EDINBURGH

1985.



**TABLE OF CONTENTS**

**TABLE OF CONTENTS**

i

**ACKNOWLEDGEMENTS**

vi

**ABSTRACT**

**ABBREVIATIONS**

ix

**CHAPTER1 GENERAL INTRODUCTION**

**1.1 INTRODUCTORY REMARKS**

1

**1.2 MITOCHONDRIAL STRUCTURE**

1

**1.3 MITOCHONDRIAL FUNCTION**

2

1.3.1 Oxidative phosphorylation

2

1.3.2 Unique features of plant mitochondria

4

1.3.3 Intermediary metabolism

5

**1.4 RELATIONSHIP BETWEEN OXIDATIVE AND PHOTOPHOSPHORYLATION IN HIGHER PLANT CELLS**

6

1.4.1 Diurnal regulation of mitochondrial function

6

1.4.2 Developmental regulation of mitochondrial function: germination and early development

7

**1.5 MITOCHONDRIAL BIOGENESIS**

8

1.5.1 Mitochondrially encoded polypeptides

9

1.5.2 Nuclearly encoded mitochondrial polypeptides

11

1.5.3 Nuclear gene copy number

13

1.5.4 Regulation of transcription

14

1.5.5 Translation and import of nuclear gene products

17

1.5.5.1 Is import co- or post translational?

18

1.5.5.2 Import pathways

18

1.5.5.3 Biogenesis of the adenine nucleotide translocator

22

**1.6 CHLOROPLAST BIOGENESIS**

23

**1.7 NUCLEAR-MITOCHONDRIAL INTERACTIONS**

25

1.7.1 Nuclear genes which control mitochondrial gene expression

26

1.7.2 Cytoplasmic male sterility in higher plants

26

**1.8 CONCLUSIONS**

29

**CHAPTER2 MATERIALS AND METHODS**

**2.1 MATERIALS**

31

2.1.1 Seed

31

2.1.2 Enzymes

31

2.1.3 Radiochemicals

31

2.1.4 Film	31
2.1.5 Bacterial strains and plasmids	31
2.1.6 Other materials	32
<u>2.2 STOCK BUFFERS AND MEDIA</u>	32
<u>2.3 GENERAL METHODS</u>	35
2.3.1 Growth of plant material	35
2.3.2 Preparation of nucleic acids	35
2.3.3 Preparation of plasmid DNA	35
2.3.4 Restriction endonuclease digestion of DNA	36
2.3.5 DNA gel electrophoresis	36
2.3.6 Transfer of DNA to nitrocellulose	37
2.3.7 Preparation of radioactively labelled DNA	37
<u>2.4 cDNA SYNTHESIS, CLONING AND CLONE IDENTIFICATION</u>	39
2.4.1 cDNA synthesis	39
2.4.2 S1 nuclease digestion and size fractionation of cDNA	39
2.4.3 Homopolymer tailing	40
2.4.4 Annealing and transformation	40
2.4.5 Ordering and maintenance of the maize cDNA library	41
2.4.6 Colony hybridisation	41
<u>2.5 CONSTRUCTION OF A MAIZE NUCLEAR DNA LIBRARY IN <math>\lambda</math>gtWES</u>	42
2.5.1 Growth of plating cells	42
2.5.2 Plating bacteriophage	42
2.5.3 Isolation of bacteriophage DNA	42
2.5.4 Preparation of $\lambda$ gtWES arms	43
2.5.5 Preparation of extracts for <i>in vitro</i> packaging	43
2.5.6 Extraction of maize nuclear DNA	44
2.5.7 Ligation and <i>in vitro</i> packaging of recombinant phage	46
2.5.8 Screening the library by plaque hybridisation	46
2.5.9 Isolation of small amounts of DNA from recombinant phage	47
2.5.10 Subcloning maize nuclear DNA into pUC plasmids	47
<u>2.6 DNA SEQUENCE ANALYSIS</u>	48
2.6.1 Cloning	48
2.6.2 Manipulation of M13 phage and transformation of <u><i>E.coli</i></u>	49
2.6.3 Preparation of nitrocellulose filter replicas	49
2.6.4 Identification of recombinant M13 clones	49
2.6.4.1 <u>In situ</u> hybridisation	49
2.6.4.2 Ordered grids	49
2.6.5 Plaque purification	50
2.6.6 Sequencing reactions and gels	50
2.6.7 Computer methods	51
<u>2.7 RNA METHODS</u>	51
2.7.1 Denaturing gels	51

2.7.2	Transfer of RNA to nitrocellulose and hybridisation conditions	52
2.7.3	S1 nuclease mapping of RNA molecules	52
2.7.3.1	Preparation of labelled DNA probes	52
2.7.3.2	DNA/RNA hybridisation	53
2.7.3.3	S1 nuclease digestion	53
2.7.3.4	Estimation of the size of S1 protected fragments	54
2.7.4	Primer extension analysis of mRNA 5' termini	54
2.7.4.1	Preparation of the primers	54
2.7.4.2	The primer extension reaction.	55

### **CHAPTER3 CONSTRUCTION OF A cDNA LIBRARY FROM MAIZE POLY A+ RNA AND IDENTIFICATION OF CLONES FOR NUCLEAR ENCODED MITOCHONDRIAL PROTEINS.**

3.1	<u>RATIONALE</u>	57
3.2	<u>SURVEY OF METHODS FOR THE CONSTRUCTION OF cDNA LIBRARIES</u>	57
3.2.1	Methods for synthesising double stranded cDNA	59
3.2.2	Strategies for cloning double stranded cDNA	60
3.2.3	Strategies for screening cDNA libraries	61
3.3	<u>RATIONALE FOR SCREENING THE MAIZE cDNA LIBRARY BY HETEROLOGOUS HYBRIDISATION</u>	63
3.4	<u>CONSTRUCTION OF A MAIZE cDNA LIBRARY</u>	64
3.4.1	Aims	64
3.4.2	Isolation of RNA from maize and synthesis of dsDNA	65
3.4.3	Optimisation of the homopolymer tailing reaction	67
3.4.3.1	Optimising the tailing of inserts	68
3.4.3.2	Optimising the tailing of the vector	68
3.4.3.3	Comparison of the efficiency of different tailing protocols in the construction of chimaeric plasmids and their introduction into <i>E.coli</i> .	71
3.5	<u>CHARACTERISTICS OF THE MAIZE cDNA LIBRARY</u>	73
3.6	<u>SCREENING THE LIBRARY FOR cDNA CLONES ENCODING MITOCHONDRIAL PROTEINS BY HETEROLOGOUS HYBRIDISATION</u>	73
3.7	<u>ISOLATION OF A cDNA CLONE ENCODING THE ADENINE NUCLEOTIDE TRANSLOCATOR OF MAIZE</u>	74
3.7.1	Deriving the hybridisation conditions	75
3.7.2	Identification of a cDNA clone, pANT-1 encoding the translocator	77
3.8	<u>OTHER cDNA CLONES RELATED TO pANT-1</u>	77
3.9	<u>DISCUSSION</u>	85

### **CHAPTER4 THE MITOCHONDRIAL ADENINE NUCLEOTIDE TRANSLOCATOR: ANALYSIS OF CLONES FROM A MAIZE cDNA LIBRARY.**

<u>4.1 RATIONALE</u>	86
<u>4.2 THE MITOCHONDRIAL ADENINE NUCLEOTIDE TRANSLOCATOR</u>	87
4.2.1 Introduction	87
4.2.2 The ANT polypeptide	87
4.2.3 Biogenesis of the adenine nucleotide translocator	89
4.2.4 Substrate specificity	91
4.2.5 Effects of energisation of the membrane on adenine nucleotide transport	91
4.2.6 Inhibitors of adenine nucleotide translocation	92
4.2.7 The mechanism of adenine nucleotide translocation	93
4.2.8 Adenine nucleotide transport and the control of respiration	96
4.2.9 Adenine nucleotide transport in chloroplasts	97
<u>4.3 DNA SEQUENCE ANALYSIS OF cDNA CLONES ENCODING THE ADENINE NUCLEOTIDE TRANSLOCATOR OF MAIZE.</u>	98
4.3.1 Recloning the 1.2kb Pst I fragment of pANT-1 into M13	98
4.3.2 Shotgun cloning restriction fragments of pANT-1 into M13	99
4.3.3 Forced cloning of the ends of the 1.2kb Pst I fragment.	101
<u>4.4 DISCUSSION</u>	104
4.4.1 The nucleotide sequence of pANT-1, a cDNA clone encoding the adenine nucleotide translocator of maize.	104
4.4.2 Amino acid homology with ANT polypeptides from other organisms	108
4.4.3 Nucleotide sequences of other maize ANT cDNA clones	109
4.4.4 RNA transcripts of the ANT genes	113
4.4.5 Codon usage in pANT-1	115
4.4.6 Predicted secondary and higher order structure of the ANT polypeptide.	117
<u>4.5 CONCLUSION</u>	120

## **CHAPTERS ORGANISATION AND EXPRESSION OF THE NUCLEAR GENES ENCODING THE ADENINE NUCLEOTIDE TRANSLOCATOR POLYPEPTIDES IN ZEA MAYS L.**

<u>5.1 RATIONALE</u>	122
<u>5.2 GENOMIC ORGANISATION</u>	122
5.2.1 Analysis of total and nuclear DNA by Southern blotting	122
5.2.2 Do chloroplast and mitochondrial genomes contain sequences homologous to the translocator?	125
<u>5.3 ISOLATION OF ANT GENOMIC CLONES</u>	125
5.3.1 Construction of a maize nuclear genomic library in $\lambda$ gtWES	125
5.3.2 Screening the library for sequences homologous to pANT-1	128
<u>5.4 CHARACTERISATION OF THE ANT GENOMIC CLONES</u>	129
5.4.1 Restriction endonuclease mapping	129
5.4.2 Nucleotide sequence analysis	132

5.4.3 Mapping the 5' point of divergence between the genomic clones	134
<b><u>5.5 COMPARISON OF THE cDNA AND GENOMIC CLONES</u></b>	<b>137</b>
5.5.1 The G2 gene encodes a different polypeptide to that encoded by pANT-1	137
5.5.2 Intervening sequences in the G2 gene	142
5.5.3 The ANT polypeptide of maize may be considerably longer than those of other organisms	145
5.5.4 Potential transcriptional control sequences	148
5.5.5 The 5' untranslated region has some unusual features.	148
5.5.6 Is the leader sequence spliced?	153
5.5.7 The 3' untranslated region	156
<b><u>CHAPTER6 GENERAL DISCUSSION</u></b>	
6.1 AIMS	158
6.2 SUMMARY OF RESULTS	158
6.3 UNRESOLVED QUESTIONS	159
6.3.1 Why are there multiple copies of the translocator gene in maize?	159
6.3.2 Is the maize translocator made as a larger precursor?	162
6.3.2.1 Homology between the G1 and G2 genes	162
6.3.2.2 Comparison of the <u>in vitro</u> synthesised product and mature protein.	162
6.3.3 Where does transcription commence?	163
6.4 FUTURE DIRECTIONS	165
<b><u>BIBLIOGRAPHY</u></b>	<b>167</b>
<b><u>PUBLICATION</u></b>	

## ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the numerous people who have taken an interest in this project. In particular to David Baulcombe (PBI Cambridge) and Udo Weinand (MPI Cologne) for their invaluable practical advice on the construction of cDNA and lambda libraries respectively, and to Mike Douglas (University of Texas Health Centre, Dallas) for the generous gift of the yeast translocator clone pYeOP1(2.6), and for communicating results prior to publication.

In Edinburgh I would like to thank all my friends and colleagues who have made my stay here both intellectually stimulating and socially rewarding. In particular a special thank you to Pete Isaac, our computing wizard, to Steve Smith and Julia Bailey-Serres, who were always willing to offer helpful advice and suggestions, and to Malcolm Hawkesford for two of the figures in this thesis. I am greatly indebted to Chris Leaver, whose interest as a supervisor was always apparent but never obtrusive.

This research was supported by an SERC postgraduate studentship and supplemented by money from the Rennie Bequest and EMBO for travel and subsistence in Cologne.

Finally I would like to dedicate this thesis to my parents in recognition of their constant support and encouragement over many years.

## ABSTRACT.

A cDNA library was constructed from polyA<sup>+</sup> RNA isolated from an inbred maize line (B37N). The library was screened by 'heterologous' hybridisation for cDNA clones representing genes which encode mitochondrial polypeptides. A cDNA clone (pANT-1) encoding the adenine nucleotide translocator (ANT) of maize was identified by virtue of hybridisation to the homologous gene from Saccharomyces cerevisiae, and the identity of the clone confirmed directly by DNA sequence analysis.

Examination of the cDNA library with DNA probes derived from pANT-1 identified two further ANT cDNA clones, one of which, pANT-2, was shown by restriction mapping and partial DNA sequence analysis to be derived from a different structural gene to that of pANT-1. The polypeptides encoded by these two genes differ by eight amino acids. The maize ANT polypeptides exhibit a high degree of amino acid sequence homology to the ANT polypeptides of Saccharomyces cerevisiae, (65%) and Neurospora crassa, (75%), but only 50% amino acid homology with the ANT protein purified from beef heart mitochondria.

A maize nuclear genomic library was constructed in a bacteriophage lambda vector, and two clones were isolated which correspond to the structural genes of pANT-1 and pANT-2. Maize is the only organism which has so far been shown to possess multiple genes for the ANT.

The nucleotide sequence of the structural gene G2, which encodes pANT-2, has been determined, together with ca. 400 nucleotides of sequence 5' to the coding region. The gene contains two introns of 95bp and 82bp which are located in different positions to those in the Neurospora crassa ANT gene. The Saccharomyces cerevisiae gene is uninterrupted. The introns in the maize G2 gene contain the consensus sequences postulated to be important for intron splicing.

The G2 gene has two potential translational starts. If the more 5' ATG



is utilised, the polypeptide is 16 amino acids longer than the ANT polypeptides of Neurospora and yeast and 22 aminoacids longer than that of beef. The implications for targeting the polypeptide to mitochondria are discussed.

There are two TATA like sequences located 5' to the coding region, 283 and 301bp upstream of the more 5' of the two ATG's. The site(s) of transcription initiation have not been accurately mapped, but the 5' untranslated region appears to be unusually long, (ca. 300 nucleotides) and contains several small inverted repeat sequences, the significance of which is unknown. The 5' untranslated region also contains a sequence of 76bp which has 57% homology to an intron found in another nuclearly encoded plant mitochondrial gene, and is delimited by potential intron-exon boundaries

The 3' untranslated region of both maize ANT genes does not contain a 5'-AATAAA-3' sequence, which supports the suggestion that this sequence is not an essential requirement for the polyadenylation of higher plant mRNA's.

**ABBREVIATIONS**

Amp	Ampicillin
Amp <sup>s</sup> , Amp <sup>r</sup>	Ampicillin sensitive, ampicillin resistant.
ANT	Adenine nucleotide translocator
ADP	Adenosine 5' diphosphate
ATP	Adenosine 5' triphosphate
ATR	Atractyloside
BKA	Bongkrekic acid
bp	Base pairs
CAT	Carboxyatractyloside
cDNA	DNA copied from mRNA
CMS	Cytoplasmic male sterility
<u>COB</u>	Structural gene for cytochrome b
<u>COX</u>	Structural gene for cytochrome oxidase subunit
cpm	Counts per minute
ct	Chloroplast
CTAB	Cetyl <sup>+</sup> methyl ammonium bromide.
DNA	deoxyribonucleic acid
DTT	Dithiothreitol
dpm	Disintegrations per minute
ds	Double stranded
EDTA	Ethylene diamine tetra acetic acid (disodium salt unless otherwise stated)
EtBr.	Ethidium bromide
* IPTG	Isopropyl thio beta D-galactoside
kb	kilobase
kD	Kilodalton
MOPS	Morpholinopropane sulphonic acid
mt	Mitochondrial
Mr	Relative molecular weight
NAD	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NaPO <sub>4</sub>	Mixture of equimolar solutions of Na <sub>2</sub> HPO <sub>4</sub> and NaH <sub>2</sub> PO <sub>4</sub> to yield a solution of the desired pH
NaPP	Sodium pyrophosphate
NTPs	Nucleosides (A,G,C,T.) 5' triphosphate
OD <sub>600</sub>	Optical density at 600nm
PEG	Polyethylene glycol
PIPES	Piperazine-NN'-bis-2-ethane sulphonic acid
pfu	Plaque forming units
RF	Replicative form

\* fw<sub>t</sub>

Fresh weight.

RNA	Ribonucleic acid
RUBP	Ribulose biphosphate
rpm	revolutions per minute
SDS	Sodium dodecyl sulphate
ss	Single stranded
TdT	Terminal deoxynucleotidyl transferase
TE80	10mM TrisHCl, 1mM EDTA pH8
Tet	Tetracycline
Tet <sup>s</sup> , Tet <sup>r</sup>	Tetracycline sensitive, tetracycline resistant.
T <sub>m</sub>	Melting temperature (double stranded nucleic acid)
Tris	Tris (hydroxymethyl) amino methane
Triton-X-100	Octylphenoxypolyethoxy ethanol
URF	Unassigned reading frame
(v/v)	Volume per volume (as a percentage)
(w/v)	Weight per volume (as a percentage)
Xgal	5-bromo 4-chloro 3-indolyl beta D-galactoside

## CHAPTER 1 GENERAL INTRODUCTION.

### 1.1 INTRODUCTORY REMARKS

The purpose of this chapter is to set the results described in Chapters 3-5 within the context of the higher plant mitochondrion and its relationship with other cellular and developmental events. To this end, mitochondrial structure and function will be briefly reviewed, with the emphasis on those aspects of plant mitochondria which differ from their animal and fungal counterparts. A comprehensive review of plant mitochondrial biochemistry is outside the scope of this thesis, and is covered in a number of text books on plant biochemistry (e.g. Hanson and Day 1980). However, the import of proteins into mitochondria and chloroplasts and interactions between nucleus and mitochondrion will be discussed.

### 1.2 MITOCHONDRIAL STRUCTURE.

Mitochondria are subcellular organelles present in the cytoplasm of all aerobically respiring eukaryote cells. A 'typical' higher plant cell may contain several hundred rod shaped or spherical mitochondria of the order of 0.5-1  $\mu\text{m}$  x 3  $\mu\text{m}$  (Clowes and Juniper 1968). All mitochondria are bounded by a double membrane which defines four compartments, the matrix, the inner membrane, the inter membrane space, and the outer membrane.

The outer membrane has a number of integral and peripheral proteins associated specifically with it. The most abundant is a protein ('porin'), which forms non-specific pores, rendering the membrane permeable to most small molecules up to ca. 5,000 molecular weight. In addition the outer membrane presumably contains specific receptor proteins which recognise polypeptides imported into the mitochondrion.

The inner membrane is extremely protein rich (ca. 80% protein by mass) and is highly convoluted. The surface area of the inner membrane appears to be correlated with the respiration rate of the mitochondria. The inner membrane is impermeable to most hydrophilic substances, and specific protein carriers or transporters mediate the selective transport of substrates between the mitochondrial matrix and intermembrane space. In addition the inner membrane contains the respiratory complexes involved in electron transport and oxidative phosphorylation.

The matrix space is also protein rich (about 50% by mass), and contains many soluble enzymes including those of the tricarboxylic acid (TCA) cycle (with the notable exception of succinate dehydrogenase, which is membrane bound), and the mitochondrial transcription and translation machinery. The mitochondrial DNA (mt.DNA) is located in the matrix, associated with the inner membrane.

### 1.3 MITOCHONDRIAL FUNCTION.

Mitochondria of higher plants, as those of fungi and animals, carry out synthesis of ATP linked to substrate oxidation, and participate in intermediary metabolism. For a general review see Tzagoloff, (1982). Electron transport in plant mitochondria is reviewed by Palmer (1976).

#### 1.3.1 Oxidative Phosphorylation.

Endogenous NADH generated by the action of the TCA cycle dehydrogenases, is re-oxidised by transfer of electrons to a series of redox carriers ( iron-sulphur proteins, flavoproteins, ubiquinones and cytochromes) located in the inner membrane. These form the electron transport chain and are associated with proteins which make up the respiratory complexes of the inner membrane, Fig 1.1. Transfer of

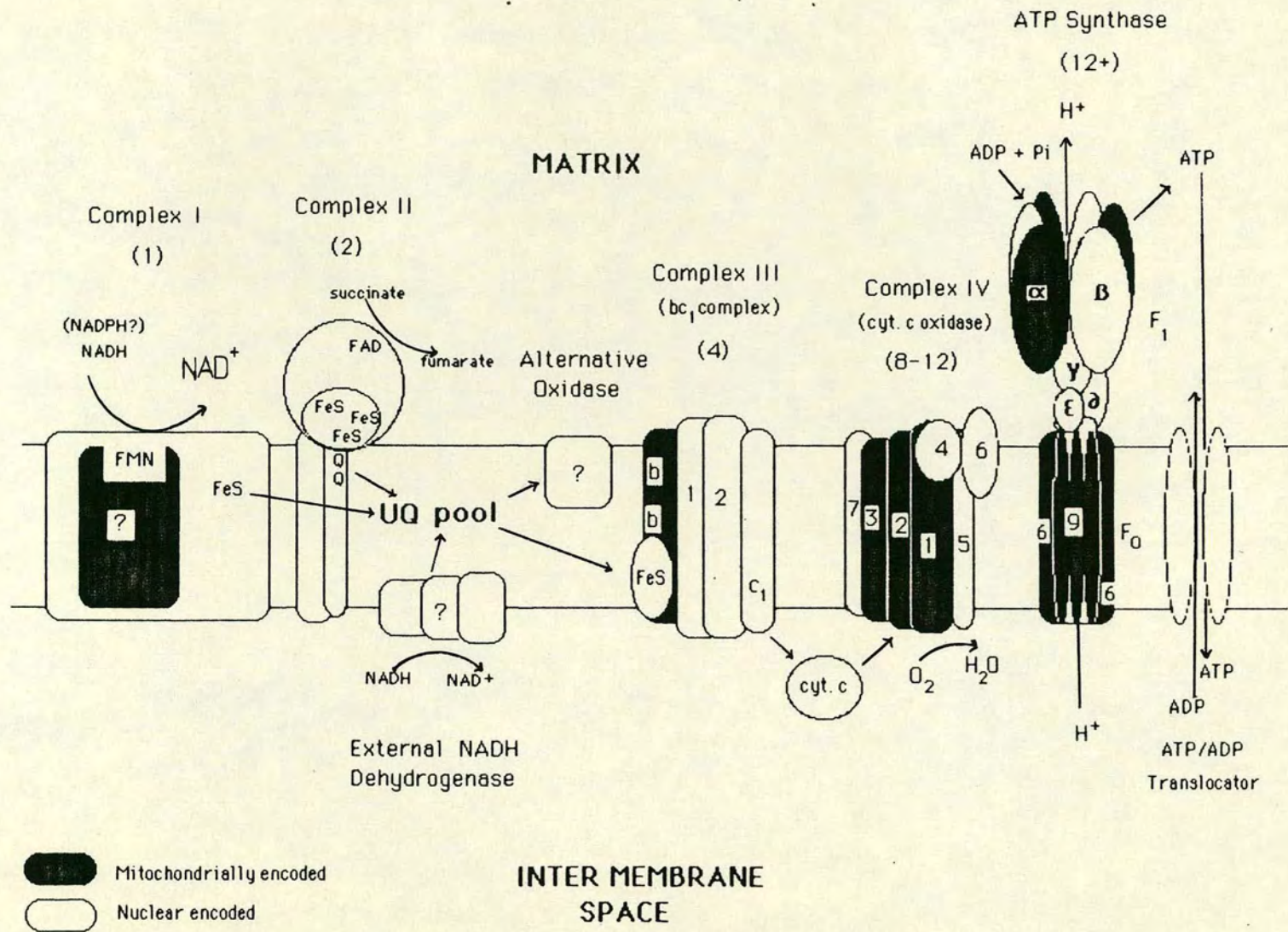


Fig.1.1 Respiratory complexes of the plant mitochondrial inner membrane

M.J.Hawkesford  
 and A.Baker 1985

electrons along the chain to cytochrome  $c$  oxidase proceeds with a decrease in free energy. Electron transport is associated with the pumping of protons from the matrix to the intermembrane space by complexes I, III, and IV, so that energy made available by electron transport is conserved in the form of an electrochemical potential gradient or proton motive force, (pmf) across the inner membrane. Typically the pmf is 180–220mV, with the matrix more negative and alkaline than the cytosol, (Moore and Bonner 1981). In coupled mitochondria, the pmf is utilised to drive synthetic reactions such as the phosphorylation of ADP, accumulation of substrates against their concentration gradients, and the import of cytoplasmically synthesised proteins. These features of mitochondria are common to animal, fungal and plant cells, although the protein subunit composition of the various complexes may vary between species.

### 1.3.2 Unique features of plant mitochondria.

Unlike animal mitochondria, plant mitochondria can oxidise external NADH (Palmer 1976), by means of an external NADH dehydrogenase located on the cytosolic face of the inner membrane (Douce *et al.*, 1973). Electrons from the oxidation of external NADH are fed into the electron transport chain at the ubiquinone pool, by-passing the first coupling site. Plant mitochondria may also be able to transfer reducing equivalents directly from the matrix to external  $\text{NAD}^+$  by means of a uni-directional transhydrogenase (Day and Wiskich 1978).

Mitochondria of higher plants, and some fungi, contain an alternative oxidase activity which, unlike cytochrome  $c$  oxidase, is insensitive to cyanide inhibition, (Solomos 1977). The alternate oxidase branches from the respiratory chain at the ubiquinone pool, by-passing complexes III and IV, two of the sites for coupling electron transport to phosphorylation. The

subunit composition and function is unknown, but it may be important in the oxidation of NADH generated by the action of the TCA cycle in the light. Under these conditions where the energy charge is high, oxidation of NADH coupled to ATP synthesis might inhibit the TCA cycle, the continued operation of which is required to produce carbon skeletons for some biosynthetic pathways, (Singh and Naik 1984).

The molecular biology of plant mitochondria exhibits a number of unusual features. These are summarised in section 1.5.1.

### 1.3.3 Intermediary metabolism.

Plant mitochondria participate in a number of metabolic pathways concerned with both degradation and biosynthesis. Several of these require co-operation with other cell organelles such as chloroplasts, glyoxysomes and peroxisomes, as well as with the cytoplasm. Some examples are:

#### (1) Gluconeogenesis in fatty seeds.

Succinate produced as a result of the glyoxylate cycle in glyoxysomes is further metabolised in the mitochondria to  $\alpha$ ketoglutarate, which is transported back to the glyoxysome to generate oxaloacetate (Cooper and Beevers 1969).

#### (2) Photorespiration in C3 plants.

Glycine, produced in the peroxisome as a result of photorespiration can undergo oxidative decarboxylation to serine, NADH,  $\text{CO}_2$  and  $\text{NH}_4^+$ . Glycine is a major substrate for mesophyll mitochondria in C3 plants, (Moore et al., 1977)

#### (3) Carbon fixation in C4 plants.

Plants which have C4 or Crassulacian Acid Metabolism fix  $\text{CO}_2$  by the action of phosphoenol pyruvate carboxylase, which converts phosphoenol pyruvate to oxaloacetate. This may subsequently be converted to malate in



the case of CAM plants and some C4 plants or aspartate in the case of other C4 plants. In CAM plants, malate is converted to pyruvate which is oxidised to  $\text{CO}_2$  in the mitochondria, while in some C4 plants which synthesise aspartate, the mitochondrial enzymes malate dehydrogenase and NAD malic enzyme are used to regenerate  $\text{CO}_2$  and pyruvate. The  $\text{CO}_2$  released is then fixed by photosynthetic dark reactions in the chloroplast.

#### (4) Amino acid biosynthesis

TCA cycle intermediates are substrates for amino acid biosynthesis, notably alanine (from pyruvate), Glutamate (from  $\alpha$ ketoglutarate) and aspartate (from oxaloacetate).

### 1.4 RELATIONSHIP BETWEEN OXIDATIVE AND PHOTO PHOSPHORYLATION IN HIGHER PLANT CELLS.

Plant cells show both developmental and diurnal changes in the major source of ATP. The first is exemplified by the transition from heterotrophy to autotrophy in early seedling development, and the second by light-dark transitions.

#### 1.4.1 Diurnal regulation of mitochondrial respiration.

The effect of photosynthesis on mitochondrial respiration in green tissues has been reviewed by Graham (1980), Singh and Naik (1984). Measurements suggest that mitochondrial respiration in the light is under tight control of the cytosolic adenylate energy charge (Hampp *et al.*, 1982), which regulates the rate of entry of ADP to the mitochondrion via the adenine nucleotide translocator (ANT). In this way a system which is sensitive and capable of rapidly responding to changes in physiological status is achieved.

#### 1.4.2 Developmental regulation of mitochondrial biogenesis in seed germination.

While diurnal changes in mitochondrial function take place too rapidly to be regulated at the level of gene expression, long term changes must rely to some extent on the de novo synthesis of mitochondrial components. During germination and early seedling development, plants are dependent upon mitochondrial oxidative phosphorylation for the catabolism of stored reserves.

De novo synthesis of mitochondria presents an apparent paradox, as nuclear encoded polypeptides are required, but many of these cannot be imported into the mitochondrion in the absence of a membrane potential (section 1.5.5), the generation of which is also dependent upon nuclear encoded subunits of the respiratory complexes.

Mitochondria from dry seeds are structurally and enzymically deficient. Upon imbibition, an increase in cyanide sensitive oxygen consumption occurs, and mitochondria become structurally more developed. Two plants exhibit extremes of mitochondrial development in early germination. In the case of pea (Pisum sativum) seeds, these changes are observed in the absence of mitochondrial and cytoplasmic protein synthesis, suggesting that upon re-hydration, assembly of pre-formed components takes place to produce respiration competent mitochondria (Morohashi and Bewley 1980). In the contrast, the number of pre-formed mitochondria in the dry seed of peanut (Arachis hypogea) is very low, and most of the early increase in mitochondrial activity is probably due to de novo synthesis of mitochondrial membranes (Morohashi et al., 1981). However, it is unlikely that mitochondria are ever totally absent.

After the initial increase in oxygen consumption upon imbibition, there is a lag period of variable duration before oxygen consumption increases

again. This may be due to temporary anaerobiosis as a result of the restriction of oxygen diffusion by the testa, or a requirement for the synthesis and assembly of new mitochondrial components. Some seeds do not show this lag, e.g. Avena fatua, whereas in others, e.g. pea, the removal of the seed coat reduces but does not abolish the lag, suggesting that oxygen availability is not the only cause.

In a number of plants the capacity of mitochondria to synthesise protein lags behind the development of cytosolic protein synthesis. This may be because the developing mitochondria lack the nuclear encoded proteins required for organellar protein synthesis, e.g. ribosomal proteins, initiation, elongation and termination factors and perhaps also polymerases and RNA maturases. Once some mitochondrial function is established and a proton motive force generated across the inner membrane, nuclear encoded subunits can be imported and the synthesis of new mitochondrial membranes can commence.

Germination is just one example of a developmental situation where there are changing demands on mitochondrial function. Other examples include the greening of etiolated tissue, fruit ripening and pollen formation. The availability of molecular probes should go some way towards resolving the temporal sequence of events which occur during differentiation, as the steady state levels of mRNA and protein for both nuclear and mitochondrially encoded polypeptides can be analysed at different developmental stages and under different physiological conditions.

### 1.5 MITOCHONDRIAL BIOGENESIS.

Mitochondrial biogenesis is a complex process requiring the co-operation of two distinct genetic systems, and is sensitive to a variety

of physiological and developmental factors. Mitochondria contain their own DNA (Luck and Reich 1964), although it only encodes a very small (<10%) but vital subset of mitochondrial proteins. Nuclear genes encode the remainder, including those polypeptides required for replication and expression of mtDNA. The lipid and sterol constituents of the mitochondrial membranes are also the product of cytoplasmic biosynthetic pathways.

### 1.5.1 Mitochondrially encoded polypeptides.

The genetic origin of the different mitochondrial proteins has been determined by a combination of techniques, including differential sensitivity of cytosolic and organelle ribosomes to protein synthesis inhibitors, genetic studies on the inheritance of mutations, and analysis of proteins synthesised in organello by isolated mitochondria. Direct sequence analysis of mtDNA and comparisons with known genes has also been valuable in identifying open reading frames in plant mitochondrial DNA. Table 1a summarises the genes encoded by mtDNA of various species. For a review see Dujon (1983), also Ise et al., (1985) and Chomyn et al., (1985) for references to the NADH dehydrogenase. The presence of ATP 6 in plant mitochondrial DNA is a personal communication from S. Levings. The organisation and expression of the mitochondrial genome in higher plants has been reviewed by Leaver and Gray (1982). The principal differences between plants and other organisms are summarised below.

(a) Plant mitochondrial DNAs are much larger than those of fungal and animal mitochondria. Plant mtDNAs range in size from ca. 200kb to 2,400kb (Ward et al 1981) compared to 16kb for the mammalian mt genome and 78kb for Saccharomyces cerevisiae. Unlike mammalian and fungal mtDNAs, higher plant mtDNA is probably not maintained as a single

circular molecule in vivo, but probably as a number of sub genomic circular and/or linear molecules (Palmer and Shields 1984, Lonsdale et al., 1984). The consequences of this more complex organisation for the expression of plant mitochondrial genes is largely unknown.

Table 1a. Summary of mitochondrially encoded genes in different organisms

<b>GENE</b>	<b><u>S.cerevisiae</u></b>	<b><u>N.crassa</u></b>	<b><u>Mammals</u></b>	<b><u>Plants</u></b>
<u>ATP 6</u>	+	+	+	+
<u>ATP 8</u>	+	+	+	?
<u>ATP 9</u>	+	+/-	-	+
<u>ATP A</u>	-	-	-	+
<u>COB</u>	+	+	+	+
<u>COX I</u>	+	+	+	+
<u>COX II</u>	+	+	+	+
<u>COX III</u>	+	+	+	?
NADH dehydrogenase	-	6 sub- units	6 sub- units	?
Ribosomal prot.	+	+	-	?
Intron coded maturases	+	+	-	?
rRNAs	2	2	2	3
tRNAs	+	+	+	+

---

(b) Plant mitochondrial DNA contains additional genes not present in the mitochondrial DNA of other organisms. The gene encoding the alpha subunit of F<sub>1</sub> ATPase is encoded in mtDNA in higher plants (Isaac et al., 1985b) while in all other eukaryotes examined to date the gene is

nuclearly encoded. The results of in organello protein synthesis by isolated mitochondria suggest that there are other as yet unidentified genes which are unique to plant mitochondria.

(c) The genetic code used by plant mitochondria differs from that of nuclear and other organelle genetic systems. The codon CGG specifies tryptophan as opposed to arginine (Fox and Leaver 1981) and TGA specifies termination, as in the nuclear genetic code, rather than tryptophan as in mammalian mitochondria (Schuster and Brennicke 1985)

### 1.5.2 Nuclear encoded mitochondrial polypeptides.

The remainder of the estimated 400+ mitochondrial polypeptides are the products of nuclear genes, transcribed, processed and modified in the nucleus, translated in the cytosol, and imported into the mitochondrion in a post translational manner.

A number of nuclear genes encoding mitochondrial proteins have been isolated and analysed, but they include only two representatives from a higher plant, the ATP/ADP translocator (Baker and Leaver 1985, and this thesis) and the beta subunit of mitochondrial  $F_1$  ATPase from Nicotiana plumbaginifolia (Boutry and Chua 1985,). Consequently this section will concern itself principally with results obtained from studies with fungi. Yeast is a particularly suitable organism for these kinds of studies, as mitochondrial biogenesis is easily manipulated via the culture conditions and it is particularly suited to the generation and analysis of mutants of mitochondrial function. Furthermore, as a single celled organism, analysis is not additionally complicated by heterogeneity due to the presence of different tissues. It remains to be seen whether regulation of the expression of nuclear genes encoding mitochondrial protein has a common basis in all eukaryotic organisms.

Table 1b summarises the nuclear genes encoding mitochondrial proteins which have been isolated and sequenced.

Table 1b.

<u>ORGANISM</u>	<u>GENE</u>	<u>REFERENCE</u>
<u>S.cerevisiae</u>	iso-1-cytochrome c	Smith <u>et al.</u> ,(1979)
	iso-2-cytochrome c	Montgomery <u>et al.</u> ,(1980)
	cytochrome <u>c</u> <sub>1</sub>	Sadler <u>et al.</u> , (1984)
	cytochrome <u>c</u> peroxidase	Kaput <u>et al.</u> , (1982)
	cytochrome <u>c</u> oxidase IV	Maarse <u>et al.</u> , (1984)
	cytochrome <u>c</u> oxidase Va + b	Cumsky <u>et al.</u> , (1985)
	cytochrome <u>c</u> oxidase VI	Wright <u>et al.</u> , (1984)
	ATP/ADP translocator	Adrian <u>et al.</u> , (1985)
	complex III genes: 14K	deHaan <u>et al.</u> , (1984)
	17K	van Loon,(1984)
	alpha subunit ATPase	M. Douglas pers. comm.
	beta subunit ATPase	Saltzgaber <u>et al.</u> , (1983)
	70kD outer membrane prot.	Hase <u>et al.</u> , (1984)
	EF-Tu (elongation factor)	Nagata <u>et al.</u> ,(1983)
	MSS-51 (splicing enzyme)	Faye and Simon (1983)
	<u>N.crassa</u>	subunit 9 ATPase
ATP/ADP translocator		Arrends and Sebald (1984)
<u>H.sapiens</u>	ornithine transcarbonylase	Horwich <u>et al.</u> , (1984)
<u>B.taurus</u>	cytochrome P450 (SCC)	Morohashi <u>et al.</u> , (1984)
<u>R.rattus</u>	cytochrome <u>c</u>	Scarpulla <u>et al.</u> , (1981)
<u>Z.mays</u>	ATP/ADP translocator	Baker and Leaver (1985)
<u>N.plumbaginifolia</u>	beta subunit ATPase	Boutry and Chua (1985)

A number of other genes have been identified and isolated on the basis of their ability to complement specific mutations, but the nucleotide sequence has not yet been reported, e.g CBP1 and 2 from yeast, required for the processing of apocytochrome b transcripts, (Dieckmann et al., 1982, McGraw and Tzagoloff 1983) and MAS-1, the protease which processes imported proteins, (Yaffe et al., 1985).

### 1.5.3 Nuclear gene copy number.

All the nuclear genes encoding mitochondrial proteins analysed to date occur in single or low copy number. Most of the yeast genes are present only once in the genome, with the exception of cytochrome c and cytochrome oxidase subunit V, (Cumsky et al., 1985), where there are two genes, both of which are expressed. In rat there are about 30 copies of the cytochrome c gene, however most are pseudogenes which represent processed copies of three alternative mRNA's (Scarpulla and Wu 1983).

Other mammalian genomes also contain multiple copies of cytochrome c (Scarpulla et al., 1982). Both of the nuclear genes isolated from higher plants are present more than once. There are two copies of the beta subunit in Nicotiana, both of which are expressed (Boutry and Chua 1985), and two or three genes for the ATP/ADP translocator in maize, of which two are known to be expressed (this thesis).

### 1.5.4 Transcription

The transcription of many nuclear genes encoding mitochondrial proteins in yeast is repressed by glucose or other fermentable carbon sources in the growth media (Periman and Mahler 1974). However, detailed studies concerning the mechanisms of induction and repression have only



been carried out on a small number of these genes.

Cytochrome c. Cytochrome c in yeast is encoded by two non allelic genes, CYC1 and CYC7. The CYC1 gene (iso-1 cytochrome c) encodes 95% of the cytochrome c in wild type yeast. CYC1 expression is repressed by glucose and stimulated by oxygen and haem. Regulation of the expression of the CYC1 gene is primarily at the level of transcription (Zitomer et al 1979). The relative rate of transcription was shown to be six-fold higher in de-repressed cells grown on raffinose, compared to repressed cells grown on glucose, while the half life of the CYC1 mRNA was shown to be similar under both growth conditions. The observed increase in CYC1 transcription in the absence of glucose is sufficient to account for the elevated levels of the protein under these growth conditions.

The nucleotide sequences involved in mediating catabolite and haem regulation of iso-1 cytochrome c have been characterised in greater detail by Guarente and co-workers (Guarente et al., 1984, Guarente and Hoar, 1984, Guarente 1984). The sites of regulation of the CYC1 gene by glucose and haem map 5' to the coding sequence. Two distinct sites are involved, named upstream activation site UAS1 and UAS2, which are located 229 and 265bp 5' to the start of transcription respectively (Guarente et al 1984). These two sites have a core of sequence homology, but also contain unique sequences. In glucose containing media nearly all the transcription is 'driven' by UAS1. Under derepressing conditions, the activity of UAS1 is increased 10 fold while that of UAS2 is increased 100 fold, so that both contribute equally to CYC1 transcription. UAS1 or UAS2 alone when placed 5' to the yeast LEU2 gene is sufficient to confer haem and glucose regulated expression. However, while UAS1 is derepressed by haem or its analogues, UAS2 cannot be derepressed by haem alone, even although haem is required for the basal rate of expression of UAS2 in the presence of

glucose. Activation of UAS1 and UAS2 is dependent upon two transacting non allelic loci, HAP1 and HAP2. A mutant of HAP1, hap1-1, prevents haem activation of UAS1 but does not affect UAS2. Conversely, a mutant of HAP2, hap2-1, causes a decrease in expression from UAS2 but not UAS1. Hence the HAP loci are proposed to encode proteins which act as positive regulatory factors at UAS1 and 2.

The UAS sequences have several features in common with viral and cellular enhancer sequences of higher eukaryotes. They appear to stimulate transcription of specific genes in cis, in response to physiological stimuli, and show a position and orientation independent effect. However, unlike enhancers, the UAS's of CYC1 are inactive when located 3' to the TATA box, (Guarente and Hoar 1984). The mitochondrial Mn superoxide dismutase (Mn SOD) of yeast also has two sequences 5' to the coding sequence which have homology to UAS1 and UAS2 (Marres *et al.*, 1985). Interestingly the Mn SOD is constitutively expressed in a mutant which also exhibits constitutive anaerobic expression of CYC1.

The CYC7 gene which encodes iso-2 cytochrome c, is independently regulated and transcription is not controlled by the intracellular level of haem. Like CYC1 several transacting loci are involved in expression. A deletion analysis of sequences 5' to the gene (Iborra *et al.*, 1985) showed that the region between -350 and -450 is required for negative regulation, and that sequences between -350 and -200 are essential for transcriptional activity. The most striking feature of this latter region is a perfect palindrome located between -309 and -295:

**5'-CCTTCTCTGAGAAGG-3'**

A similar palindrome consisting of a pyrimidine rich sequence followed by a purine rich one is found in the 5' untranslated region of the maize and Neurospora adenine nucleotide translocator genes, (section 5.5.5).

Subunits of complex III. The steady state levels of the mRNA's for seven nuclear encoded subunits of complex III in yeast, change co-ordinately when grown on glucose or a derepressing carbon source (van Loon *et al.*, 1983). However, when individual cloned subunits are introduced into yeast cells on multi copy plasmids, overproduction of the individual subunit is observed (van Loon *et al.*, 1983, van Loon *et al.*, 1984). In the case of the 11kD protein, although mRNA steady state levels are increased 15-30 fold, the steady state level of the protein is only increased 5-10 fold. Translational controls and elevated rate of protein turnover were suggested as a means whereby partial compensation for gene dosage is achieved (van Loon *et al.*, 1983). Similar experiments with the core II (40kD protein) and the Mn superoxide dismutase (originally thought to be the Rieske iron sulphur protein, van Loon *et al.*, 1984, but see Marres *et al.*, 1985), showed that over production of these individual subunits did not affect the levels of the other complex III subunits, and that the over produced subunits were imported into mitochondria. The authors conclude that:

- (1) although the synthesis of the different subunits of complex III is normally tightly coupled, this is not obligatory
- (2) mRNA levels are probably the most important factor in determining protein levels for these subunits
- (3) over production of individual subunits does not markedly influence mitochondrial function and
- (4) that import of the subunits into mitochondria occurs even although the excess subunits are not assembled.

Other yeast nuclear genes. Szekely and Montgomery (1984) showed that the mRNA levels of the alpha and beta subunits of mitochondrial F<sub>1</sub> ATPase and the adenine nucleotide translocator were co-ordinately regulated in

response to glucose. A mutant was described (C149), in which these three genes are insensitive to derepression, but levels of cytochrome  $c$  mRNA were not affected. This suggests that not all nuclear genes for mitochondrial proteins in yeast are regulated by the same mechanism.

#### 1.5.5 Translation and import of nuclear gene products.

The mRNA's which encode the products of nuclear genes destined for the mitochondrion are translated on cytosolic ribosomes, along with the mRNA's encoding proteins for all the other subcellular compartments. The means by which the correct subset of polypeptides are targeted to the appropriate organelle is fundamental to the differentiation of eukaryotic cells. It is clear that the information required for recognition by the appropriate compartment must reside within the polypeptide. Amino acid sequences have been identified which, when fused to heterologous coding sequences, will direct foreign polypeptides to the nucleus (Kalderon *et al.*, 1984), the mitochondrion (Hurt *et al.* 1985 and Horwich *et al.*, 1985), the chloroplast (van den Broeck *et al.*, (1985), Schreier *et al.*, 1985), and the secretory pathway via the endoplasmic reticulum (Walter *et al.*, 1984, Lingappa *et al.*, 1984). However in most cases the molecular mechanisms of protein transport are not known in detail.

The import of proteins into mitochondria has been studied principally in fungal systems, and is comprehensively reviewed by Reid (1984) and Hay *et al.*, (1984). In this section I will briefly summarise the main steps in this process, and seek to draw comparisons with the import of proteins into chloroplasts, which appears to share several features of the mitochondrial import pathway. The biogenesis of the adenine nucleotide translocator will be discussed in more detail.

#### 1.5.5.1 Is import co- or post-translational?

Experiments performed in vitro and in vivo with a variety of organisms have demonstrated that mitochondria import polypeptides post-translationally. Thus there is no obligate coupling of transport to protein synthesis. However, under some circumstances, cytoplasmic ribosomes have been reported associated with the mitochondrial outer membrane, and these are enriched for mRNA's encoding mitochondrial proteins (Ades and Butow 1980). It is unclear whether these results reflect the situation in vivo, i.e. that a proportion of some polypeptides may be imported co-translationally, or whether they are an artifact of the experimental manipulations (Suissa and Schatz 1982).

#### 1.5.5.2 Import pathways

(1) The matrix and inner membrane proteins. Most, but not all, of the polypeptides destined for the matrix and inner membrane are synthesised as precursors which have a slower electrophoretic mobility than the mature protein on SDS-PAGE. This has been demonstrated for polypeptides from a variety of organisms (Maccacchini et al., 1979, Gietl and Hock 1982). Nucleotide sequence analysis of the cloned genes and comparison with the amino terminal sequence of the mature protein has shown that in all cases to date, the additional coding sequence is located at the amino terminus. The pre piece is removed upon or shortly after import into the mitochondrion, but transport is not dependent upon processing (Zwizinski and Neupert 1983). The cleavage is carried out by specific metal ion dependent protease(s) located in the mitochondrial matrix (Bohni et al., 1980, McAda and Douglas 1982). The activity processes all mitochondrial precursors tested but not denatured precursors or cytoplasmic proteins. Despite the high specificity for mitochondrial precursors, the protease(s)

do(es) not show species specificity. A mitochondrial lysate from maize will correctly process yeast precursors, and yeast mitochondria import and process the Neurospora crassa subunit 9 of  $F_0$  ATPase, despite the fact that the homologous protein from yeast is encoded in mtDNA (Schmidt et al., 1983).

These results suggest that the protease recognises a domain in the precursor rather than a specific amino acid sequence. Indeed there is very little amino acid sequence homology around the known cleavage sites, and in some gene fusions where the authentic site is lost, cleavage may take place at a secondary site (Hurt et al., 1984).

The first twelve amino acids of the cytochrome oxidase subunit IV precursor have been shown to be both necessary and sufficient for the re-targeting of a cytoplasmic protein to the mitochondrial matrix. <sup>(Hurt et al., 1985)</sup> In the case of polypeptides synthesised without amino terminal extensions, the targeting information must presumably reside within the mature polypeptide, as is the case with the secreted polypeptide ovalbumin (Lingappa et al., 1978, 1979). Whether some proteins which are synthesised as precursors also contain essential targeting information within the mature protein is uncertain. Experiments by Douglas et al., (1984) which utilised gene fusions between the beta subunit of ATPase and beta galactosidase, suggested that a substantial amount of the mature protein may be required for the targeting of beta galactosidase to the mitochondrion. However, caution must be exercised in the design and interpretation of such gene fusion experiments. Recognition of precursor polypeptides by mitochondria almost certainly involves protein-protein interactions in three dimensions, and the fusing together of heterologous polypeptide sequences could result in the aberrant folding of the hybrid precursor protein such that it is not recognised or recognised less

efficiently by mitochondrial receptors.

Apart from the optional presence of an amino terminal extension, import of all matrix and inner membrane proteins appears to proceed via the same route, with the following sequence of events:

(1) Tight specific and reversible binding to the outer membrane, presumably via receptor proteins (Riezman et al., 1983)

(2) Translocation dependent upon a membrane potential across the inner membrane (Gasser et al., 1982a, Schleyer et al., 1982), and processing where appropriate

(3) Assembly into biologically active form.

It is not known whether these proteins use the same class of receptor or different ones.

(2) Inter membrane space proteins. Proteins destined for the inter membrane space, with the exception of cytochrome c, share part of the import pathway with matrix and inner membrane proteins. These polypeptides, e.g. cytochrome c<sub>1</sub>, cytochrome b<sub>2</sub> and cytochrome c peroxidase are made as precursors and their import is energy dependent. They have quite long and complex presequences, including a region of non polar amino acids capable of spanning a lipid bilayer as an alpha helix (Kaput et al., 1982, Sadler et al., 1984). It has been demonstrated that these precursors are processed in two steps, the first catalysed by a matrix protease, and the second by a less well defined intermembrane space protease (Gasser et al., 1982b, Reid et al., 1982). The intermediate is attached to the inner membrane but with the bulk of the polypeptide projecting into the inter membrane space. Thus it is proposed that transport across the inner membrane is initiated, but arrested by the hydrophobic 'stop transfer' sequence. The matrix protease processes the

amino terminus of the polypeptide projecting into the matrix, and the second protease releases the mature protein to the inter membrane space. In the case of cytochrome  $c_1$ , the final maturation step is dependent upon the covalent attachment of haem (Ohashi et al., 1982).

(3)Outer membrane proteins. Less is known about the biogenesis of the outer membrane in comparison with any other sub mitochondrial compartment. Outer membrane proteins are not made as higher molecular weight precursors, with the exception of rat liver porin (Shore et al., 1981). However, in the case of the 70kD outer membrane protein, the targeting information has been shown to reside at the extreme amino terminus (Hase et al., 1984). Unlike matrix, inner membrane and inter membrane space polypeptides, the import of outer membrane polypeptides is not dependent upon a membrane potential across the inner membrane, (Gasser and Schatz 1983). In yeast it is also insensitive to mild trypsin treatment of the mitochondria, which abolishes import of inner membrane proteins.

(4)Cytochrome c. Cytochrome  $c$  appears to be imported by a different route to all the other proteins examined to date. It is localised on the inner membrane side of the inter membrane space and interacts with cytochrome  $c_1$  in the respiratory chain. The polypeptide is not made as a higher molecular weight precursor, although the apoprotein (lacking haem) was demonstrated to possess a different conformation to the holoprotein (Korb and Neupert 1978). Import of cytochrome  $c$  does not require a membrane potential, and it has a different receptor to the other imported polypeptides (Hennig et al., 1983).



The import of a number of proteins into mitochondria is stimulated by factor(s) present in reticulocyte lysate, (Miura et al., 1983, Ohta and Schatz 1984). The 'factor(s)' are sensitive to trypsin (Ohta and Schatz 1984) and high concentrations of RNase (Firgaira et al., 1984), which suggests that a ribonucleoprotein may be involved.

#### 1.5.5.3 Biogenesis of the adenine nucleotide translocator.

The ATP/ADP translocator is imported by the same pathway as other proteins destined for the inner membrane, but it is one of the few which is not made as a larger precursor (Zimmerman et al., 1979). Some of the smaller subunits of the yeast complex III also apparently lack cleavable N terminal extensions. However, since it has not proved possible to compare the amino acid and gene sequences for the translocator from any one organism, processing at the amino terminus cannot entirely be ruled out.

The translocator is quite a hydrophobic polypeptide, and the precursor can be identified as high molecular weight aggregates in aqueous solution (Zimmerman and Neupert 1980). The precursor and mature polypeptide have different properties which suggest that the protein takes up an alternative conformation upon import. The extra mitochondrial precursor binds to hydroxyapatite, but the mature protein does not. Furthermore, the precursor imported in vitro acquires the properties of the mature protein (Schleyer and Neupert 1984).

A role for the adenine nucleotide translocator in mitochondrial biogenesis can be envisaged. Yeast petite strains, which contain defective mtDNA, are still able to import polypeptides even although they lack mitochondrially synthesised proteins and cannot carry out electron transport. However import is blocked by bongkreikic acid, an inhibitor of the translocator. Thus the translocator may be able to generate a

membrane potential sufficient for the import of proteins, merely by the electrogenic exchange of  $ATP^{4-}$  and  $ADP^{3-}$ , as directed by their concentration gradients. This mechanism could be important in the early stages of mitochondrial biogenesis, before electron transport becomes fully active. However this cannot be essential as the *S. cerevisiae* *opi* mutant, which lacks a functional translocator, is capable of assembling a competent respiratory chain (Kovac et al., 1967)

### 1.6 CHLOROPLAST BIOGENESIS.

Like mitochondria, chloroplasts also have their own genetic system but are dependent upon nuclear gene products to provide many protein components of the organelle. Chloroplast biogenesis has been reviewed by Ellis (1981).

Chloroplasts and mitochondria appear to share many common features in their biogenesis. All the imported chloroplast proteins which have been studied to date are synthesised as higher molecular weight precursors, imported post translationally and processed upon import to the mature size (Cashmore et al., 1985). As with import of proteins into the mitochondrial matrix, the transport of chloroplast precursors into the stroma is energy dependent (Grossman et al., 1980). The import of proteins into the chloroplast envelope has not been reported.

Relatively few imported chloroplast proteins have been studied, and by far the greatest amount of work has been done on the small subunit of ribulose bisphosphate carboxylase. The leader sequence of this polypeptide has been shown to be sufficient to direct the bacterial gene product neomycin phosphotransferase into the chloroplast (van den Broeck et al., 1985, Schreier et al., 1985). The genes encoding four proteins which are imported into chloroplasts have been sequenced, the small subunit of RuBP

carboxylase (Bedbrook et al.,1980, Berry-Lowe et al.,1982, Broglie et al.,1983, Smith et al.,1983 ), the chlorophyll a/b binding protein (Dunsmuir et al. 1983), ferredoxin (Smeekens et al.,1985a), and plastocyanin (Smeekens et al., 1985b) The presequences of these polypeptides do not have any amino acid homology, with the exception of the sequence G<sup>R</sup>/K V near the cleavage site of the first three. However, like the presequences of imported mitochondrial proteins, they have an excess of basic over acidic residues and a relatively high threonine and serine content.

The stromal located protease has some similar features to the mitochondrial matrix protease(s) (Robinson and Ellis 1984). It is specific for chloroplast precursors, and is neither polypeptide or species specific. Both stromal (small subunit) and thylakoid (plastocyanin) polypeptides are substrates. The protease is metal ion dependent as is the mitochondrial enzyme, both being inhibited by 1,10 phenanthroline.

Import of proteins into chloroplasts and mitochondria do differ with respect to the energy source required. Mitochondrial import requires a transmembrane potential, ATP alone will not suffice if the ATPase is inhibited by oligomycin (Gasser et al., 1982a). Chloroplasts however can import proteins in the presence of an uncoupler, so long as ATP is present (Grossman et al.,1980).

As both chloroplasts and mitochondria carry out electron transport and phosphorylation reactions, DNA replication, transcription and protein synthesis, many of their proteins are homologous or analogous. Given the considerable similarities in the processes of protein import by both organelles, and the analogous features of the proteins to be transported, how do polypeptides become associated with the correct organelle? To resolve this question it may be necessary to characterise in much greater detail than has been possible so far, the receptors in the mitochondrial

outer membrane and chloroplast envelope respectively.

### 1.7 NUCLEAR-MITOCHONDRIAL INTERACTIONS

Nuclear gene products are required not only when a net synthesis of mitochondrial material is occurring but continuously, to compensate for turnover of mitochondrial components and to regulate mitochondrial activities. In addition some nuclear gene products may be required at specific stages in the organisms life cycle. In this section I will emphasise the importance of nuclear-mitochondrial interactions, and discuss a number of examples from higher plants and fungi where this relationship appears to be defective.

Nuclear gene products are involved in probably every mitochondrial activity. To those functions considered in section 1.3 can be added

mtDNA replication

recombination

transcription

mRNA processing

mt protein synthesis

metabolite transport

import, processing and assembly of cytoplasmically synthesised polypeptides.

Because of the complex nature of nuclear mitochondrial interactions, nuclear mutants which affect these processes are often pleiotropic and therefore difficult to recognise. Furthermore, because many of them will be expected to severely perturb mitochondrial function they will be lethal or conditional lethal.

### 1.7.1 Nuclear genes controlling mitochondrial gene expression.

In *S. cerevisiae*, there are a number of respiration defective mutants which are inherited in a Mendelian fashion even although they lack mitochondrially encoded gene products such as cytochrome b or cytochrome oxidase subunits I and II (Dieckmann et al., 1982, McGraw and Tzagoloff 1983, Pillar et al., 1983, Pratje et al., 1983, Faye and Simon 1983) Other nuclear gene mutations are able to suppress mitochondrial mutations (Contamine and Bolotin-Fukuhara 1984, Dujardin et al., 1983). There is also an example of a nuclear mutation which blocks the expression of a mitochondrial gene but can itself be suppressed by a mtDNA rearrangement (Muller et al., 1984, Costanzo and Fox 1985).

Most of these mutations define nuclear encoded components of the mitochondrial splicing system. The CBP1 and CBP2 genes are involved in the complex processing of the apocytochrome b transcript (Dieckmann et al., 1982, McGraw and Tzagoloff 1983). The NAM 2-1 mutation suppresses the mitochondrial mutation box7, and restores correct splicing of the COX I and COB transcripts. The NAM 2-1 allele requires the presence of the COX I al4 intron in order to function and is thought to activate a latent maturase encoded by this intron (Dujardin et al., 1983).

### 1.7.2 Cytoplasmic male sterility in higher plants.

In higher plants most mutations affecting nuclear-mitochondrial interactions will probably be lethal due to the inability of plants to survive in the absence of functional mitochondria. Cytoplasmic male sterility (CMS), may represent a class of non lethal mutations which are due to perturbation of 'normal' nuclear- mitochondrial interactions. CMS in higher plants has been reviewed by Leaver and Gray (1982).

Cytoplasmic male sterility is a trait which results in the inability to

produce or shed viable pollen. A number of commercially important plant species exhibit this phenotype and it has been widely exploited in the production of  $F_1$  hybrids in plant breeding. In some cases the earliest observable phenotypic effect is the degeneration of mitochondria in the tapetal layer of the developing anthers, which suggests that the CMS trait may be due to a failure of mitochondrial biogenesis or function. The CMS phenotype is maternally inherited, and is associated with an alteration in mtDNA restriction patterns relative to male fertile lines, and with the synthesis of characteristic variant polypeptides by isolated mitochondria. The CMS phenotype can be 'restored' to fertility by nuclear restorer genes, and this forms the basis of classification of CMS lines into different groups. In this regard the action of restorer genes in suppressing a mitochondrial mutation is analogous to the yeast nuclear suppressors described in section 1.7.1. However, the nature of the CMS mutation(s) remains to be elucidated, as does the mechanism of action of restorer genes.

Recent evidence suggests that crosses which generate incompatible combinations of nuclear and cytoplasmic genotype (i.e. male steriles), result in recombination of the mitochondrial genome. In Sorghum bicolor, the variant polypeptide synthesised by one male sterile line (Kafir nucleus/9E cytoplasm = K/9E) was identified as a larger (42kD vs 38kD) form of cytochrome oxidase subunit I (Dixon and Leaver 1982). Analysis of the COX I gene from male sterile and male fertile sorghum lines (Bailey-Serres et al., in preparation) has shown that synthesis of a larger CO I polypeptide is due to a rearrangement within the coding region which generates an extended open reading frame at the 3' end of the gene. Sorghum lines which synthesise the 42kD CO I variant have the gene located on a 10.4kb Eco R1 generated restriction fragment, as opposed to a

4.3kb Eco R1 fragment in the lines that make the 38kD CO I. The two genes have different characteristic patterns of transcription. The nuclear genotype in Sorghum affects the organisation of the mitochondrial genome. The introduction of 9E genes into the nucleus by backcrossing results in the relocation of the COX I gene on the 4.3kb Eco R1 fragment and the synthesis of the 38kD form of the CO I polypeptide. The points of sequence divergence between the genes encoding the 38kD and 42kD forms of CO I are flanked by small direct repeats. These sequences may be involved in recombination events which alter the genomic environment and expression of the Sorghum COX I gene in different nuclear backgrounds. The synthesis of a 42kD CO I polypeptide is not correlated with CMS, but may be a consequence of recombination events which impair the function of the mitochondrial genome at crucial times in development.

In Zea mays CMS is also associated with rearrangements of the mt genome adjacent to the structural gene sequences COX I, COX II and ATP A (Isaac et al., 1985, Dawson, 1983 PhD. thesis, Isaac et al., 1985 submitted). In one specific form of CMS, S, the COX I gene undergoes rearrangement at a point just 5' to the coding sequence, to generate a series of Bam HI fragments which contain complete copies of the gene (Isaac et al., 1985). These are present in different stoichiometries depending on the nuclear background. Like COX I in Sorghum, the rearrangements adjacent to COX I in CMS maize involve repeated sequences, but there is no alteration in the pattern of transcription or the size of the CO I polypeptide. The repeated sequence represents 186bp of the 208bp terminal inverted repeat of the S plasmids, low molecular weight linear DNAs which are characteristic of S type cytoplasms, and which are known to integrate into the main mitochondrial genome. Integration events are also known to be influenced by the nuclear background. Thus the sequence 5' to COX I may be a target

site for integration or recombination with S plasmid sequences.

The question remains as to why the CMS phenotype only becomes evident at anthesis. In the laboratory, seeds from CMS lines germinate and grow as vigorously as those from fertile lines, and in the case of mitochondria from seedlings of male sterile Sorghum lines, contain a normal cytochrome oxidase activity despite the presence of a larger subunit I. It is possible that mitochondria from CMS lines, although defective in some yet undefined function, are capable of supporting normal growth and development. However, during pollen formation other factors may impose an additional stress upon the mitochondria which they can no longer accommodate. This could be due to an increased demand for cellular ATP, the production of some developmental specific compound to which CMS mitochondria are especially sensitive, or perhaps the unleashing of uncontrolled recombination triggered by some developmental event. The resolution of these questions will probably require the study of mitochondria from developing anthers. Hitherto this has been difficult due to the problems of obtaining sufficient material for in organello protein synthesis, or ethidium bromide staining of mtDNA. However, less material is required for hybridisation analysis and the availability of gene probes will allow an investigation of mitochondrial and nuclear gene expression at the time that the defect becomes apparent.

### 1.8 CONCLUSIONS

The preceding sections have dealt with the central role of mitochondria in cellular metabolism, and the complex interplay between organelles necessary for their biogenesis and continued function. The disruption of intracellular communication may have drastic consequences for the whole organism, as with male sterility. As yet comparatively little



is known about the molecular basis of these interactions between organelles, and most of the data available are from fungal systems. While mitochondria of higher plants are in many respects similar to those of their animal and fungal counterparts, they differ in detail, and have some unique biochemical activities. Furthermore the development and physiology of plants is quite different, and consequently the demands and constraints on their mitochondria may be subtly different. One major difference between plant and animal cells is the presence in the former of another semi autonomous organelle, the chloroplast. Like mitochondria, chloroplasts have their own genetic system, but are dependent on nuclear gene products to provide many of the protein components of the organelle. This would seem to add a further level of complexity to the problem of cyto-differentiation in higher plants compared to fungi and mammals.

The work presented in this thesis is concerned with the isolation and analysis of nuclear genes encoding a mitochondrial polypeptide from a higher plant, Zea mays L. This gene, encoding the adenine nucleotide translocator, is one of the few constitutively expressed genes to be isolated from a higher plant. The essential role of this protein in mitochondrial metabolism (section 4.2) means that expression of the translocator gene can be used as a marker of mitochondrial biogenesis, and the availability of specific gene probes will form the basis for future experiments aimed at understanding the mechanism and co-ordination of mitochondrial biogenesis in higher plants.

## CHAPTER2 MATERIALS AND METHODS

### 2.1 MATERIALS

#### 2.1.1 Seed.

Maize (Zea mays L.) seed, nuclear genotype B37, with the N (fertile) cytoplasm, was kindly supplied by Pioneer Hi-bred International, Des Moines, Iowa, USA.

#### 2.1.2 Enzymes.

Restriction enzymes were purchased from Boehringer Mannheim, Bethesda Research Laboratories, Amersham International plc., P&S Biochemicals, NBL and BCL.

Klenow fragment of Escherichia coli DNA polymerase I was from Boehringer or Amersham. Calf Thymus Terminal Deoxynucleotidyl Transferase was from Amersham, T4 DNA ligase from Boehringer and Reverse Transcriptase from Life Sciences Inc.

#### 2.1.3 Radiochemicals.

Alpha  $^{32}\text{P}$  dCTP (410 Ci/ mmol) in aqueous solution, and  $^{35}\text{S}$  Methionine (1040 Ci/ mmol) were purchased from Amersham International plc.

#### 2.1.4 Film

X ray film was Dupont Cronex 4 , Fuji RX (sequencing gels) or Kodak X-omat A.R. for genomic Southern.

#### 2.1.5 Bacterial Strains and Plasmids.

E.coli strains:

294 (  $\text{Pro}^- \text{End}^- \text{A}^- \text{r}_k^- \text{m}_k^- \text{sulIII}^+ \text{Thi}^-$ ) (Smith et al., 1983)

JM101 ( $\Delta(\text{lac pro}), \text{thi}, \text{Sup E}, \text{F}'\text{traD36}, \text{proAB lac i}^q \text{z}$ )

(Messing 1979).

BHB2600 [803  $\text{Sup E}^+ \text{Sup F}^+ \text{r}_k^- \text{m}_k^- \text{met}^-$ ] Hohn (1979)

ED8654 ( $\text{met}^- \text{supE supF hsd R}^- \text{m}^+ \text{s}^+ \text{trpR red}^+$ )

BHB 2688 [N205 rec A<sup>-</sup> ( imm434 cIts<sub>b</sub>2 red3 Eam4 Sam7 $\lambda$ )]  
Hohn (1979)

BHB2690 [N205 recA<sup>-</sup> ( imm 434 cIts<sub>b</sub>2 red3 Dam15 Sam7 $\lambda$  )]  
Hohn (1979)

Plasmid vectors:

pAT153 Amp<sup>r</sup>Tet<sup>r</sup> Twigg and Sherratt (1980)

pUC 9 Amp<sup>r</sup> Messing and Vieira (1982a)

Bacteriophage vectors:

$\lambda$ gtWES. $\lambda$ B. Leder *et al.*, (1977)

M13 mp8 Messing and Vieira (1982b)

mp11

mp18 Yanisch-Perron *et al.*, (1985).

mp19 " "

### 2.1.6 Other materials

Nitrocellulose was purchased from Schleicher and Schull. Sigma type II agarose was used for genomic Southern. Hybond N nylon membranes were purchased from Amersham and used for Southern blots according to the manufacturer's instructions.

## 2.2 STOCK BUFFERS AND MEDIA.

All solutions and reagents were stored at room temperature unless otherwise specified.

Bacterial and bacteriophage media.

L. Broth 10g/l Difco Bactotryptone, 5g/l Difco Bacto yeast extract, 10g/l NaCl pH7.2

LB Agar L Broth plus 15 g/l Difco Bacto agar.

CY Medium 10 g/l casamino acids, 5 g/l Difco Bacto yeast extract, 3 g/l NaCl, 2 g/l KCl pH7. Autoclave then add 25ml 1M Tris HCl pH7.4 and 5ml 1M  $MgSO_4$  per litre.

CY Agar CY medium plus 15 g/l Difco Bacto agar.

Top Agarose 10 g/l Difco Bacto tryptone, 5 g/l NaCl, 7 g/l agarose.

BBL top layer 10 g/l Baltimore Biological Laboratories trypticase, 5 g/l NaCl, 6.5 g/l Difco Bacto agar.

Minimal Plates 15 g/l Difco Bacto agar, 2 g/l  $(NH_4)_2SO_4$ , 14 g/l  $K_2 HPO_4$ , 6 g/l  $KH_2PO_4$ , 1 g/l  $Na_3$ citrate, 0.2 g/l  $Mg SO_4$ , 0.2% glucose, 1mg/l Thiamine hydrochloride.

LTB 20mM Tris HCl pH8, 20mM NaCl, 1mM EDTA.

IPTG 24 mg/ml dissolved in dimethyl formamide. Stored at  $-20^{\circ}C$

XGal 20 mg/ml dissolved in dimethyl formamide " "

#### Antibiotics

Tetracycline 10mg/ml stock solution in ethanol. Stored at  $-20^{\circ}C$

Ampicillin 50 mg/ml stock suspension in sterile double distilled water. Stored at  $-20^{\circ}C$ .

Chloramphenicol Added as a solid to bulk cultures, 150 ug/ml. Stored at  $4^{\circ}C$ .

#### Electrophoresis buffers

10X TBE 890mM TrisHCl, 890mM boric acid, 25mM EDTA pH8.3.

10X TAE 400mM TrisHCl, 20mM EDTA, 200mM Na acetate pH8

10XMOPS 250mM MOPS, 50mM Na acetate, 10mM EDTA pH7.

#### Hybridisation buffers

20xSSC 3M NaCl, 0.3M sodium citrate pH7.

20XSSPE 200mM  $NaH_2PO_4$ , 3M NaCl, 20mM EDTA pH7.

20XSET 3M NaCl, 0.6M TrisHCl, 20mM EDTA pH8.

100X Denhardt's solution 2% BSA, 2% Ficoll, 2% Polyvinyl pyrrolidine.  
(stored at  $-20^{\circ}\text{C}$ ).

Southern transfer solutions.

Denaturing 0.5M NaOH, 1.5M NaCl

Neutralising 0.5M TrisHCl, 3M NaCl pH7

Restriction enzyme buffers.

10X High 1M NaCl, 500mM TrisHCl pH7.5, 100mM  $\text{MgCl}_2$ , 10mM DTT.

10x Medium 500mM NaCl, 500mM TrisHCl pH7.5, 100mM  $\text{MgCl}_2$ , 10mM DTT.

10x Low 500mM TrisHCl pH7.5, 100mM  $\text{MgCl}_2$ , 10mM DTT.

10X HindIII 600mM NaCl, 100mM TrisHCl pH7.5, 70mM  $\text{MgCl}_2$ .

10X EcoRI 500mM NaCl, 100mM TrisHCl pH7.5, 100mM  $\text{MgCl}_2$ .

10X Sma I 200mM KCl, 100mM TrisHCl pH8, 100mM  $\text{MgCl}_2$ , 10mM DTT.

10X Ligation 500mM Tris HCl pH7.5, 100mM  $\text{MgCl}_2$ , 100mM DTT, 5mM ATP.

Honda medium.

2.5% Ficoll, 5.0% Dextran T40, 3mM  $\text{MgCl}_2$ , 25mM TrisHCl pH7.1, 1mM

2-mercaptoethanol.

## 2.3 GENERAL METHODS.

### 2.3.1 Growth of plant material.

Maize seed was <sup>surface</sup>sterilised by immersion in a 1:15 dilution of sodium hypochlorite (stock=11-14% available chlorine) for 10 min. and imbibed in running tap water overnight. The imbibed seed was germinated on cellulose wadding saturated with 0.2mM CaCl<sub>2</sub> at 28-29°C in total darkness for 60h.

### 2.3.2 Preparation of nucleic acids.

Coleoptiles were harvested directly into liquid N<sub>2</sub> or onto aluminium foil spread on dri-ice, and either stored at -80°C or extracted immediately. Total nucleic acids were extracted by the method of Leaver and Ingle (1971), and RNA by the method of Chirgwin *et al.*, (1979).

Total nucleic acids were fractionated into RNA and DNA by two cycles of salt precipitation with 2.5M NaCl according to Haffner *et al.*, (1978). Poly A<sup>+</sup> containing RNA was recovered from total RNA by elution from an oligo dT cellulose column (Aviv and Leder 1972). "Poly A<sup>+</sup> enriched" RNA is the bound fraction eluted from the column by 10mM Tris-HCl pH7.5, and contains ca. 10% rRNA as judged by non-denaturing agarose gel electrophoresis. To obtain "Poly A<sup>+</sup>" RNA, the enriched fraction was reapplied to the column and eluted as before.

### 2.3.3 Preparation of plasmid DNA.

Crude nucleic acid extracts of *E.coli* strains harbouring recombinant plasmids were prepared by lysing the cells with hot SDS as described by Barnes (1977).

Small scale plasmid DNA preparations were obtained by the alkaline SDS lysis method of Birnboim and Doly (1979). For large scale isolation of

plasmid DNA the lysozyme/triton lysis method of Katz et al., (1973) was employed. Bacteria were grown at 37°C in L. Broth supplemented with 1% (w/v) glucose and the appropriate antibiotic until an OD<sub>600</sub> of 1.0 was reached. Chlor amphenicol was added to a final concentration of 150ug/ml and the culture amplified overnight. Cleared lysate was prepared according to Katz et al., and the DNA purified in CsCl/ EtBr density gradients (Radloff et al. ,1967).

#### 2.3.4 Restriction endonuclease digestion of DNA.

Plasmid DNA was digested at a concentration of 10-100ng/ul in volumes of 10-30ul, with 2-3U of enzyme per ug of DNA. Low, medium and high salt buffer were used for most enzymes, according to Maniatis et al., (1982), with the exception of Sma I, which requires KCl. Digestions were performed at the recommended temperature for 1-4h., and terminated by heating at 65°C for 10min. or by phenol extraction.

Nuclear DNA was digested in larger volumes, typically 5-10ug in 100ul with 5U of enzyme per ug of DNA. With both Eco RI and HindIII, more complete digestion was observed using buffers recommended by the manufacturers, as opposed to medium salt buffer.

#### 2.3.5 DNA gel electrophoresis.

Flat-bed agarose slab gels were run in TAE buffer. DNA fragments >10kb were resolved on 0.8% (w/v) agarose gels, 1-10kb on 1% gels and <1kb on 1.5% gels. Vertical acrylamide gels (19:1 acrylamide:bis) prepared according to Dingman and Peacock (1968) were used to resolve small fragments of DNA. Gels were stained in 1ug/ml EtBr for 15-30 min., destained in distilled water, and photographed under short wave UV light (254nm, Ultraviolet products Inc.) The film was Kodak 2451 technical pan

or Ilford HP5.

### 2.3.6 Transfer of DNA to nitrocellulose.

DNA was transferred to nitrocellulose by the method of Southern (1975). In order to facilitate transfer of large fragments of DNA, gels were soaked in 0.25N HCl for 5min prior to denaturation. Denaturation and neutralisation were for 40-60min each. Plasmid DNA in gels treated in this manner could be efficiently transferred to nitrocellulose in 2-3h. Gels containing digests of high molecular weight nuclear DNA were allowed to transfer overnight. In some instances it was useful to obtain two identical filters from the same gel. This was achieved by means of the "double dry blot" technique, where the gel is sandwiched between two pieces of nitrocellulose (Palmer et al., 1982).

### 2.3.7 Preparation of radioactively labelled DNA.

#### a) Nick-translation.

Nick-translation of DNA (Rigby et al., 1977) was carried out as described by Maniatis et al., (1982), except that a typical reaction was 30ul in volume, and contained 0.1-0.5ug of DNA. Unincorporated nucleotides were removed by centrifugation of the reaction mixture through a 1ml column of Sephadex G-50 (Maniatis et al. (1982). Using 10-20 uCi alpha <sup>32</sup>P dCTP (410Ci/mmol), 20-50% of the label could be incorporated, depending on the DNA preparation, and specific activities of  $5 \times 10^7$  cpm/ug could be routinely obtained.

#### b) Second strand synthesis.

Strand-specific probes were obtained by synthesising a radioactive complementary strand to single stranded M13 clones of known sequence. Both sequencing (Messing et al., 1981) and probe primers (Hu and Messing



1982) were used to prime second strand synthesis by AMV Reverse Transcriptase. Primer (2.5ng) was annealed to the template (ca. 150ng) by incubating at 60°C for 20-30 min. in 10mM MgCl<sub>2</sub>, 10mM TrisHCl pH8.5 in a volume of 10ul. The annealing mix was allowed to cool slowly to 42°C and centrifuged briefly. 1.5ul of 10x RT salts (500mM TrisHCl pH8.3, 200mM KCl, 80mM MgCl<sub>2</sub>, 100mM DTT), 1ul each of 1mM dATP, dGTP and dTTP, 10-30uCi alpha <sup>32</sup>P dCTP and 8U AMV Reverse Transcriptase were added, and the reaction incubated at 42°C for 1-2h. Unincorporated nucleotides were removed as previously described (Section 2.3.7a).

c) Labelling restriction fragments with random hexanucleotide primers

In order to synthesise probes of very high specific activity, the random primer method of Feinberg and Vogelstein (1984) was used. DNA was digested with the appropriate restriction endonuclease(s) and separated on a 1% (w/v) low gelling temperature agarose gel in 1XTBE. The gel was stained in 1ug/ml EtBr for 10 min., and the desired band excised in the minimum volume of agarose. The agarose containing the DNA was weighed and double distilled water added in the ratio 3ml per g agarose. The DNA/agarose was denatured by boiling for 7 min and labelled exactly as described by Feinberg and Vogelstein (1984). With this protocol probes with specific activities in excess of 10<sup>9</sup>cpm /ug could be obtained.

d) Preparation of end labelled DNA fragments for molecular weight markers

Plasmid DNA was digested with a suitable restriction endonuclease (one leaving a 5' protruding end) and the fragments labelled by filling in the 3'recessed end with reverse transcriptase. Digests were phenol extracted, ethanol precipitated and resuspended in 1x RT buffer (section 2.3.7b). Nucleotides, enzyme and label were added and the reaction incubated as for second strand synthesis.

## 2.4 cDNA SYNTHESIS, CLONING, AND CLONE IDENTIFICATION.

### 2.4.1 cDNA synthesis.

The method used was essentially that of Efstradiatis *et al.*, (1976), as modified by Buell *et al.*, (1978), which relies on the formation of a hairpin loop at the 5' end of the first strand to prime second strand synthesis, followed by S1 nuclease digestion of the double stranded DNA products.

First strand reaction contained 4ug poly A<sup>+</sup> RNA per 50ul incubation mix (50mM TrisHCl pH8.3, 20mM KCl, 8mM MgCl<sub>2</sub>, 10mM DTT, 1mM each dATP, dTTP, dGTP, dCTP, 150ng oligo dT) 5-10uCi alpha <sup>32</sup>PdCTP was added to facilitate detection of the product. The incubation also contained 30U placental RNase inhibitor and 30U AMV Reverse Transcriptase. The reaction mix was incubated 10 min on ice then 60 min at 42<sup>0</sup>C. After boiling for 2 min followed by rapid cooling on ice, 50ul 2x second strand buffer (200mM HEPES-KOH, 2mM each dNTP, 6mM KCl) and 15U of the Klenow fragment of DNA polymerase 1 of *E.coli* were added and the reaction incubated at 15<sup>0</sup>C for 5h. Unincorporated nucleotides were removed by gel filtration through Sephadex G-50. The size of the 1st and 2nd strands was estimated by electrophoresis of glyoxal treated aliquots of the reaction through 1%(w/v) agarose gels in 10mM NaPO<sub>4</sub> (section 2.7.1).

### 2.4.2 S1 nuclease digestion and size fractionation of cDNA.

The double stranded cDNA was resuspended in 200ul of S1 digestion buffer (0.3M NaCl, 0.03M NaOAc pH4.5, 4.5mM ZnSO<sub>4</sub>, 5% (v/v) glycerol) and incubated with 400U S1 nuclease at 37<sup>0</sup>C for 60min. followed by 10 min at 10<sup>0</sup>C.

S1 digested cDNA was size fractionated by centrifugation through sucrose gradients. The gradients were 12.5%(w/v) sucrose dissolved in 150mM NaCl, 10mM TrisHCl pH8, 1mM Na<sub>2</sub>EDTA, and were formed by freezing 5ml aliquots of this solution in 5ml polyallomer centrifuge tubes and allowing them to thaw for 30 min. at 37<sup>0</sup>C. After S1 digestion the cDNA was ethanol precipitated, the pellet resuspended in 100ul of TE80 and applied to the top of the gradient, which was centrifuged at 50Krpm for 4h. at 20<sup>0</sup>C in the Beckmann SW50.1 rotor. The gradients were fractionated by displacement through a hole in the bottom of the tube and the fractions Cerenkov counted to localise the labelled cDNA. The size of each fraction was estimated by running a small aliquot on a 1.5%(w/v) agarose mini gel in TAE. The cDNA recovered from the gradient was resuspended in sterile double distilled water for homopolymer tailing.

#### 2.4.3 Homopolymer tailing.

Optimal conditions for the tailing reaction (section 3.4.3) were found to be 100ng double stranded cDNA, 100mM sodium cacodylate, 1mM dCTP, 0.2mM DTT, 1mM CoCl<sub>2</sub>, 10U calf thymus terminal transferase, incubated at 37<sup>0</sup>C for 10min. Vector tailing was performed as above except that 2ug Pst 1 digested pAT153 and 1mM dGTP replaced the cDNA and dCTP, and the incubation was carried out at 37<sup>0</sup>C for 5 min. Reactions were terminated by the addition of 2ul of 0.5M Na<sub>2</sub>EDTA pH8.

#### 2.4.4 Annealing and Transformation.

1ul of cDNA from the tailing reaction and up to 5ul from the vector tailing reaction were added directly to 50ul of 100mM NaCl, 10mM TrisHCl pH8, 1mM Na<sub>2</sub>EDTA. The annealing mix was incubated at 65<sup>0</sup>C for 10min and

allowed to cool overnight to 4°C. Competent cells of E. coli strain 294 were prepared and transformed by the procedure of Dagert and Ehrlich (1979). The transformation mixes were plated on L-broth plates containing 100ug/ml ampicillin.

#### 2.4.5 Ordering and maintenance of the maize cDNA library.

Amp<sup>S</sup> Tet<sup>r</sup> colonies were tooth picked onto 6x8 arrays on L Broth plates containing 10ug/ml tetracycline. A permanent replica of the library was frozen in microtitre dishes at -80°C as described by Gergen et al.,(1979). The library contains ca. 2,500 colonies.

#### 2.4.6 Colony hybridisation.

Filter replicas of the cDNA library were prepared on Whatmann 541 paper as described by Gergen et al.,(1979). Filters (7x5cm) were probed under 'heterologous' conditions. Prehybridisation was performed in 50%(v/v) formamide, 5x SSC pH7, 150ug/ml denatured herring sperm DNA in a total volume of 4ml per filter on a rocking table at 25°C for 2-16h. Hybridisation was for 48h. at the same temperature in fresh buffer containing the denatured radioactively labelled yeast probe. Filters were washed 4x30min. in 2xSSC at room temperature, air dried and exposed to pre-flashed X-ray film at -80°C, with intensifying screens for up to three weeks. (Laskey and Mills 1977). Colony hybridisations using cloned maize DNA fragments were performed in 4xSSC/ 10xDenhardt's solution/ 0.1%(w/v)SDS/ 150ug/ml denatured herring sperm DNA at 65°C, and washed in 2xSSC 2x30min. at room temperature.

## 2.5 CONSTRUCTION OF A MAIZE NUCLEAR DNA LIBRARY IN $\lambda$ gtWES

### 2.5.1 Growth of plating cells

A single colony of BHB2600 was used to inoculate a 5ml L Broth 'overnight' and grown with shaking at 37°C to stationary phase. This was used to inoculate 200ml L Broth containing 0.1% maltose and grown overnight 37°C with shaking. Cells were harvested by centrifugation, 5Krpm x 5 min. and resuspended in 0.01 vol. 10mM MgSO<sub>4</sub> (100X plating cells). Alternatively cells were grown to an OD<sub>600</sub> of 0.5, centrifuged, and resuspended in 0.5 vol of 10mM MgSO<sub>4</sub> (2X plating cells).

### 2.5.2 Plating bacteriophage

Serial dilutions of phage in 10mM MgSO<sub>4</sub> were adsorbed to 100ul of 2X plating cells in thin walled glass tubes for 15min at 29°C. Five ml aliquots of molten BBL top agar (cooled to 50°C) were added and the contents of each tube plated onto dry 90mm L-broth plates. The plates were incubated for 8-14h at 37°C.

### 2.5.3 Isolation of bacteriophage DNA

$\lambda$ gtWES DNA was prepared from liquid cultures. BHB2600 cells were grown to an OD<sub>600</sub> of 0.4 in 100ml L Broth containing 10mM MgSO<sub>4</sub>. Phage were added to a multiplicity of infection (m.o.i.) of 1, (ca. 4x10<sup>9</sup> pfu) and the culture incubated at 37°C with vigorous shaking until lysis was complete (ca. 4h.). At this point the culture was no longer turbid, and aggregates of bacterial debris were visible. 0.2ml of chloroform were added and the culture incubated with shaking for a further 25 min. The culture was cleared by centrifugation 10Krpm x 10 min in the Sorvall GSA

rotor. Lambda DNA was purified by the PEG precipitation method of Yamamoto et al., (1970).

#### 2.5.4 Preparation of $\lambda$ gtWES arms

0.5ml of the phage suspension removed from the CsCl equilibrium gradient was diluted 10 fold with 10mM Tris HCl pH8/ 0.1% (w/v)SDS. 1mg of proteinase K was added and the phage incubated for 20min at 37<sup>0</sup>C, followed by 10 min at 50<sup>0</sup>C.. The digest was extracted with an equal volume of 1:1(v/v) phenol/chloroform and the organic layer back extracted with 4ml of TE80. The pooled aqueous phases were re-extracted with 1:1(v/v) phenol/chloroform, and ethanol precipitated.

150ug of DNA was digested to completion with 150U of restriction endonuclease EcoR1. Digests were heated to 65<sup>0</sup>C for 5min , chilled in ice water, and loaded directly onto sucrose gradients. The gradients were 11.5 ml, linear 5-30% sucrose in 1M NaCl, 20mM TrisHCl pH8, 10mM Na<sub>2</sub>EDTA, and were poured in polyallomer tubes. The DNA was applied to the top of the gradient in a volume of 500ul, 75ug of DNA per gradient, and centrifuged for 16h. at 28Krpm and 15<sup>0</sup>C in the Beckmann SW41 rotor. The gradients were fractionated into 0.5ml aliquots and those fractions containing the arms identified by agarose gel electrophoresis of small aliquots. The desired fractions were pooled and ethanol precipitated.

#### 2.5.5 Preparation of extracts for in vitro packaging

E.coli strains BHB2688 and BHB2690 are temperature sensitive lysogens, and act as 'head' and 'tail' donors (Hohn 1979). Strains were checked for inability to grow at 37<sup>0</sup>C.

Single colonies were streaked onto CY plates and grown overnight at 30<sup>0</sup>C. Bacteria from one plate were resuspended in 5ml L Broth which was

used to inoculate 500ml CY medium. 1.5 litres of BHB2688 and 500ml of BHB2690 were incubated at 32°C with shaking, until the OD<sub>600</sub> reached 0.5. The lysogens were induced by incubation in a 45°C water bath for 20 min., and the cultures returned to a 37°C incubator for a further 2.5h. Cells were harvested by centrifugation, 5Krpm x 5 min at 4°C in the Sorvall GSA rotor, resuspended in 10ml L Broth per litre of original culture, transferred to 30ml Corex tubes and re centrifuged 5Krpm x 5 min in the SS-34 rotor.

The pellet from BHB2690 was used to make the sonicated extract (SE). It was resuspended in 3.5ml Buffer A (20mM Tris HCl pH8, 3mM MgCl<sub>2</sub>, 0.05%(v/v) 2-mercaptoethanol, 1mM Na<sub>2</sub>EDTA) and sonicated, setting 8 for 3x30s on ice, using the fine probe of an 'MSE soniprep'. The sonicated extract was centrifuged for 10 Krpm x 5 min at 4°C in the Sorvall SS-34 rotor, and the supernatant which was highly viscous, was frozen in liquid N<sub>2</sub> in 200ul aliquots and stored at -80°C until use.

The pellet from the BHB2688 culture was used to prepare the freeze thaw lysate (FTL). It was resuspended in 6ml 10% sucrose, 50mM Tris HCl pH7.5, in a Corex tube. 75ul of a 2 mg/ml solution of lysozyme in 25mM Tris HCl pH7.5 was added and the tube frozen in liquid N<sub>2</sub>. The lysate was allowed to thaw slowly on ice for 1h., then centrifuged 20Krpm x 60 min. at 5°C in the SS-34 rotor in 15ml thick walled polypropylene tubes. The supernatant was frozen in 100ul aliquots in liquid N<sub>2</sub> and stored at -80°C.

### 2.5.6 Extraction of maize nuclear DNA

Total DNA was prepared from etiolated coleoptile tissue as described by Weinand and Feix (1980) except that the CTAB precipitation step was

omitted.

Nuclear DNA was extracted from nuclei prepared by a modification of the method of Luthe and Quatrano (1980). All procedures were carried out on ice unless otherwise specified. Coleoptiles (ca.15g) which had been germinated in the dark for 60h. (section 2.3.1) were ground in an equal volume of Honda medium (Honda et al., 1966) in a precooled pestle and mortar. The suspension was filtered through 20u mesh nylon which was washed with several volumes of Honda medium. Nuclei were pelleted by centrifugation  $1000g \times 5 \text{ min. } 4^{\circ}\text{C}$ , and the pellet washed in Honda medium. The nuclear pellet was resuspended in 10ml Honda medium containing 1%(v/v) Triton-X-100 to lyse the organelles, and the nuclei were pelleted as before. This pellet was resuspended in 5ml 0.2M sucrose, 25mM TrisHCl pH7.4, 3mM  $\text{MgCl}_2$ , underlaid with a 10ml cushion of 0.3M sucrose, 25mM TrisHCl pH7.4, 3mM  $\text{MgCl}_2$ , and centrifuged  $275g \times 6 \text{ min.}$  The pellet obtained was washed by resuspension in cushion buffer. Nuclei were resuspended in 18ml lysis buffer (30mM TrisHCl pH8, 10mM  $\text{Na}_2\text{EDTA}$ ). Sarkosyl and protease K were added to final concentrations of 1%(w/v) and 50ug/ml respectively, and incubated  $37^{\circ}\text{C}$  with gentle shaking for 2-3h. Nuclear DNA was purified by centrifugation in CsCl equilibrium density gradients containing 1g CsCl per ml of original volume and 500ug/ml EtBr. Centrifugation was for 24h at 38K and  $15^{\circ}\text{C}$  in the Sorvall TVT865B vertical rotor. The DNA was removed from the gradients with a bent pasteur pipette and diluted with 4 volumes of TE80. EtBr was removed with n-butanol which had been equilibrated with TE80 saturated with CsCl. The DNA was precipitated with 2 volumes of ethanol, washed twice in 70% ethanol and resuspended in 10mM TrisHCl, 0.1mM EDTA pH8.



### 2.5.7 Ligation and in vitro packaging of recombinant phage.

20ug total maize DNA was digested to completion with 100U of the restriction endonuclease Eco R1, and the digest terminated by heating at 65°C for 10 min. This DNA was ligated to 20ug  $\lambda$ gtWES arms according to the following protocol.

10ug Eco R1 digested maize DNA and 10ug  $\lambda$ gtWES arms were incubated together at 55°C for 5 min. in 20mM Tris HCl pH7.6, 10mM MgCl<sub>2</sub> in a total volume of 115ul. Dithiothreitol and ATP (pH'd to 7.5) were added to final concentrations of 10mM and 1mM respectively, and the ligation reaction allowed to proceed for 60 min. at 29°C. <sup>in the presence of 10U of T4 DNA ligase.</sup> The reaction was terminated by addition of 100ul double distilled water, 20ul 10%(w/v)SDS, 10ul 0.4M Na<sub>2</sub>EDTA pH8, and the DNA recovered by ethanol precipitation.

Ligated DNA was resuspended in 50ul of double distilled water for packaging. 70ul Buffer A (section 2.5.7) 14ul buffer M (50mM Tris HCl pH7.5, 30mM Spermidine HCl pH7.5, 60mM Putrescine HCl pH7.5, 20mM MgCl<sub>2</sub>, 6mM ATP pH7.5, 25mM 2-mercaptoethanol), 60ul SE and 50ul FTL (section 2.5.7) were added, mixed, and the reaction incubated 10 min on ice followed by 15 min at 29°C. 500ul of (100X) BHB2600 plating cells (section 2.5.1) were added, and allowed to adsorb for 15 min. at 29°C. , and 25ul aliquots of the reaction plated with top agarose on dry L Broth plates.

### 2.5.8 Screening the library by plaque hybridisation.

The library was plated at a density of ca.10,000 plaques per plate on 60 x 90mm petri dishes. Plaque lifts were performed according to the method of Benton and Davis (1977), Maniatis et al., (1982) (see also section 2.6.3), and the filters hybridised with nick-translated plasmid

containing the maize ATP/ADP translocator cDNA clone pANT-1. Nitrocellulose filters were prehybridised in 6x SSPE, 0.1%(w/v) SDS, 10x Denhardt's solution, 100ug/ml denatured herring sperm DNA at 65<sup>0</sup>C for >30 min. Hybridisation was performed at 65<sup>0</sup>C in 3x SSPE, 0.1% (w/v) SDS, 10x Denhardt's solution, 100ug/ml denatured herring sperm DNA and the denatured probe (specific activity ca. 10<sup>8</sup>cpm/ug). Filters were washed in 2xSSPE, 0.1% (w/v) SDS for 2 x15 min at 65<sup>0</sup>C, air dried and exposed to x-ray film at -80<sup>0</sup>C with intensifying screens for three days.

Positive plaques were picked into 1ml 10mM MgSO<sub>4</sub> replated and screened as before. The process was repeated until pure cultures of phage were obtained.

#### 2.5.9 Isolation of small amounts of DNA from recombinant phage.

DNA was prepared from plate lysates of plaque purified recombinant phage as described by Maniatis et al., (1982), with the exception that 10mM MgSO<sub>4</sub> was used in place of SM buffer throughout.

#### 2.5.10 Subcloning maize nuclear DNA into pUC plasmids.

The pUC vectors (Messing and Vieira 1982a) are plasmids derived from pBR322, but carrying part of the lac Z gene and the M13 polylinker with the multiple cloning sites. The selection is for ampicillin resistant colonies and recombinants give rise to white colonies when plated upon media containing IPTG and Xgal.

The plasmid pUC9 was linearised with restriction endonuclease Eco R1, and ligated to Eco R1 digested DNA from the recombinant phage minipreps. As the concentration of miniprep DNA was not known accurately, three different ratios of plasmid to phage DNA were used in the ligation. Aliquots of the ligation reaction were used to transform E. coli strain.

MC1022 and the transformants plated on L-broth plates containing 100ug/ml Amp, and which had been spread with 30ul of 20mg/ml Xgal and 20ul of 24mg/ml IPTG diluted in 200ul of L broth.

Transformants exhibiting the white phenotype were picked at random and plasmid DNA prepared from them by the method of Birnboim and Doly (1979). The desired plasmids were identified by restriction endonuclease digestion with Eco R1 and hybridisation to the pANT-1 probe in Southern blots.

## 2.6 DNA SEQUENCE ANALYSIS

The chain terminator method of DNA sequence analysis (Sanger et al., 1977) was used in conjunction with the filamentous phage M13 vectors developed by Messing (Messing et al., 1977; Gronenborn and Messing 1978; Messing and Vieira 1982; Messing 1983). For a comprehensive review of M13 cloning and sequencing methods see Messing (1983).

### 2.6.1 Cloning.

At the outset of a sequencing project, plasmids containing DNA to be sequenced were digested with tetranucleotide recognising restriction enzymes e.g. Tag 1, Alu 1, Hae3, Sau3A, Msp 1, and shotgun cloned into M13 vectors. M13 recombinants containing the desired sequences were detected by plaque hybridisation (Benton and Davis 1977) with the appropriate hybridisation probe. At later stages in the sequencing project, specific fragments were 'force cloned' into the appropriate vector (Messing 1983).

When no convenient restriction sites were available for sequencing, Bal 31 nuclease was used to generate deletions from a known restriction site (Poncz et al., 1982).

### 2.6.2 Manipulation of M13 phage and transformation of E coli.

Competent JM101 cells were prepared and transformed by the method of Dagert and Ehrlich (1979). Recombinant phage cultures were grown and the single stranded DNA prepared as described by Messing (1983), except that L Broth was used as the growth medium.

### 2.6.3 Preparation of nitrocellulose filter replicas.

Plates containing phage 'plaques' were overlaid with precut circles of nitrocellulose. Orientation marks were provided by puncturing the filters with a needle. The nitrocellulose was left in contact with the plates for 2 min., then peeled off and laid 'plaque' side up on a piece of Whatmann 3MM paper soaked in Southern denaturing solution. After 2 min., the filters were transferred to a second piece of 3MM paper soaked in neutralising solution for a further 2 min., and finally to a piece of 3MM paper soaked in 2x SSPE. Filters were dried on 3MM paper and baked 60 min. at 80°C under vacuum.

### 2.6.4 Identification of recombinant M13 clones.

#### 2.6.4.1 In situ hybridisation.

Nitrocellulose filters were prepared directly from the plates obtained as a result of the transformation as described above. Filters were hybridised with the appropriate <sup>32</sup>P labelled DNA fragment. This approach was used to screen large numbers of plaques for relatively rare recombinants.

#### 2.6.4.2 Ordered grids.

Minimal plates were overlaid with 5ml BBL top agar containing 200ul JM101 plating cells, which were prepared by growing a 1 in 100 dilution of a fresh overnight culture for 2h. at 37°C. White plaques from the

transformation were picked onto duplicate plates in 6x8 arrays using sterile tooth picks, and the plates incubated for 12-16h at 37°C. Plaque lifts were then made from each plate as described above. Multiple lifts (up to three) were made from each plate if more than one probe was to be employed. The principal advantage of ordered grids is the ease with which hybridising plaques can be identified and located.

### 2.6.5 Plaque purification

Positive plaques were picked into 1ml of LTB with a sterile tooth pick, and the phage suspension diluted 1 in 1000 with LTB. One microlitre of the 1 in 1000 dilution was added to 100ul JM101 plating cells, 10ul Xgal and 7ul IPTG, and plated with 1ml BBL top agar on 3.5cm minimal plates. This procedure reliably gave single well separated plaques which could be used for the preparation of single stranded template DNA for sequencing.

### 2.6.6 Sequencing reactions and gels.

Sequencing reactions were carried out according to Sanger et al., (1977), Sanger et al., (1980), except that the synthetic M13 sequencing primer (5'TCCCAGTCACGACGT3') was used to prime the reactions, which were carried out in 1.5ml Eppendorf tubes. The labelled nucleotide was alpha<sup>32</sup>PdCTP (410Ci/mmol).

Sequencing gels were 8% (w/v) polyacrylamide (38:2 acrylamide:bis), 8M urea, dissolved in 1x TBE. Gels were poured between glass plates 30cm x 40cm separated by 0.4mm plasticard spacers. Typically 10 clones (40 reactions) were loaded onto this size of gel, and electrophoresed at 65W constant power for 2h. (short run) or 6h. (long run). Gels were fixed by immersion in 10% (v/v) acetic acid for 10 min., rinsed briefly in tap water and dried onto Whatmann 3MM paper. Dried gels were exposed to Fuji x-ray

film for 16h. to 3days.

### 2.6.7 Computer methods.

DNA sequence was assembled on the Apple IIe microcomputer. Sequences read from gels were entered into the computer and checked for homology to any other sequence already entered by means of the program FINDMATCH. This program takes a short sequence e.g. 10 bases, from the beginning and end of a longer sequence and searches for its occurrence in any other sequence (or its complement), specified. In this way overlapping and complementary sequences can rapidly be identified and combined into longer sequences. The programmes of the University of Wisconsin Genetics Computer Group (UWGCG) Wisconsin package were used for further analysis of completed DNA sequences.

## 2.7 RNA METHODS.

### 2.7.1 Denaturing gels

#### (a) Glyoxal gels (McMaster and Carmichael 1977)

Samples were made 50% (w/v) DMSO, 10mM NaPO<sub>4</sub>pH7, and 1M with respect to glyoxal, incubated 1hr at 50°C and cooled on ice. 0.1 vol sample buffer (50%w/v glycerol, 10mM NaPO<sub>4</sub> pH7, 1mg/ml bromophenol blue) was added and the samples electrophoresed through agarose gels in 10mM NaPO<sub>4</sub> pH7 with continuous recirculation of buffer.

#### (b) Formaldehyde gels (Lehrach *et al.*, 1977)

Samples were made 50%(w/v) formamide, 6% formaldehyde, 1x MOPS buffer, heated at 65°C for 5min and cooled on ice. 0.2 vol of sample buffer (50% w/v formamide, 1mg/ml bromophenol blue) was added, and the samples electrophoresed through agarose gels containing 6% formaldehyde



in 1x MOPS buffer.

### 2.7.2 Transfer of RNA to nitrocellulose and hybridisation conditions.

RNA was transferred to nitrocellulose as described by Thomas (1980). Filters were prehybridised in 50% (w/v) formamide, 5xSET, 0.05M NaPO<sub>4</sub> pH7, 0.1% NaPP, 0.1%SDS, 10x Denhardt's solution, 100ug/ml denatured herring sperm DNA at 42<sup>0</sup>C. Hybridisation was performed in the same buffer but including the probe, and with the concentration of Denhardt's reduced to 1x, for 12-18h at 42<sup>0</sup>C. The first wash was in 2xSET, 0.1%SDS, 0.1%NaPP at 55<sup>0</sup>C for 15 min., and subsequent washes (2-3) in 2xSSC, 0.1%SDS, 0.1%NaPP at 55<sup>0</sup>C for 15 min. Filters were air dried and exposed to X-ray film at -80<sup>0</sup>C with intensifying screens.

### 2.7.3 S1 Nuclease mapping of RNA molecules.

S1 nuclease mapping was carried out as described by Sharp et al., (1980).

#### 2.7.3.1 Preparation of labelled DNA probes

M13 subclones covering the presumptive start of transcription were uniformly labelled with <sup>32</sup>P by second strand synthesis, such that the labelled strand was complementary to the RNA. The standard second strand synthesis conditions for Klenow were used, except that after 1h incubation at 37<sup>0</sup>C the reaction was 'chased' by the addition of 8ul of a 2mM solution of non radioactive dNTP's. The unincorporated nucleotides were removed by passing the reaction through a G50 centrifuged column, and the <sup>32</sup>P labelled DNA precipitated with ethanol. The pellet was washed with 70% ethanol, dried and counted. The pellets were resuspended in 10mM TrisHCl 1mM EDTA pH8 at 1-3x10<sup>5</sup>cpm/ul.

### 2.7.3.2 DNA /RNA hybridisation.

Four parallel reactions were set up for each probe. The first contained the labelled DNA alone, the second labelled DNA plus 10ug yeast tRNA, and the third and fourth labelled DNA, 10ug yeast tRNA and 5 and 10ug maize poly A+ RNA respectively.  $3-5 \times 10^5$ cpm of labelled DNA were used per reaction, and each set of four reactions with the same probe had equal numbers of input counts.

The DNA, RNA and tRNA were mixed together in 1.5ml eppendorf tubes and dried in a speed vac. The pellets were resuspended in 2ul of 5x hybridisation buffer (2M NaCl, 200mM PIPES-NaOH pH6.4, 5mM EDTA) and 8ul of de-ionised formamide, heated to 82°C for 10 min., transferred directly to a waterbath which had been stabilised at the desired temperature, and incubated submerged for 12-20h. The temperature for hybridisation was determined from the empirical formula of Howley *et al.*, (1979)

$$T_m = 81.5 + 16.6 \log_{10} M + (0.41 \times \%GC) - (0.72 \times F)$$

Where M = the molar concentration of Na<sup>+</sup>

%GC = the GC content of the probe

F = the % of formamide

The formula was derived for DNA/DNA hybridisation. As DNA/RNA hybrids are more stable in formamide than DNA/DNA hybrids, the hybridisations were performed at  $T_m + 11$  as this reduces the extent to which the probe self associates ( J. Bailey-Serres pers. comm.)

### 2.7.3.3 S1 nuclease digestion.

S1 nuclease was diluted with ice cold S1 buffer (250mM NaCl, 30mM Na acetate pH4.6, 1mM ZnSO<sub>4</sub>), to a concentration of 250U/ml and kept on ice.



100ul of the ice cold S1 mix plus 2ul of 10mg/ml salmon sperm DNA were added to each of the RNA containing hybridisations whilst they were still in the waterbath. The tubes were immediately removed, placed on ice for 10 secs., briefly vortexed and incubated at 37°C. After 30 min. digestion the reactions were returned to ice and ethanol precipitated in the presence of 10ug tRNA. The pellets were resuspended in 2mM ammonium acetate, 5mM trisHCl pH8, 0.5mM EDTA pH8, and reprecipitated. The precipitates were washed with 70% ethanol, dried and counted. Aliquots of the S1 reactions were electrophoresed through standard sequencing gels alongside a sequencing ladder of a clone of known sequence. Gels were fixed, dried and exposed as for sequencing gels.

#### 2.7.3.4 Estimation of the size of S1 protected fragments.

The fragments which were unique to the poly A<sup>+</sup> containing reactions, and which showed an increase in intensity at the higher poly A<sup>+</sup> concentration were considered to be true S1 protected fragments. The size of these fragments was measured by their position relative to the sequencing reaction. The value obtained for the S1 fragments had to be increased by an amount corresponding to the distance from the M13 primer binding site to the cloning site used in the construction of the clone used as the probe, as this sequence is not complementary to the RNA and is removed by S1 digestion, but is present in the products of the sequencing reaction used for the size markers.

#### 2.7.4 Primer extension analysis of mRNA 5' termini.

##### 2.7.4.1 Preparation of the primers.

An M13 clone containing the sequences desired for primers was

labelled by second strand synthesis as in the preparation of S1 probes. After the chase reaction, the now double stranded DNA was digested with the appropriate restriction endonuclease(s) to generate the desired fragment(s), which were isolated from a polyacrylamide gel by the crush soak method of Maxam and Gilbert (1977 ). The isolated fragments were subjected to secondary digests to generate more specific primers, and the products of this reaction were separated by electrophoresis through an 8M urea, 6% (38:2 acrylamide:bis) polyacrylamide gel. The desired single stranded primers were detected by autoradiography, excised from the gel and extracted by the crush soak method.

#### 2.7.4.2 The primer extension reaction.

Labelled primer (ca.  $10^3$  cpm) was mixed with 10ug polyA<sup>+</sup> RNA or 10ug yeast tRNA, and dried down. The nucleic acids were resuspended in 4.4ul of annealing buffer (100mM Tris HCl pH8, 100mM MgCl<sub>2</sub>), heated to 80°C for 5 min., and allowed to cool in air to room temperature. Reverse transcriptase mix was prepared immediately before use as follows:

- 2.5ul 1.0M Tris HCl pH8.3
- 5.0ul 500mM KCl
- 5.0ul 80mM MgCl<sub>2</sub>
- 5.0ul 1mM DTT
- 5.0ul 40mM NaPP
- 1.0ul 1mM dATP
- 1.0ul 1mM dGTP
- 1.0ul 1mM dTTP
- 1.0ul 1mM dCTP
- 1.0ul 40U Reverse Transcriptase

5.6ul of this mix was added to each reaction to give final concentrations

of 144mM Tris HCl, 50mM KCl, 52mM MgCl<sub>2</sub>, 100uM DTT, 4mM NaPP, 100uM dNTP's, and the reaction was incubated at 40<sup>0</sup>C for 25 min. Reactions were stopped by freezing or the addition of formamide dye. Primer extension reactions were electrophoresed through sequencing gels and sized from sequencing reactions in the adjacent lanes.

### CHAPTER3 CONSTRUCTION OF A cDNA LIBRARY FROM MAIZE POLY A<sup>+</sup>RNA, AND IDENTIFICATION OF CLONES FOR NUCLEAR ENCODED MITOCHONDRIAL PROTEINS.

#### 3.1 RATIONALE

In order to study many of the processes involved in mitochondrial biogenesis, the availability of defined gene probes is an essential prerequisite. Such probes will allow an investigation of the location, number and organisation of genes involved in the biogenesis of mitochondria, and their expression under defined environmental conditions or at particular stages of development.

The isolation of cDNA clones is the first step in the isolation of the nuclear genes themselves. DNA sequence analysis of these genes will permit comparison with other known gene sequences in an attempt to identify common regulatory elements, or sequences involved in directing the translation product to specific sub cellular compartments. Recent advances in plant transformation techniques mean that it will now be possible to test the functional significance of these sequences *in vivo*, by the reintroduction of altered genes into plant cells.

#### 3.2 SURVEY OF METHODS FOR THE CONSTRUCTION OF cDNA LIBRARIES.

Construction of a cDNA library involves the enzymatic synthesis of single stranded cDNA complementary to the mRNA, subsequent conversion of single stranded to double stranded DNA, and cloning into a plasmid or bacteriophage vector. Since the original method of Efstratiadis *et al.*, (1976) a number of variations have appeared designed to circumvent particular problems inherent in the method. These have largely been concerned with improving the overall yield of cDNA from RNA, and the

efficiency with which the 5' end of the mRNA molecule can be cloned.

Table 3a is a summary of the different procedures.

TABLE 3a

**METHODS OF cDNA LIBRARY CONSTRUCTION.**

**I) Enzymatic Conversion of poly A+ mRNA to double stranded cDNA.**

**A) 1st strand synthesis.**

- 1) Prime by oligo (dT)12-18.<sup>1,2,4</sup>
- 2) Prime by hybridization to T tailed vector.<sup>3</sup>
- 3) Additives to inhibit RNase activity.
  - a) RNasin.
  - b) Na pyrophosphate.

**B) 2nd Strand Synthesis:**

- 1) Prime by 1st strand hairpin loop.<sup>1</sup>
- 2) Prime by oligo (dG)12-18 hybridized to 3' C tail.<sup>2</sup>
- 3) "Nick-translation" of RNA/DNA hybrid.<sup>3,4</sup>

**II) Cloning of cDNA into a vector to make library.**

- A) Homopolymeric tailing of cDNA and vector.
- B) Ligation of DNA linkers to cDNA.
- C) Synthesis of cDNA directly into vector.<sup>3</sup>

1= Efstratiadis et al., (1976) 2= Land et al., (1981)

3= Okayama and Berg (1982) 4= Gubler and Hoffman (1984)

---

### 3.2.1 Methods for synthesising double stranded cDNA.

Synthesis of the first strand of the cDNA is usually primed by oligo dT hybridised to the 3' polyA tail of the mRNA, either as an oligomer, or as an poly dT 'tail' attached to a vector primer. AMV reverse transcriptase is used to copy the mRNA sequence, extending the oligo dT primer.

The single stranded cDNA has the ability to form a hairpin loop at the 3' end and so prime the synthesis of a complementary second strand, either by reverse transcriptase or E.coli DNA polymerase 1 (Efstratiadis et al., 1976). The resulting double stranded product is covalently joined at one end, and must be digested with a single strand specific nuclease (e.g S1) to yield blunt ended molecules suitable for cloning. S1 nuclease digestion frequently results in the loss of sequences representing the 5' end of mRNAs, and also the hairpin priming has been found to result in the generation of artifactual sequences (e.g. Fagan et al., 1980). To overcome these problems several procedures have been developed which eliminate this step.

(1)The method of Land et al.,(1981,1983) uses calf thymus terminal deoxynucleotidyl transferase to add dC tails to the 3' end of the first strand, which allows priming of second strand synthesis with oligo dG. The claimed efficiency is  $10^4$ - $10^5$  transformants per microgram of RNA.

(2)Okayama and Berg (1982) used poly dT tailed plasmid vector annealed to mRNA as a primer for first strand synthesis, then tailed the cDNA/RNA hybrid with dCTP, and annealed a dG tailed linker to this. The linker has at its other end a sticky end complementary to the free end of the vector primer, so that the vector becomes cyclised on annealing. The RNA strand is converted to DNA with E.coli RNase H , DNA polymerase 1 and DNA ligase. This method is claimed to be highly efficient, yielding  $10^5$  transformants per microgram RNA, with a very high recovery of 5'

terminal sequences. The major disadvantage is the time consuming preparation of the vector primers, although these are now commercially available.

(3) The method of Gubler and Hoffman (1983) uses the standard oligo dT priming technique for first strand synthesis, then the RNA replacement step of Okayama and Berg. The cDNA can then be cloned into the vector by the method of choice.

### 3.2.2 Strategies for cloning double stranded cDNA.

The vector primed methods result in the cDNA being synthesised directly into the vector. Methods such as those of Efstratiadis, Land and Gubler and Hoffman yield double stranded cDNA which can be cloned in a number of ways. Most simply the ds-cDNA can be ligated directly into a vector digested with a restriction enzyme which gives blunt ends. In practice this method may be less useful as;

(1) Blunt ended ligations are relatively inefficient.

(2) Excising the inserted cDNA is difficult unless it was cloned into a polylinker which contains flanking restriction sites

(3) Multiple inserts may be obtained, complicating further analysis.

The most commonly used methods are the ligation of oligo nucleotide linkers, containing specific restriction sites, to the cDNA (Ulrich *et al.*, 1977), or the addition of homopolymer tails, Jackson *et al.*, (1972) Maniatis *et al.*, (1976). The major disadvantage of oligo nucleotide linkers is that the procedure involves ligation of the linkers to the cDNA followed by digestion with the restriction enzyme that cleaves within the linkers, in order to generate the 'sticky ends' for ligation. Thus if this restriction site is also present in the cDNA, incomplete clones will be obtained. This

can be circumvented in the case of restriction enzymes which are sensitive to methylation, by first treating the cDNA with the appropriate methyl transferase. Double inserts may be obtained with linkers, particularly if the size of the cDNA is small, but is less of a problem than with blunt end ligation.

Homopolymer tailing is probably the most widely used cloning method for cDNA. Usually calf thymus terminal deoxynucleotidyl transferase (terminal transferase) is used to add dC residues to the 3' ends of the cDNA and dG residues to the 3' ends of the vector. The complementary ends are annealed and are ligated upon transformation into E.coli. If the vector is cut with Pst 1 to generate the substrate for tailing, the Pst 1 site is regenerated. The disadvantages of homopolymer tailing are that excessive lengths of G-C tail may lead to insert instability when propagated in  $recA^+$  strains of E. coli, and that G-C tails interfere with sequencing by the dideoxy chain terminator method. (section 4.3.2)

### 3.2.3. Strategies for screening cDNA libraries.

A number of approaches have been developed which may exploit particular properties of the sequences to be identified. Some of these are outlined below. The method of screening may dictate the vector to be used in construction of the library.

#### (1) Colony Hybridisation.(Grunstein and Hogness 1975)

Bacterial colonies containing the recombinant plasmids are grown on nitrocellulose filters or other similar solid support. The colonies are lysed and the DNA immobilised on the filters which are then hybridised with the probe.

The probe employed will depend upon the gene(s) of interest. Genes which exhibit developmental or organ specific expression, or are inducible



by environmental conditions can be identified by a +/- hybridisation. Labelled RNA, or cDNA synthesised from RNA from induced (+) and uninduced (-) cells can be used to probe duplicate filters, and the colonies which hybridise more intensely to the (+) probe selected for further characterisation. Alternatively the hybridisation of labelled (+) RNA can be competed with unlabelled (-) RNA, and clones which are least affected by the competitor RNA selected (e.g. van Loon et al., 1982). In some instances RNA can be significantly enriched for the desired product e.g. by fractionation in sucrose gradients or immunoprecipitation of polysomes. If the amino acid sequence of the gene product is known, synthetic oligonucleotides corresponding to the predicted nucleotide sequence may be employed. Alternatively, if the homologous gene from another organism has been cloned, it can be used to probe the library under conditions of reduced stringency which will take into account the sequence divergence between the same gene from different organisms.

(2) Hybrid select translation. (Ricciardi et al., 1979)

If an antibody is available, batches of recombinant clones can be used to hybrid select complementary mRNA which is subsequently translated in a cell free system, and the products immunoprecipitated. Although this method has been used for primary screening (e.g. Viebrock et al., 1982) it is laborious for the screening of large libraries. This method is often used to confirm the identity of clones picked as a result of other types of screening.

(3) Direct immunological screening of expression libraries. (Broome and Gilbert 1978)

An alternative method of utilising an antibody as a probe requires the sequences encoded in the cDNA to be expressed as a fusion protein. The cDNA must be cloned into an expression vector in the correct orientation

and reading frame to make a fusion protein which has antigenic determinants recognised by the antibody. Several plasmid and bacteriophage expression vectors are now available which provide the necessary transcription and translation signals for expression of the foreign protein in E. coli. (e.g. Young and Davis 1983, Ruther and Muller-Hill 1982)

(4) In vivo complementation of mutants.

At present this is applicable mainly to bacteria and fungi where there are well characterised mutants and efficient transformation systems. A mutant in the structural gene of interest is transformed with DNA from the wild type, in the case of yeast the DNA is cloned in a shuttle vector which can replicate in both yeast and E. coli (Beggs, 1978). Colonies which exhibit co-segregation of the wild phenotype and vector borne markers are selected for further characterisation.

### 3.3 RATIONALE FOR SCREENING THE MAIZE cDNA LIBRARY BY HETEROLOGOUS HYBRIDISATION.

In Saccharomyces cerevisiae the phenomenon of glucose repression has been exploited to identify nuclear genes encoding mitochondrial proteins. Transcription of these genes is repressed in the presence of glucose, and this has been used as the basis of a competition hybridisation assay (van Loon et al., 1982). In higher plants there is insufficient knowledge of the effects of environmental factors on the synthesis of mitochondrial components to permit the adoption of a similar approach. Screening methods relying on antibodies were also not possible as at the time none were available to specific plant mitochondrial proteins.

Heterologous hybridisation has been used successfully to identify a number of maize mitochondrial genes. Fox and Leaver (1981) used a clone

for cytochrome oxidase subunit II (COX II) from yeast to identify the maize equivalent, and Dawson *et al.*,(1984) and Isaac *et al.*,(1985) identified apocytochrome b (COB) and COX I by the same method. This approach has also been used to identify a number of other genes which are apparently well conserved between even quite distantly related organisms. For example *Saccharomyces* elongation factor Tu cross hybridised with an *E.coli* *tufB* gene probe, Nagata *et al.*,(1983), Rat cytochrome *c* with a yeast probe, Scarpulla *et al.*,(1981), and homeo box sequences from a number of organisms including humans, mice, chickens and toads have been identified with *Drosophila* gene probes. McGinnis *et al.*,(1984), Carrasco *et al.*,(1984).

Clearly a major limitation is the availability of a gene probe for the sequence of interest. For some genes the homology with their counterparts from other organisms is too low, as with some of the mitochondrially encoded polypeptides which are now believed to be part of the rotenone sensitive NADH dehydrogenase. In other cases plant mitochondria may have unique polypeptides e.g. nuclear restorer genes, which suppress the mitochondrially inherited CMS phenotype, for which there may not be an equivalent in animal or fungal cells (section 1.7.2). However, as many of the polypeptides of the mitochondrial inner membrane exhibit considerable cross species homology and as probes for several of these were available from both yeast and *Neurospora*, this approach was the one taken in this study.

### 3.4 CONSTRUCTION OF A MAIZE cDNA LIBRARY.

#### 3.4.1 Aims

The aim of constructing the library was to identify and characterise specific cDNA clones which could then be used

- (1) to screen a lambda nuclear genomic library and identify the

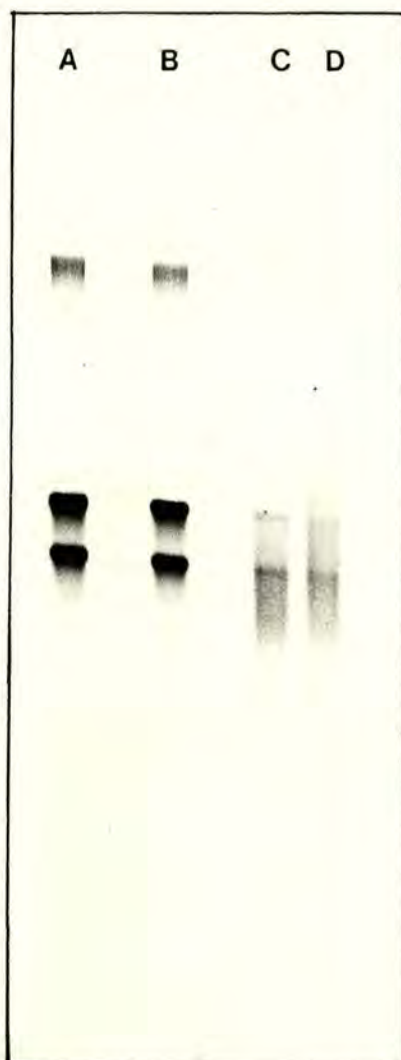
corresponding nuclear gene and

(2) as hybridisation probes to assess steady state mRNA levels during mitochondrial biogenesis.

Consequently the recovery of full length clones was not initially a prime consideration, although clearly it would be advantageous in the long term. Several of the more recent methods described above have been published since this piece of work was initiated, therefore the method employed was in essence that of Efstratiadis et al.,(1976) as modified by Buell et al.,(1978).

#### 3.4.2 Isolation of RNA from maize and synthesis of dsDNA.

RNA was extracted from maize coleoptiles, germinated for 60h in total darkness, during which time mitochondrial biogenesis is very active (V. Jones pers. comm.), and chloroplast development is retarded. RNA was extracted by two methods, the guanidinium thiocyanate method of Chirgwin et al.,(1979), and the phenol-detergent method of Leaver and Ingle (1971), and used to prepare poly A<sup>+</sup> RNA. The two RNA preparations were compared by electrophoresis in non-denaturing agarose gels, Fig.3.1. A low level of rRNA contamination was present in both samples after two cycles of elution from oligo dT cellulose. The guanidinium thiocyanate extracted RNA appeared to contain a larger proportion of high molecular weight RNA (larger than the 25S rRNA), but this may reflect differing degrees of secondary structure between the two samples. No difference could be detected in the spectrum of polypeptide products of these RNA's when translated in a wheat germ in vitro protein synthesising system. As the phenol-detergent method consistently yielded more RNA than the guanidinium thiocyanate method, (average 19 ug polyA RNA per g fwt tissue cf. 5 ug/g fwt) it was routinely used to prepare RNA for cDNA



**Fig.3.1** Comparison of RNA extracted by the guanidinium thiocyanate and phenol detergent methods.

Ethidium bromide stained agarose gel. Lane A 5ug total RNA; B, 5ug polyA<sup>-</sup>RNA; C, 5ug polyA<sup>+</sup>RNA extracted by the phenol-detergent method; D, 5ug polyA<sup>+</sup>RNA extracted by the guanidinium thiocyanate method.

An equal volume of loading buffer (23mM TrisHCl pH7.5, 2.3mM EDTA, 0.25%(w/v) SDS, 20%(w/v) sucrose, 42%(w/v) urea, 0.5mg/ml xylene cyanol, 0.5mg/ml bromophenol blue) was added to each sample which was heated to 50<sup>0</sup>C for 2 min., rapidly cooled on ice and electrophoresed in a 1.1%(w/v) agarose gel in 10mM sodium phosphate buffer pH7. Buffer was recirculated during electrophoresis.

synthesis, in vitro translation and transcript analysis by Northern blotting and S1 mapping.

cDNA was synthesised as described in section 2.4.1. Samples of the  $^{32}\text{P}$  labelled first strand and double stranded cDNA before and after S1 treatment were treated with glyoxal and fractionated by electrophoresis through a 1% (w/v) agarose gel (section 2.7.1), and autoradiographed to determine the size of products and the efficiency of the S1 treatment. After S1 digestion, the cDNA was size fractionated in sucrose gradients. This removes small fragments of cDNA which would otherwise be over represented in the library, and selects for larger cDNA's. Aliquots of one tenth of each of the fractions from the gradient were analysed by electrophoresis in agarose mini gels. The fraction from the bottom of the gradient (which contained the longest cDNA molecules), had a median size of around 1600bp, and the next fraction was only slightly smaller. There was considerable overlap in size between the fractions.

### 3.4.3 Optimisation of the homopolymer tailing reaction

Homopolymer tailing is one of the most critical steps in construction of a library, so the reaction must be calibrated to add the correct number (15-20) residues to cDNA and vector. Deviation from this optimum causes a sharp decrease in the number of transformants obtained, (Peacock et al., 1981)

The 3' protruding ends of a double stranded DNA molecule are the favoured substrate for terminal transferase. The addition of divalent ions such as  $\text{Co}^{2+}$  or  $\text{Mn}^{2+}$  makes it possible for the enzyme to utilise blunt or recessed 3' ends, Nelson and Brutlag (1979), Deng and Wu (1981), Maniatis et al., (1982). It has been suggested that  $\text{Co}^{2+}$  favours the addition of dCMP tails and  $\text{Mn}^{2+}$  the addition of dGMP, although  $\text{Mn}^{2+}$  may improve the

tailing of blunt ends, Deng and Wu (1981).

#### 3.4.3.2 Optimising the tailing of inserts

The approximate rate of the addition of dC residues to small blunt ended DNA molecules was measured in a time course experiment using Alu I generated restriction fragments of lambda DNA, Fig.3.2. This gives an average figure for the number of residues added per 3' end, but the calculation requires the assumption that all molecules are tailed. The kinetics of terminal transferase suggest that the enzyme dissociates between rounds of dNTP addition (Michelson and Orkin 1982). The higher efficiency of tailing of 3' protruding ends compared with blunt ends means that molecules tailed early in the incubation will have their tails extended in preference to the enzyme initiating on untailed molecules. Therefore this assumption is probably not valid, and may lead to an under estimation of tail length.

#### 3.4.3.2 Optimising the tailing of the vector

Vector tailing was assayed by measuring the change in mobility in polyacrylamide gels of tailed restriction fragments. Pst I digested pAT153 was tailed with dG residues, then digested with Hae III and separated by electrophoresis in a 6% polyacrylamide gel, alongside a Pst/Hae III double digest. Fig.3.2. The restriction fragments containing the Pst I 'sticky ends' (122bp and 145bp ) increased in mobility relative to those in the untailed digest, and an estimate of the number of tails added (15 and 25 residues respectively, ) could be obtained by calculating the size increase. The disappearance of the 122 and 145bp restriction fragments and their replacement by discrete (if faint ) fragments at 37 and 70 bp (indicated by arrows in Fig.3.2) indicate that a very high proportion of the plasmid molecules are tailed, and that the reaction is



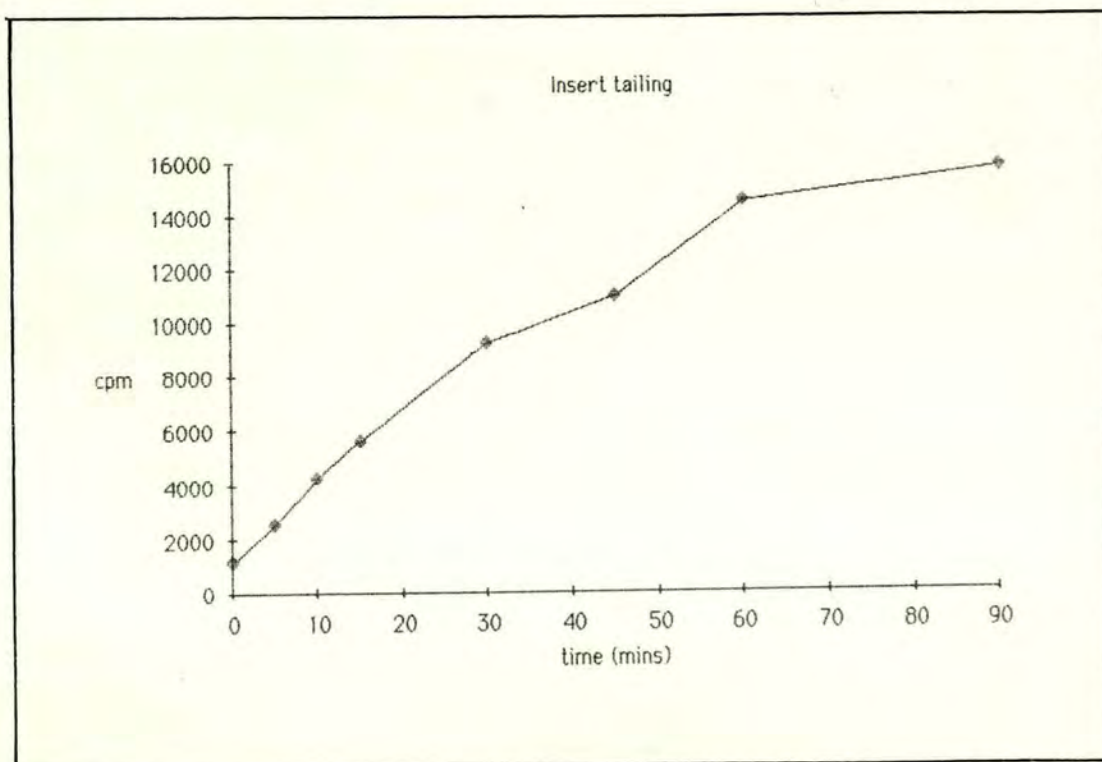
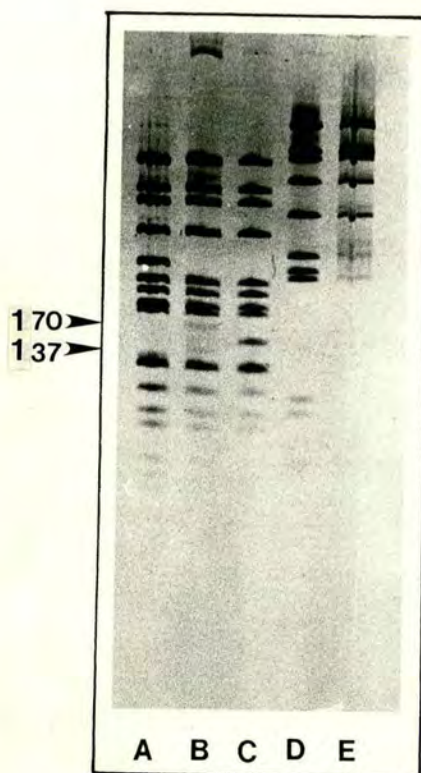


Fig.3.2 Analysis of homopolymer tailing.

Upper panel: 6%(w/v) polyacrylamide gel stained with ethidium bromide. Lanes are as follows: A, pAT153 digested with HaeIII; B, pAT153 digested with PstI, tailed with dGTP for 5 min. as described in materials and methods, then digested with HaeIII; C, pAT153 digested with HaeIII and PstI; D, pBR322 digested with AluI and E, pBR322 digested with AluI and tailed with dCTP for 10 min. as described in materials and methods. The two bands indicated with arrows are those which correspond to the two tailed fragments (lane B).

Lower panel: A graph illustrating the incorporation of  $^{32}\text{P}$  dCTP into homopolymer tails by terminal transferase using AluI generated fragments of lambda DNA as a substrate.

Fig.3.2



self limiting. In contrast restriction fragments derived from an Alu 1 digest of lambda DNA, tailed and run on the same gel do not give discrete fragments of higher molecular weight, but merely smear. This suggests that the enzyme can add infinitely long tails to the inserts.

### 3.4.3.3 Comparison of the efficiency of different tailing protocols in the construction of chimaeric plasmids and their introduction into E. coli.

In order to overcome the problems associated with calculating both the rate of tailing and the fraction of molecules tailed, and to test the annealing and transformation steps, a library was constructed using lambda DNA as mock inserts. Pst 1 digested vector DNA and Alu 1 generated lambda fragments were tailed under a variety of conditions, annealed pairwise and used to transform E.coli strain 294. The number of transformants per microgram of recombinant DNA and the percentage Amp<sup>r</sup> colonies were scored, Table 3b. The number of transformants/ug obtained varied over an order of magnitude from  $1.1 \times 10^4$  -  $9.8 \times 10^4$ . The best results were obtained with vector tailed for 10 min. in the presence of  $\text{Co}^{2+}$ , and insert for 5 min. in the presence of  $\text{Mn}^{2+}$ . However the difference between the highest and lowest values in the first two rows and three columns is only a factor of two, and this may not be significant given that only a small number of colonies (100-300) obtained in a single experiment were scored. Doubling the amount of insert DNA did not increase the number of colonies obtained on a per ug basis. The conditions recommended by Michelson and Orkin (1982) for efficient insert tailing gave consistently poor results in this experiment (column 4). The tailing conditions used for construction of the maize cDNA library were:

pAT153- 10 min., 37<sup>0</sup>C, 10mM  $\text{Co}^{2+}$  100ng/ul DNA

cDNA - 5 min., 37<sup>0</sup>C, 10mM  $\text{Co}^{2+}$  5ng/ul dscDNA.

TABLE 3b.

INSERT VECTOR	1	2	3	4
A	$5.9 \times 10^4$ cols./ug 50%Amp <sup>s</sup>	$7.65 \times 10^4$ cols/ug N.D.	$7.35 \times 10^4$ cols/ug N.D.	$1.15 \times 10^4$ cols/ug 24%Amp <sup>s</sup>
B	$4.05 \times 10^4$ N.D.	$8.70 \times 10^4$ 63%Amp <sup>s</sup>	$7.70 \times 10^4$ 76%Amp <sup>s</sup>	$1.30 \times 10^4$ N.D.
C	$3.35 \times 10^4$ 52%Amp <sup>s</sup>	$7.20 \times 10^4$ N.D.	$9.80 \times 10^4$ N.D.	$3.30 \times 10^4$ 43%Amp <sup>s</sup>
D	$6.30 \times 10^4$ 89%Amp <sup>s</sup>			
A <sub>2</sub>		$5.80 \times 10^4$ 74%Amp <sup>s</sup>		
B <sub>2</sub>		$6.40 \times 10^4$ N.D.		

Uncut pAT153  $4 \times 10^5$  colonies per ug.

Vector A alone - <250 colonies

No DNA - no colonies.

A<sub>2</sub>, B<sub>2</sub>, vectors A and B annealed with double the amount of insert DNA

N.D. - not determined.

#### Tailing protocols

1. 100mM Na cacodylate, 1mM CoCl<sub>2</sub>, 0.2mM DTT, 1mM dCTP, 10U TdT, 37°C, 5 min.

2. As above but incubated for 2 min

3. As 1. but with 10mM MnCl<sub>2</sub> in place of CoCl<sub>2</sub>.

4. 120mM Na Cacodylate, 1mM CoCl<sub>2</sub>, 0.1mM DTT 5uM dCTP, 500ug/ml BSA, 12°C, 20 min.

A. As 1 except dGTP was substituted for dCTP

B As A, except incubation was for 10min.

C As A, except incubation was for 10min at 20°C.

D Commercially tailed pBR322.

### 3.5 CHARACTERISTICS OF THE MAIZE cDNA LIBRARY.

The annealed cDNA/pAT153 transformed into E. coli strain 294 gave rise to colonies at a frequency of  $5.5 \times 10^3$  per ug recombinant DNA. Tailed vector alone gave less than  $10^2$  colonies per ug, and no colonies were observed in the absence of exogenously added DNA. Supercoiled pAT153 gave  $2.8 \times 10^7$  colonies /ug. The library was expanded to a total of 2,500 clones, of which 76% were Tet<sup>r</sup>, Amp<sup>s</sup>.

One plate was selected at random from the library and SDS lysates made of 32 colonies (Barnes 1977). The lysates were fractionated by electrophoresis through 1% (w/v) agarose gels and the size of the plasmids estimated relative to undigested plasmids of known size. Twenty-five out of the thirty-two colonies contained detectable plasmids. The mobility of these plasmids relative to pAT153 suggested that they contained cDNA inserts of 250 to 900bp.

### 3.6 SCREENING THE LIBRARY FOR cDNA CLONES ENCODING MITOCHONDRIAL PROTEINS BY HETEROLOGOUS HYBRIDISATION.

A number of factors have to be taken into consideration when using a gene probe from another organism to isolate the homologous gene from a higher plant by DNA/DNA hybridisation.

Firstly, there is the predicted homology between the probe and the target sequence. This will determine the melting temperature of a heteroduplex formed between the probe and the target sequence, and hence the hybridisation conditions which must be employed to detect that sequence. As the true level of homology is unknown, a "best guess" figure has to be employed, and the stringency of the hybridisation conditions determined empirically.

Secondly, higher plants present an additional complication in that

some mitochondrial proteins have counterparts in the chloroplast, e.g. ATPase subunits, cytochromes. Neither higher plant mitochondrial or chloroplast RNAs are polyadenylated (C.J. Leaver, unpublished observations, Wheeler and Hartley 1975) and therefore should not be represented in the cDNA library. However, in some instances the analogous mitochondrial and chloroplast proteins are both nuclearly encoded and their mRNAs might have been cloned. In addition the possibility of either mitochondrial or chloroplast RNA's being present in the poly A<sup>+</sup> fraction due to the presence of A-rich tracts elsewhere in the transcription unit cannot be ruled out. In order to try and reduce this problem RNA was extracted from young etiolated tissue where mitochondrial biogenesis is very active but chloroplasts are relatively undifferentiated.

Another consideration is that E.coli contains proteins analogous ( and homologous) to many mitochondrial proteins, which could lead to a high background in low stringency colony hybridisations, particularly if the probe sequence happened to be more closely related to the E.coli gene than the plant gene.

### 3.7 ISOLATION OF A cDNA CLONE ENCODING THE ADENINE NUCLEOTIDE TRANSLOCATOR OF MAIZE

The adenine nucleotide translocator (ANT) has a number of features which make it of interest for our studies. The structure and properties of the ANT are reviewed in section 4.2, but the main features are summarised below.

(1) It is a protein with a well defined and important function. It mediates the exchange of ATP and ADP between mitochondrial matrix and cytosol and is the exclusive means by which the nucleotide pools of these two compartments are linked. Its specificity for adenine nucleotides

determines the specificity of ATP synthase, Klingenberg (1985), and it is a point where control may be exerted over the rate of oxidative phosphorylation. In higher plants where there are developmental and diurnal fluctuations in the contributions of photophosphorylation and oxidative phosphorylation to cellular ATP levels, the role of the translocator may be particularly important.

(2) It is an abundant polypeptide comprising ca. 10% of the protein of the mitochondrial inner membrane as estimated from Coomassie blue staining, and therefore is likely to be represented in the library.

(3) The protein has been purified from beef heart and Neurospora crassa mitochondria, and its properties have been subject to extensive investigation (see Vignais 1976, Klingenberg and Heldt 1982 for reviews). Kinetic and inhibitor data from maize mitochondria (Earnshaw 1977), imply that the plant protein has similar properties to the better characterised mammalian and fungal translocators.

(4) DNA and protein sequence data show that the translocator has regions of high amino acid homology between such diverse organisms as Neurospora and cow (Arrends and Sebald 1984, Aquila et al., 1982).

(5) Neither higher plant chloroplasts or E.coli have a homologous protein. Higher plant chloroplasts have an adenine nucleotide exchange activity (Heldt 1969) but the characteristics differ from that of the mitochondrial translocator and it is unlikely to be the same protein. Most of the ATP synthesised in chloroplasts is exported by the extremely active triose phosphate carrier, which exchanges phosphoglycerate and dihydroxyacetone phosphate (Heber 1974).

### 3.7.1 Deriving the hybridisation conditions.

The melting temperature of a DNA-DNA duplex under defined conditions

can be calculated according to the formula

$$T_m = [ 81.5 + 16.6 \log_{10} M + 0.41(GC) - 0.72F ] ^\circ C.$$

where  $T_m$  = the melting temperature

M = The molar concentration of sodium

(GC)= %GC composition of the DNA

F= % of formamide in the hybridisation buffer.

(Howley et al., 1979)

From the nucleotide sequence of the Neurospora gene (Arrends and Sebald 1984), the percentage GC content of the coding region is 58.5%. This gives a  $T_m$  in 50% formamide and 5xSSC of 68.1 $^\circ$ C.

Comparison of the amino acid sequences of the beef and Neurospora translocators gives an overall amino acid sequence homology of 50%, although there are two regions of the polypeptide where 15 or more consecutive amino acids are identical. In these regions the nucleotide sequence homology must be at least 66% taking into account the redundancy of the genetic code. Given that a 1% mismatch reduces  $T_m$  of a heteroduplex by 1 $^\circ$ C (Bonner et al., 1973), a maximum temperature of 34.1 $^\circ$ C should allow these regions to hybridise. Since the degree of homology between the maize gene and the yeast probe was unknown, a hybridisation temperature of 25 $^\circ$ C, which corresponds to an overall nucleotide sequence homology of 57%, was employed. This is similar to the level of nucleotide sequence homology shown to exist between maize and yeast mitochondrially encoded genes (Fox and Leaver 1981, Dawson et al., 1984, Isaac et al., 1985a).

#### MAIZE MITOCHONDRIAL GENE

	COB	COX I	COX II
% homology to <u>S.cerevisiae</u>	52	58	56



### 3.7.2 Identification of a cDNA clone, pANT-1, encoding the translocator

The maize cDNA library was screened by hybridisation with a 2.6kb Bam H1 fragment of yeast DNA which contains the entire translocator gene (O' Malley et al., 1982). The initial colony hybridisations were ambiguous, with a number of colonies binding the probe to different extents. Forty four of the more intensely hybridising colonies were picked and crude nucleic acid extracts were prepared by the method of Barnes (1977). The plasmid DNA was separated from E.coli chromosomal DNA and RNA by electrophoresis through 1%(w/v) agarose gels in TAE buffer, transferred to nitrocellulose and hybridised again with the same probe under the same conditions, Fig.3.3. Colony 12/1 G2 is clearly positive. Plasmid DNA was prepared from this colony by the method of Birnboim and Doly (1979) and digested with Pst 1. The digested DNA was fractionated by electrophoresis in a 1% (w/v) agarose gel and transferred to nitrocellulose. The probe hybridised specifically to the ca.1150bp Pst 1 fragment representing the inserted cDNA, Fig.3.4. This cDNA clone was designated pANT-1. Sequence analysis (section 4.3) subsequently confirmed this to be a clone for the ANT.

### 3.8 OTHER cDNA CLONES RELATED TO pANT-1.

The filter replicas of the cDNA library were recycled to remove bound probe as described by Gergen et al., (1979), and rehybridised with an M13 clone derived from pANT-1, containing a 225bp Sau 3A fragment designated S16, which is contained entirely within the coding region of the gene, towards the carboxy terminus, Fig.3.5. This M13 clone had been sequenced in its entirety and shown to contain no other sequences apart from those derived from the translocator gene. The clone was labelled by second strand synthesis to a specific activity of  $4 \times 10^7$  dpm/ug and hybridised with the filters under 'homologous' conditions but washed

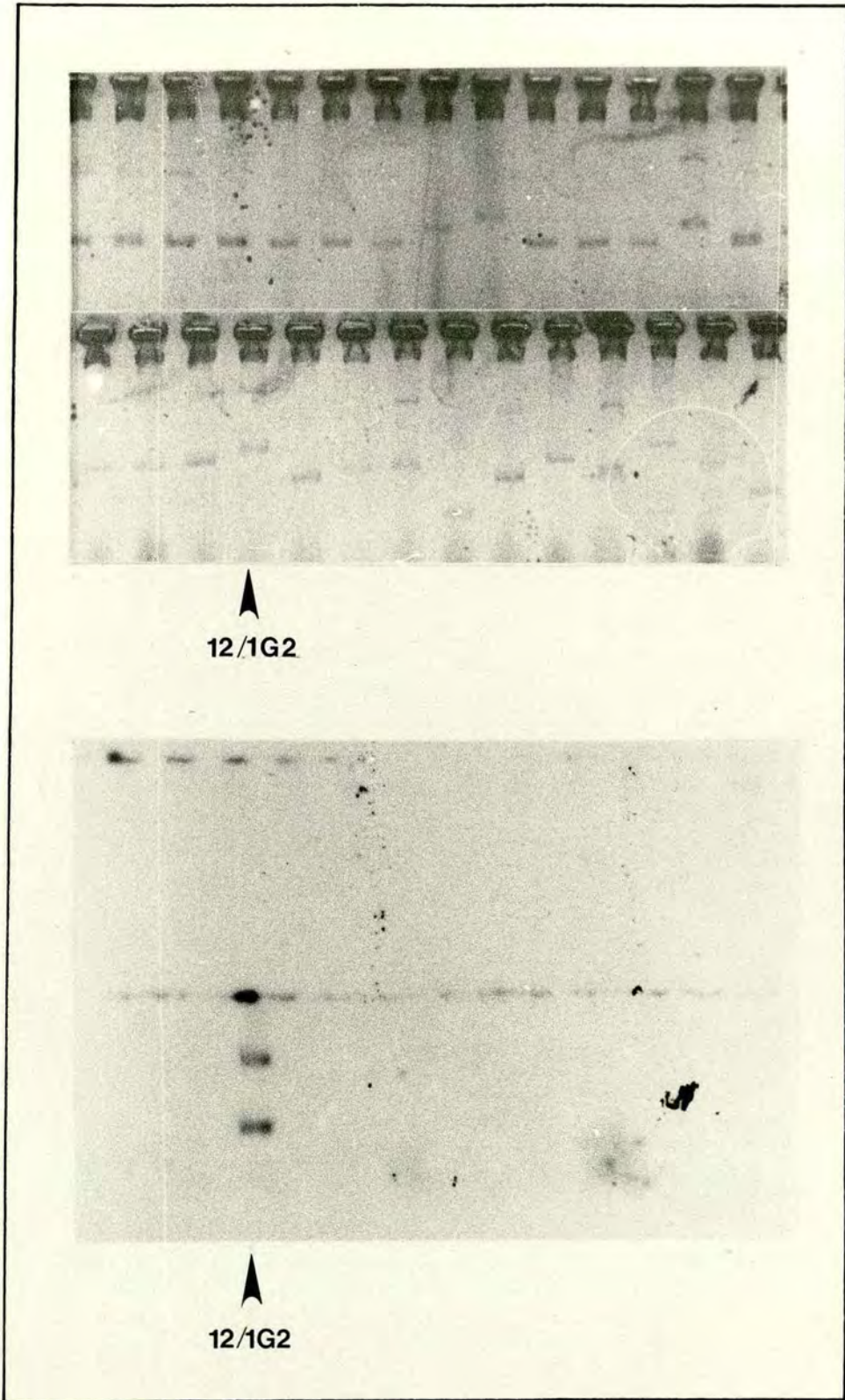


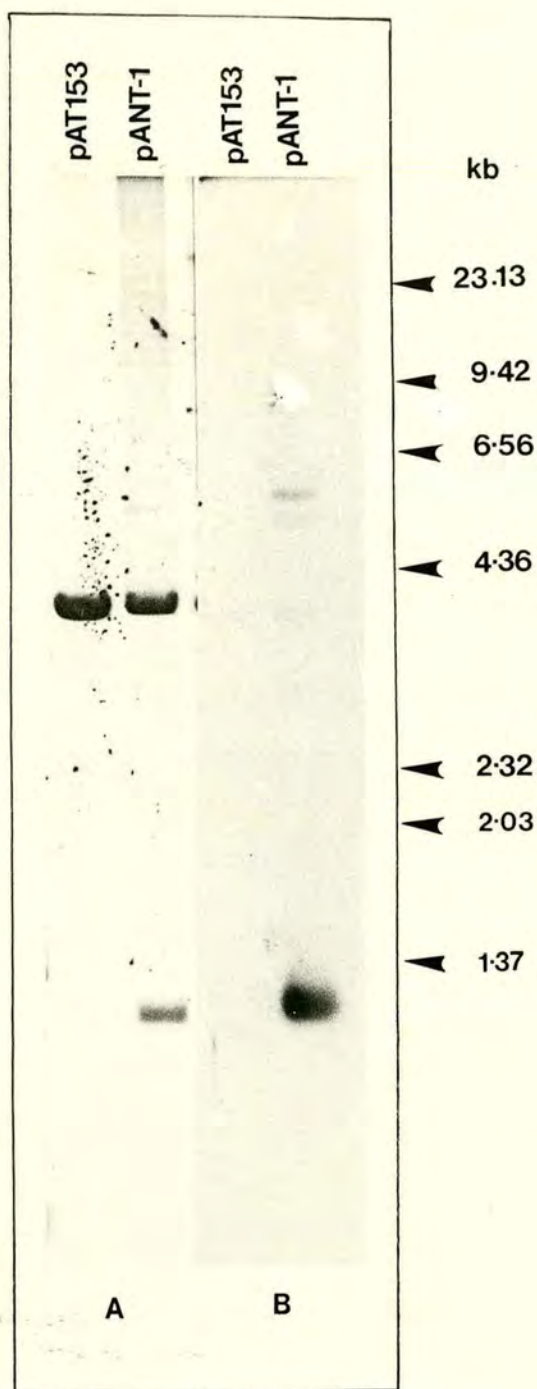
Fig. 3.3 Identification of pANT-1 by hybridisation with the yeast ATP/ADP translocator

Upper panel: 1%(w/v) agarose gel of crude nucleic acid extracts from single colonies from the maize cDNA library visualised by ethidium bromide staining.

Lower panel: autoradiograph of the gel shown above. The DNA was transferred to nitrocellulose and hybridised with the  $^{32}\text{P}$  labelled 2.6kb BamHI fragment from pYeOP1(2.6) in 50% formamide, 5xSSC, 10x Denhardt's solution, 150ug/ml herring sperm DNA at 25 $^{\circ}$ C. Washes were in 2xSSC at room temperature. Clone 12/1G2 (arrowed) = pANT-1.

Fig.3.3



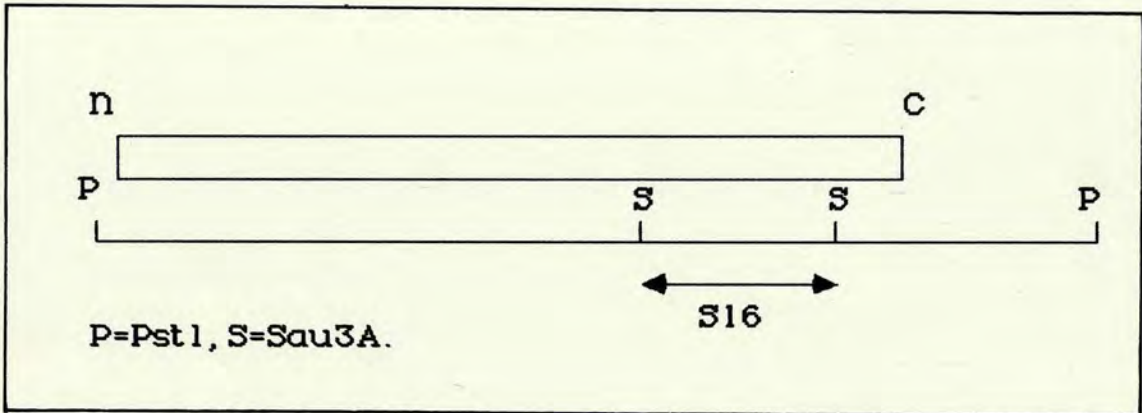
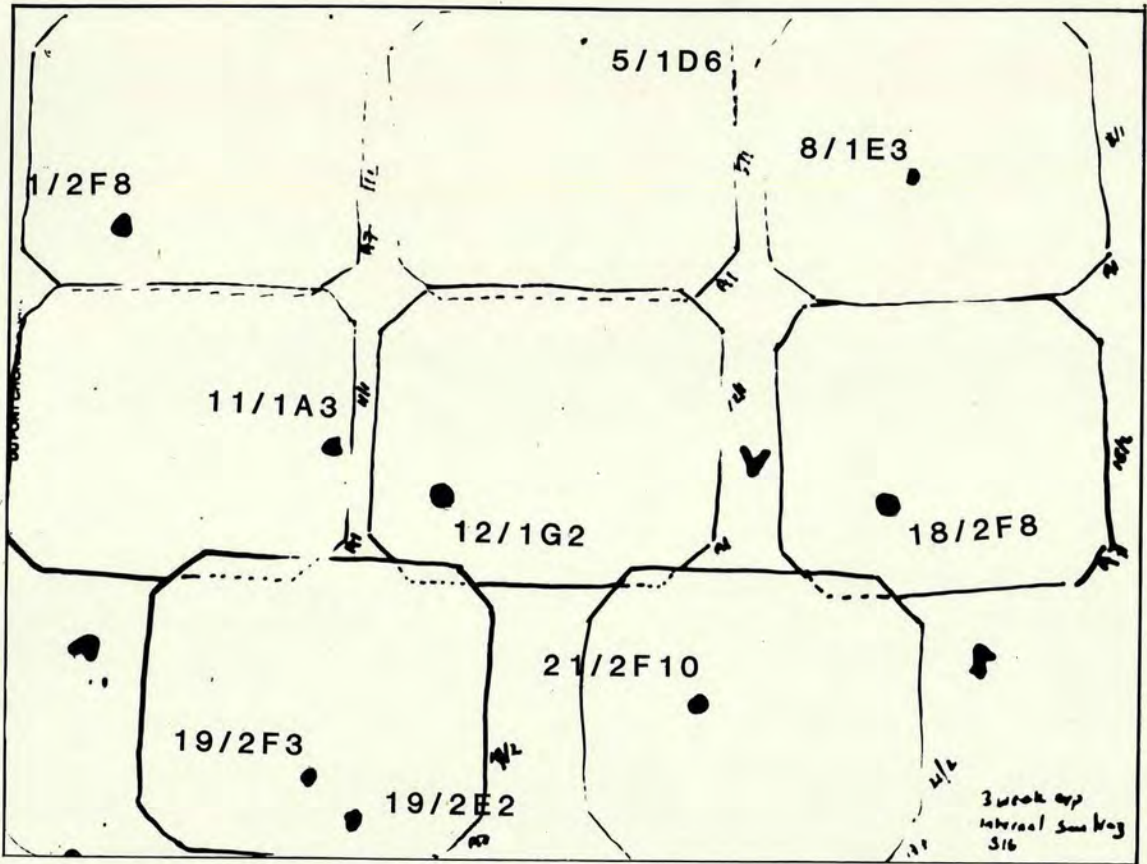


**Fig.3.4** Estimation of the size of the cDNA insert in pANT-1

Plasmids pAT153 and pANT-1 were digested to completion with PstI.

Left panel: Ethidium bromide stained 1%(w/v) agarose gel.

Right panel: Autoradiograph of the gel after transfer of the DNA to nitrocellulose and hybridisation with the 2.6kb BamHI fragment from pYeOP1(2.6). Hybridisation conditions are as described in the legend to Fig.3.3. Size markers are from a HindIII digest of  $\lambda$ cl857, and Sau3A digest of pBR322.



**Fig.3.5** Identification of additional ANT clones by colony hybridisation

Clone S16 was labelled with  $^{32}\text{P}$  by second strand synthesis (sp. act.  $4 \times 10^7 \text{dpm}/\mu\text{g}$ ) and hybridised with the filters under homologous conditions ( $4 \times \text{SSC}$ ,  $10 \times \text{Denhardt's}$   $0.1\%(\text{w/v})\text{SDS}$ ,  $250 \mu\text{g}/\text{ml}$  herring sperm DNA) at  $62^\circ\text{C}$ . Washes were in  $2 \times \text{SSC}$  at room temperature.

Upper panel: Autoradiograph of the filters.

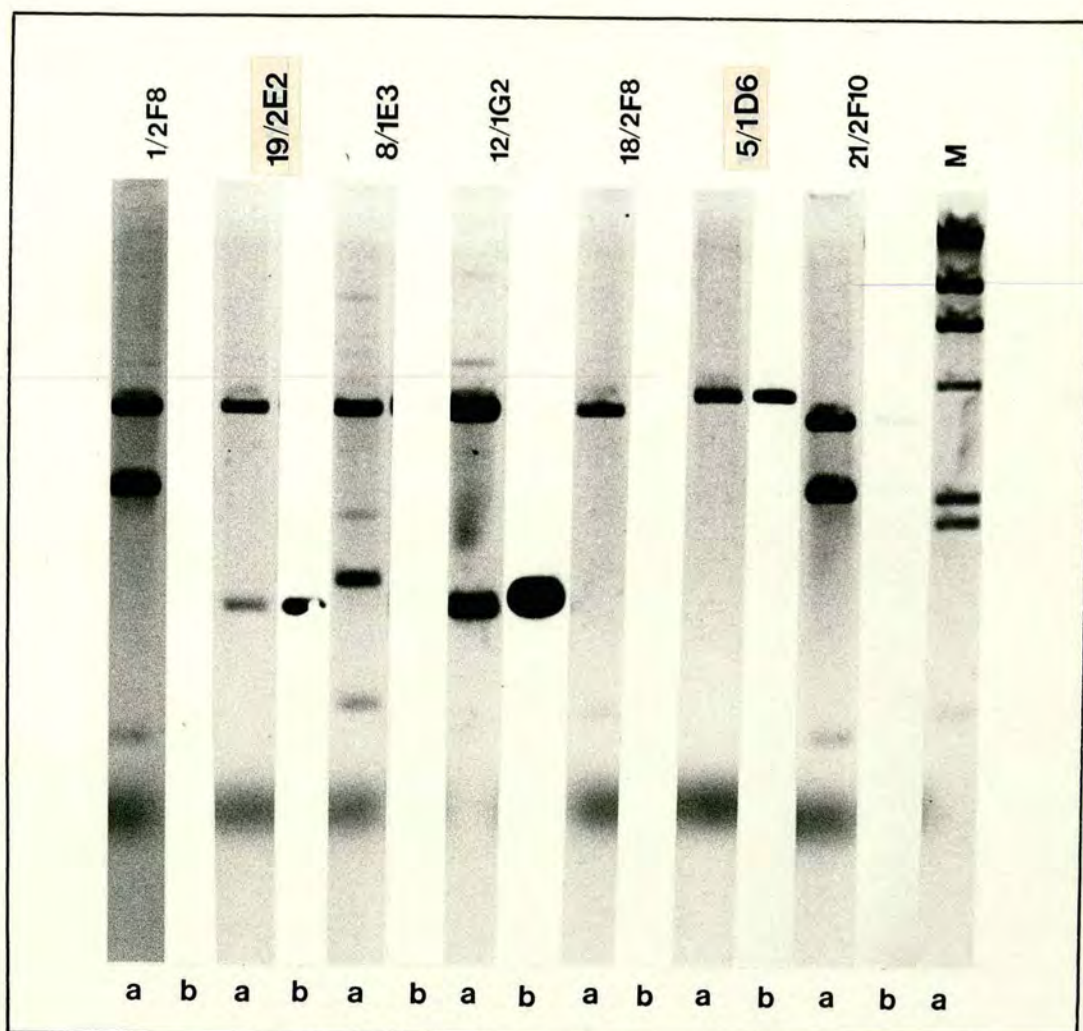
Lower panel: Origin of the S16 subclone of pANT-1

relatively non-stringently. The autoradiograph of the filters is shown in Fig.3.5.

Nine intensely hybridising colonies were observed, while two others gave weaker signals but were significantly above background. Plasmid DNA was prepared from single colonies of the clones by the method of Birnboim and Doly (1979), digested with restriction endonuclease PstI and separated by electrophoresis through a 1%(w/v) agarose gel in TAE buffer. The DNA was transferred to nitrocellulose by capillary blotting (Southern, 1975) and hybridised with the S16 probe (ca.  $2 \times 10^7$  dpm/ug)

Only three of the clones hybridise to this probe, Fig.3.6; 5/1D6 (=pANT-2) the insert of which cannot be excised by PstI, 12/1G2 (=pANT-1) which serves as a positive control in this experiment, and 19/E2 (=pANT-3) which contains a similar sized PstI insert to pANT-1. The reason why only three out of the nine positives from the colony hybridisation should hybridise to the probe in this experiment is unclear, however this result was repeatable with subsequent plasmid preparations from the same colonies.

Plasmid DNA from these three clones was digested with the restriction endonucleases AluI and HaeIII, and electrophoresed through a 6% polyacrylamide gel alongside similar digests of the vector, pAT153, Fig.3.7. All three clones clearly contain different restriction fragments as well as many common ones. Some differences may arise as the result of the cloned cDNA's being of different lengths, therefore fragments containing both vector and insert sequences will vary in size, and other fragments may be missing altogether due to cDNA synthesis terminating before the 5' end of the mRNA is reached. There is also the possibility, that these three clones may show restriction site polymorphism which implies that they are the transcripts of different but related genes.



**Fig.3.6** Characterisation of the additional ANT cDNA clones.

Plasmid DNA was extracted from each of the colonies identified in Fig.3.5 (11/1A3 and 19/2F3 failed to grow), digested with *Pst*I and separated by electrophoresis in a 1%(w/v) agarose gel. The DNA was transferred to nitrocellulose and hybridised with the S16 probe in 3xSSPE, 10xDenhardtts, 0.1%(w/v)SDS, 50ug/ml herring sperm DNA at 65°C, and washed in 2xSSPE, 0.1%(w/v)SDS at 65°C. (a) = ethidium bromide stained gel, (b) = autoradiograph. M =  $\lambda$ HindIII size markers (23.13, 9.42, 6.56, 4.34, 2.32, 2.03, 0.56kb)





**Fig.3.7** Comparison of the restriction endonuclease digest patterns of the cDNA clones pANT-1, pANT-2, and pANT-3.

6% polyacrylamide gel stained with ethidium bromide. Lanes A-D contain DNA from pANT-1, pANT-2, pANT-1, and pANT-3 respectively digested with AluI. Lanes E-H contain DNA from the same clones digested with HaeIII.

### 3.9 DISCUSSION.

The cDNA library has been shown to contain clones for nuclear encoded mitochondrial polypeptides, and can be used for the isolation of additional genes in the future. Heterologous hybridisation has proved a useful tool for identifying genes where a probe for a homologous gene is available from another organism. This however, along with the reliance of a reasonably high degree of nucleotide sequence homology between the two genes from different organisms, is its major limitation. There may be components of higher plant mitochondria which are not homologous to yeast or mammalian mitochondria. The genes involved in the restoration of cytoplasmic male sterility are an example. Other polypeptides may have counterparts in other organisms which have an insufficient degree of sequence homology to detect by this relatively insensitive method. For example the small nuclear encoded subunits of complexes III and IV of the respiratory chain are not so well conserved between different species. Thus in order to identify less well conserved or unique plant genes alternative methods of identifying clones will have to be developed.

**CHAPTER 4. THE MITOCHONDRIAL ADENINE NUCLEOTIDE  
TRANSLOCATOR: ANALYSIS OF CLONES FROM A MAIZE cDNA  
LIBRARY.**

**4.1 RATIONALE.**

Mitochondrial biogenesis is dependent upon contributions from both the nucleocytoplasmic and the organellar genetic machinery. A complete understanding of mitochondrial biogenesis requires an knowledge of the location and expression of the genes encoding mitochondrial proteins, the mechanisms by which the activity of the two genetic systems are co-ordinated and regulated, and the way in which the components of different genomic origin are directed to and assemble into a biologically active form. Of equal importance is an understanding of the way in which mitochondrial biogenesis and activity is integrated with other cellular processes that may be dependent upon, and in turn influence, these events.

In the case of higher plants, a number of mitochondrially encoded genes have been isolated, sequenced and subjected to transcriptional analysis (Fox and Leaver 1981, Hiesel and Brennicke 1983, Bonen et al., 1984, Kao et al., 1984, Dawson et al., 1984, Boer et al. 1985, Schuster and Brennicke 1985, Moon et al., 1985, Dewey et al., 1985, Isaac et al., 1985a, b) Physical maps of mt DNA's from Brassica campestris and Zea mays have been constructed ( Palmer and Shields 1982, Lonsdale et al., 1984 ), and the location of the genes determined on the maize map, (Dawson et al., in preparation). The results of these studies indicate that the organisation and expression of the plant mitochondrial genome, whilst sharing basic features with fungal and animal mitochondria, may be fundamentally different in a number of ways (section 1.5.1)

In contrast, comparatively little attention has been paid to the nuclear

genes which encode ca.90% of the proteins in plant mitochondria. This chapter describes the characterisation of a cDNA clone for a nucleary encoded mitochondrial polypeptide from maize, the adenine nucleotide translocator (ANT).The isolation and characterisation of clones encoding this polypeptide will provide a starting point for a detailed study of the role of nuclear genes in mitochondrial biogenesis in higher plants.

## 4.2 THE ADENINE NUCLEOTIDE TRANSLOCATOR.

### 4.2.1 Introduction.

The mitochondrial inner membrane contains a number of specific protein carriers which permit the controlled exchange of otherwise impermeant molecules between the matrix space and cytosol (reviewed by Wiskich 1977, LaNoue and Schoolwerth 1979). Fig.4.1. These carriers are the means by which the biochemical processes of mitochondria and the rest of the cell are integrated, while maintaining distinct environments within the different subcellular compartments. The adenine nucleotide translocator is the best characterised of these transport proteins and has been the subject of extensive research for two decades. Much information has been obtained concerning the structure and function of the protein from a variety of sources including mammals, fungi and plants (see Vignais 1976, Klingenberg and Heldt 1982, Klingenberg 1985 for reviews), but only in recent years have the increasingly powerful techniques of molecular biology been applied to the study of this polypeptide which is so vitally important to the energy economy of the eukaryotic cell.

### 4.2.2 The ANT polypeptide

The adenine nucleotide translocator is probably the most abundant mitochondrial protein. In a variety of species it has a  $M_r$  of 30,000-33,000

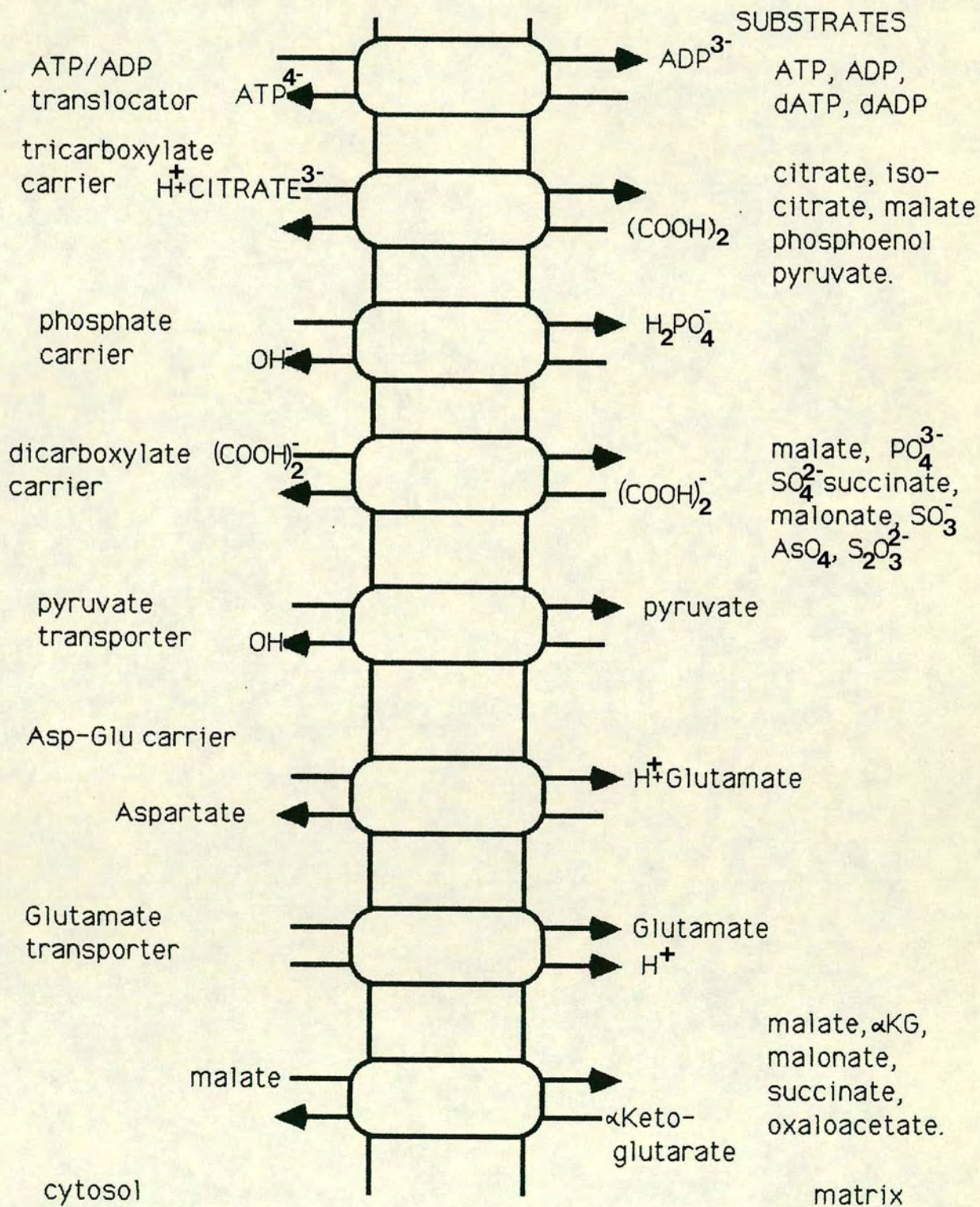


Fig.4.1. Transport systems of the mitochondrial inner membrane.

daltons as estimated from SDS-PAGE and in maize mitochondria is 5-10% of the total protein as estimated by Coomassie blue staining. This is in good agreement with the figure of 13% quoted for beef heart mitochondria (Klingenberg, 1985). The functional protein is a dimer which spans the inner membrane, and there is one binding site per dimer for adenine nucleotides and inhibitors of transport. The translocator can be readily purified from mitochondria (Kramer *et al.*, 1977 Aquila *et al.*, 1978), and has been reconstituted into phospholipid vesicles of defined composition (Kramer and Klingenberg 1977) The amino acid sequence of the beef heart mitochondrial protein has been determined directly (Aquila *et al.*, 1982). It contains 297 amino acids and has a calculated molecular mass of 38,870 daltons. The nucleotide sequence of the cloned cDNA and gene have been determined from Neurospora crassa (Arrend and Sebald 1984) and now from a higher plant, Zea mays L. (this thesis).

The data available concerning the structural and functional properties of the protein can now be correlated with the predicted primary and higher order structure, to yield a more complete picture of the adenine nucleotide translocator as a functional entity. In addition the availability of cloned gene sequences and antibodies to use as probes will enable the organisation and expression of the gene(s) encoding the translocator to be studied, and should add considerably to the understanding of its biogenesis.

#### 4.2.3 Biogenesis of the adenine nucleotide translocator.

The vast majority (ca. 90%) of mitochondrial proteins are the products of nuclear genes. Several lines of evidence suggested that the adenine nucleotide translocator was one of these.

(1) Yeast petite mutants, which are respiration deficient as a result of

the loss of large portions of their mt DNA, contain a wild type adenine nucleotide translocator as judged by its sensitivity to ATR and BKA (Groot et al., 1975).

(2) Mutations which affect transport of adenine nucleotides and yield strains which are unable to grow on non fermentable carbon sources (e.g. *op1*), are inherited in a Mendelian fashion (Kovac et al., 1967). More recently the gene for the adenine nucleotide translocator of *Saccharomyces cerevisiae* has been cloned by complementation of the *op1* mutation with plasmids containing restriction fragments of wild type nuclear DNA, (O'Malley et al., 1982)

The nucleotide translocator is encoded in nuclear DNA, transcribed in the nucleus, the mRNA translated in the cytoplasm and the protein imported into mitochondria in a post translational manner. The translocator protein has been synthesised by translation of mRNA in a cell free system and imported into *Neurospora crassa* mitochondria in vitro (Zimmerman et al., 1979, Zimmerman and Neupert 1980). The precursor synthesised in vitro has the same  $M_r$  on SDS-PAGE as the mature protein (33,000 daltons), suggesting that it is not subjected to extensive proteolytic processing upon import. This is in contrast to many other imported mitochondrial inner membrane proteins which are synthesised with an amino terminal 'pre-piece' shown to be important for targeting to the mitochondrion. Import of the translocator shows the same sensitivity to uncouplers of mitochondrial oxidative phosphorylation as do other proteins, indicating that a proton motive force across the inner membrane is required.

#### 4.2.4 Substrate specificity.

Early experiments indicated that when radioactively labelled ATP or ADP was added to mitochondrial suspensions they rapidly equilibrated with the endogenous adenine nucleotide pool (Pfaff *et al.*, (1965), demonstrating the presence of a specific transport system. By pre-loading mitochondria with [<sup>3</sup>H] ADP, then adding exogenous [<sup>14</sup>C] ADP, it could be shown that the stoichiometry of exchange is 1:1. The specificity of the system is extremely high. Only ADP and ATP are transported efficiently; dATP and dADP are slightly active whilst AMP is almost entirely excluded. Other nucleotide di- and tri- phosphates are also inactive in transport, (Pfaff and Klingenberg 1968). The substrate specificity of the adenine nucleotide translocator is higher than that of ATP synthase, so it is the former which determines the high substrate specificity of oxidative phosphorylation in intact mitochondria (Klingenberg, 1985).

#### 4.2.5 Effects of energisation of the inner membrane on adenine nucleotide transport.

Adenine nucleotide transport is not energy dependent per se, as the binding energy of the substrate provides the energy for translocation. In uncoupled mitochondria ATP and ADP are exchanged in all possible combinations e.g. ATP for ATP, ATP for ADP, ADP for ATP and ADP for ADP, to an extent dependent on the concentrations of the nucleotide species in the different compartments. However, in coupled mitochondria, the presence of a membrane potential, such that the matrix is negative with respect to the outside, allows the export of ATP against its concentration gradient due to its extra negative charge in relation to ADP, and this accounts for ca.75% of the exchange (Klingenberg 1985). Purified



translocator protein reconstituted into phospholipid vesicles also shows membrane potential dependent control of the exchange of adenine nucleotides, (Kramer and Klingenberg, 1980).

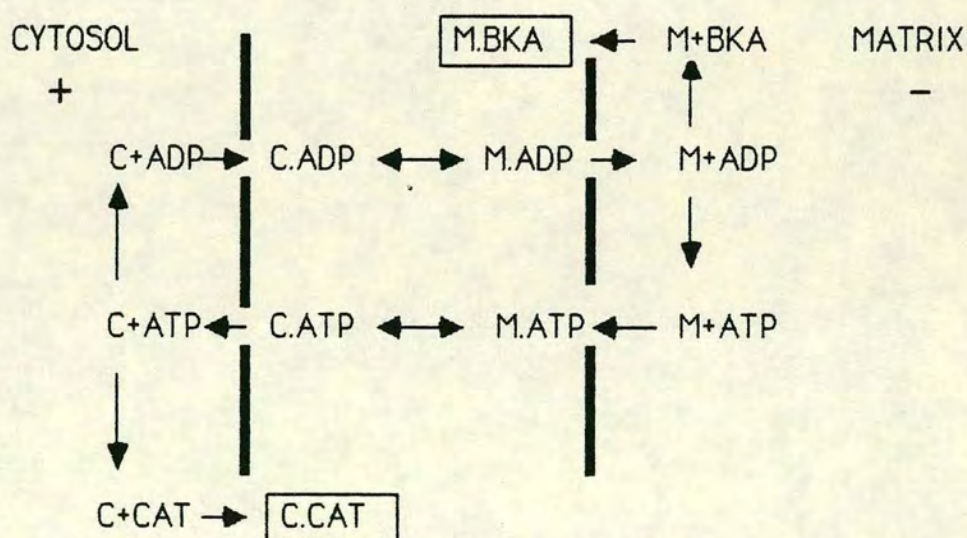
#### 4.2.6 Inhibitors of adenine nucleotide translocation.

Considerable insight into the mechanism of adenine nucleotide translocation has been gained from studies with the specific inhibitors atractyloside (ATR),(Chappell and Crofts 1965) carboxyatractyloside (CAT), and bongkreikic acid (BKA) (Henderson and Lardy 1970). These compounds inhibit adenine nucleotide translocation in mitochondria from all sources, but higher concentrations are required to achieve the same degree of inhibition of the plant mitochondrial translocator (Earnshaw 1977). ATR and CAT interact with the translocator from the cytosolic side and displace bound ADP. The inhibition of translocation by CAT is virtually irreversible, while ATR can be displaced by adenine nucleotides. Bongkreikic acid inhibits nucleotide transport after a time lag and in a pH and temperature dependent fashion, implying that it is required to cross the membrane in order to reach its site of action. In contrast, inhibition of adenine nucleotide transport in submitochondrial particles (where the membranes form inside out sealed vesicles) is immediate and not dependent upon temperature or pH. These results indicate that the translocator has an asymmetric orientation within the membrane, with different binding sites exposed on opposite sides of the membrane. Other compounds which inhibit the translocator are long chain acyl CoA derivatives. These act as competitive inhibitors, and may play a role in the regulation of ATP/ADP exchange in vivo, (Morel et al., 1974). All the compounds which inhibit the adenine nucleotide translocator have three net negative charges, as does ADP. This suggests that a cluster of three

positive charges may be a salient feature of the binding site created by the protein.

#### 4.2.7 Mechanism of adenine nucleotide transport.

The ANT is postulated to be able to alternate between two different conformations. In the 'c' state the nucleotide binding site is oriented towards the cytosol, where it can interact with adenine nucleotides or inhibitors such as CAT or ATR. Upon binding an adenine nucleotide, the protein undergoes a conformational change, resulting in the transport of the nucleotide and the reorientation of the binding site to the matrix. This is the 'm' state and is defined by the ability of the translocator to bind BKA. In the 'm' state the ANT can bind matrix adenine nucleotides and transport them to the cytosol. CAT and BKA are able to 'lock' the translocator in the 'c' and 'm' states respectively, which permits the investigation of their individual properties. The transport cycle is illustrated in Fig.4.2.



SCHEMATIC DIAGRAM SHOWING THE OPERATION OF THE TRANSLOCATOR IN THE 'PRODUCTIVE' MODE.

There is considerable evidence for changes in conformation of the protein upon binding and transport of adenine nucleotides. These include:

- 1) Different immunological cross reactivity of the translocator in the "m"- and "c"-states (Buchanan et al., 1976).
- 2) Changes in sensitivity of the protein to proteases (Aquila et al., 1978)
- 3) Unmasking of -SH groups sensitive to alkylation upon transition between states (Klingenberg and Appel 1980)
- 4) Unmasking of arginine residues important for transport (Klingenberg and Appel 1980, Block et al., 1981)
- 5) Changes in circular dichroism spectra (Klingenberg, 1985)
- 6) Changes in the accessibility of some lysine side chains to modification with pyridoxal phosphate.(Bogner et al., 1983)
- 7) Different binding of fluorescent adenine nucleotide derivatives to the different states of the translocator (Klingenberg et al., 1984)
- 8) Differential sensitivity to ultra violet light (Block et al., 1979)

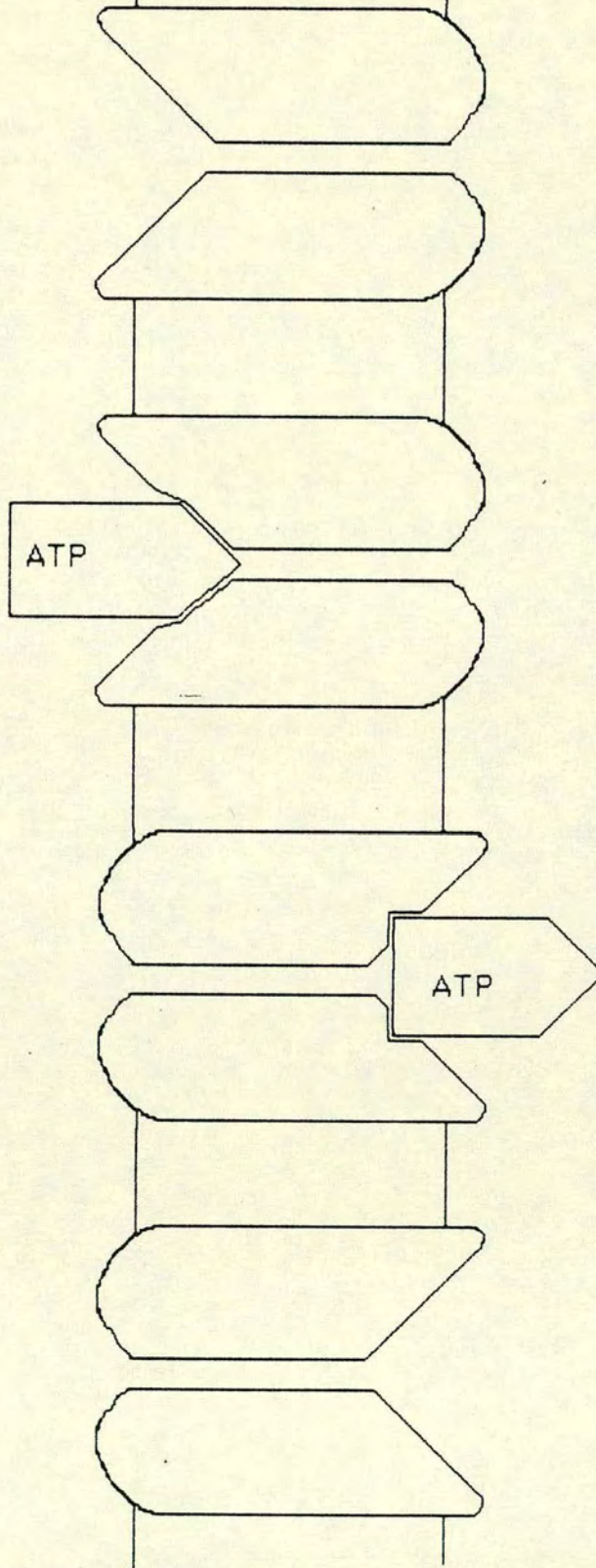
As one molecule of inhibitor is bound per dimer, both subunits are postulated to contribute to the binding site, perhaps by forming a relatively hydrophilic 'pore' through the membrane. Access to the pore would be permitted from one side of the membrane only, depending on the orientation of the translocator, Fig.4.3.

Several attempts have been made to identify residues involved in the binding site. Walker et al.,(1982) have identified a sequence in the beef heart mitochondrial translocator which may be part of a common nucleotide binding domain shared with other nucleotide binding proteins, while Bogner et al.,(1983) have studied the accessibility of lysine residues in the presence and absence of bound nucleotides and inhibitors, and Block et al., 1979, 1981 and Klingenberg et al., 1984, have used UV and

INTER MEMBRANE  
SPACE

MATRIX

C-State



m-State

Fig.4.3 "Gated pore" model of adenine nucleotide transport  
(After Klingenberg, 1985)

fluorescent probes. However as yet there is no definitive answer.

#### 4.2.8 Adenine nucleotide translocation and the control of respiration.

Mitochondrial oxidative phosphorylation is the major source of ATP in aerobically growing eukaryotic cells. Most of the enzymic processes which utilise ATP occur outside the mitochondria. In addition many key enzymes in metabolic pathways are regulated via the levels of adenine nucleotides e.g. phosphofructokinase, pyruvate kinase. The adenine nucleotide translocator forms the exclusive link between the ATP producing and ATP consuming processes. However, whether transport of adenine nucleotides is the rate limiting step in respiration is highly controversial.

Isolated mitochondria in the absence of ADP consume little oxygen and produce little ATP (state 4 respiration). Upon the addition of ADP to the incubation medium it is transported into the mitochondria via the translocator, oxygen consumption is increased and ATP is synthesised (state 3). The ratio of state3/state4 is the respiratory control ratio, and is a measure of the integrity of the mitochondria. However in vivo many factors may influence rates of respiration and so the resting and active states of isolated mitochondria may be misleading. Recent experiments (Duszynski et al., (1982), Groen et al., (1982), Tager et al., (1983), Baggetto et al., (1984)) have attempted to answer this question by measuring the control strength (Kacser and Burns 1973) of various steps in respiration. The sum of the control strengths of all the steps in a metabolic pathway is 1, and the higher the control strength of any individual step, the greater its control over flux through the pathway. The control strengths of the different steps varies according to the rate of respiration (Groen et al., 1982). In state 4 most of the control is exerted via the passive permeability of the inner membrane to protons ('proton leak'). However in

state 3 the adenine nucleotide translocator (control strength 0.29) and the dicarboxylate carrier (control strength 0.33) make the most significant contribution to the control of respiration rate. The conclusions drawn from these experiments is that the adenine nucleotide translocator exerts a significant control over state 3 respiration but is not the sole rate determining step. In higher plants the role of the translocator may be particularly important during light-dark transitions, as the cytosolic energy charge, and hence the availability of ADP, is an important way of controlling the rate of phosphorylation by mitochondrial ATP synthase.

#### 4.2.9 Adenine nucleotide transport in chloroplasts

Chloroplasts also contain an adenine nucleotide translocation activity, (Heldt, 1969), but it differs from the mitochondrial one in that (i) it is insensitive to CAT, (ii) it is inhibited by uncouplers, (iii) it shows a marked preference for ATP as a substrate and (iv) the measured translocation rate is much slower than the mitochondrial translocator. It is insufficiently active to export the ATP formed by photophosphorylation to the cytoplasm, and probably functions to supply ATP to the chloroplast in the dark. Recently a CAT insensitive adenine nucleotide translocator has been reported from rat liver mitochondria (Austin and Aprille 1984). It is similar to the chloroplast nucleotide translocator in that ATP is the preferred substrate. The translocation rate is an order of magnitude less than the CAT sensitive translocator, and its function is unknown. It is not known whether there is any relationship between this protein and the chloroplast nucleotide translocator.

#### 4.3 SEQUENCE ANALYSIS OF cDNA CLONES ENCODING THE ADENINE NUCLEOTIDE TRANSLOCATOR OF MAIZE.

The clone pANT-1 was identified in a maize cDNA library using the homologous nuclear gene from yeast as a probe in colony hybridisations (section 3.7.2). In the absence of an antibody to confirm the identity of the clone by hybrid select translation (Ricciardi *et al.*, 1979) and immunoprecipitation, and in order to obtain information concerning the degree of similarity between plant and animal nuclear genes encoding mitochondrial proteins, the nucleotide sequence of the clone was determined.

##### 4.3.1 Re-cloning the 1.2kb Pst I fragment of pANT-1 into M13mp8.

By cloning the entire cDNA into M13 in both orientations it was hoped that the nucleotide sequences at the ends of the clone could be determined very rapidly. Thus allowing positive identification of the clone as the ATP/ADP translocator and allowing an estimation to be made of the extent of the mRNA copied into cDNA.

Plasmid DNA from pANT-1 was digested to completion with restriction endonuclease PstI and ligated to PstI digested mp8 DNA. White 'plaques' from the resulting transformation were plaque purified, and single stranded and double stranded replicative form (RF) DNA prepared from a total of 12 clones. The RF DNA was digested with PstI and subjected to electrophoresis through a 1% agarose gel. None of these clones contained an insert which co-migrated with the cDNA insert of pANT-1, although several phage contained inserts identical in size to the vector pAT153. A number of clones contained no detectable inserts.

The 1.2kb PstI fragment of pANT-1 was purified by electroelution from a preparative agarose gel and ligated to PstI digested mp8. Replicative

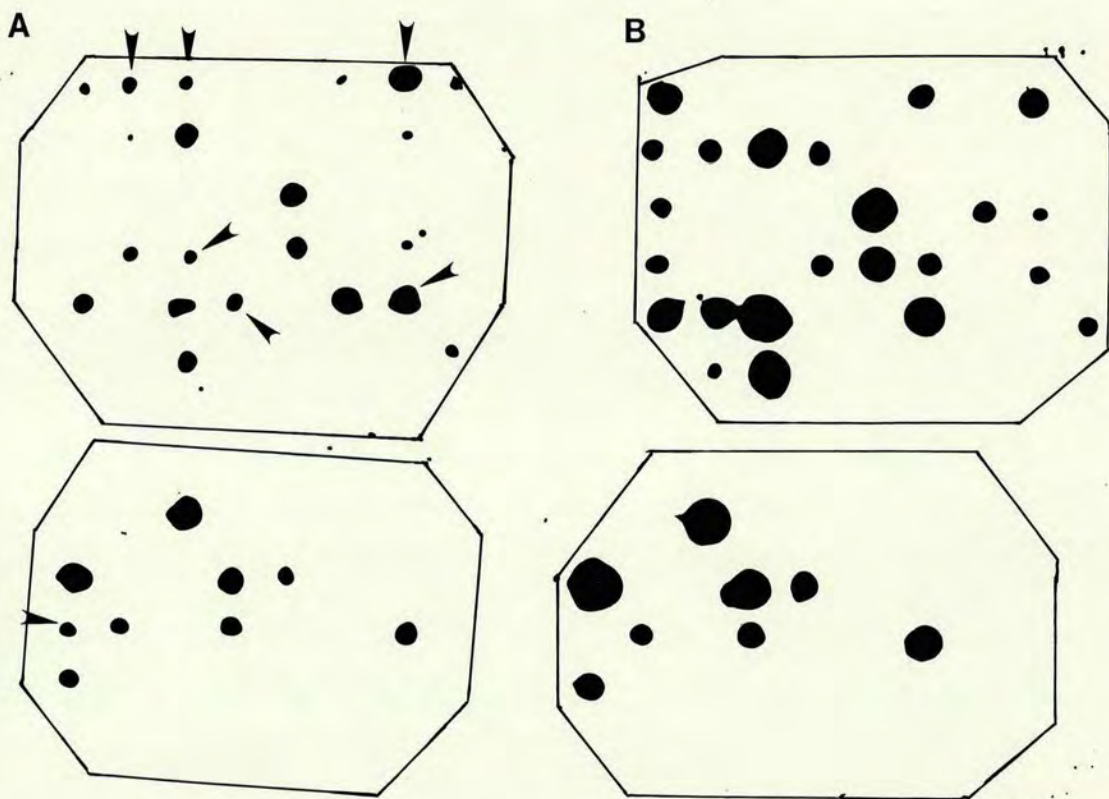
form DNA was prepared from 10 recombinant phage and digested with PstI as before. Again none of the recombinants contained a 1.2kb insert.

The inability to clone the 1.2kb PstI fragment suggests that it may not be stably maintained in M13. This may be due to the presence of the homopolymer tails at either end of the cDNA, which might lead to deletion of the insert via homologous recombination between these sequences. (JM101, the M13 host strain used, is recombinationally active.) This might explain the recovery of phage with the recombinant phenotype, i.e. white plaques, but which apparently lacked an insert. Paradoxically, the clone pANT-1 is stably maintained in E. coli strain 294 which is also recA<sup>+</sup>.

#### 4.3.2 'Shotgun cloning' restriction fragments of pANT-1 into M13 vectors.

Comparison of restriction endonuclease digests of pANT-1 and pAT153 determined that there were restriction sites for AluI, HaeIII, TaqI and Sau3A within the 1.2kb insert of pANT-1. No sites were found for the hexanucleotide recognising enzymes EcoR1, HindIII, SalI, XhoI or BamHI. Plasmid DNA from pANT-1 was digested separately with each of the enzymes AluI, HaeIII, TaqI and Sau3A and ligated with the appropriately digested mp8 (SmaI for AluI and Hae III fragments, AccI for TaqI, and BamHI for Sau3A). Duplicate nitrocellulose filters were made from the phage plaques (section 2.6.3). One set was hybridised with nick translated pANT-1, and the duplicate with nick translated pAT153. Plaques which hybridise to pANT-1 but not with pAT153 are predicted to contain inserts derived from the cDNA, Fig.4.4. This screening strategy led to the rejection of clones which contain both pAT153 and cDNA sequences, i.e. those which span the PstI sites in pANT-1. In order to recover these fragments, the purified 1.2kb PstI fragment was nick translated and used as a probe.





**Fig.4.4** Screening recombinant phage by hybridisation.

White plaques were picked onto 6x8 grids and duplicate sets of filters made as described in materials and methods. One set (A) was hybridised with nick translated pANT-1 and the other (B) with nick translated pAT153. Hybridisation conditions were as described in the legend of Fig 3.5. Phage which hybridise to pANT-1 but not pAT153 (arrowed) contain inserts derived from the cDNA.

### 4.3.3 Forced cloning of the ends of the 1.2kb Pst I fragment.

The type of shotgun cloning strategy outlined above results in an initial very rapid accumulation of unique sequence, but the frequency with which new sequence is obtained declines as the accumulation of sequence data approaches completion and more recombinant clones have to be screened in order to detect the 'rare' clones not previously isolated. In addition some sequences may be under represented due to a lack of convenient restriction sites or as a result of instability in M13. Other clones may be overlooked initially as the recombinant phage may be blue. This can occur if a small DNA fragment is inserted in the lac gene such that the reading frame is maintained, or if the insert contains a sequence which is functional as a promoter in E.coli.

In the case of pANT-1, the major difficulty encountered was obtaining clones derived from the ends of the cDNA, containing the homopolymer tails. These clones were recovered less frequently than expected, and almost invariably in the orientation such that the plasmid sequences were adjacent to the sequencing primer binding site. The presence of the homopolymer tails causes the sequencing reaction to terminate, or to generate spurious bands ('laddering'), probably a result of secondary structure in the template, so no useful sequence could be determined distal to the tracts of polydG or polydC, Fig4.5. Attempts to use Bal31 nuclease digestion (Poncz et al., 1982) to remove the homopolymer tails was largely unsuccessful, as the activity of the enzyme could not be sufficiently controlled to remove just the homopolymer tracts and not adjacent sequences of interest. However, some Bal31 clones were obtained which were useful in providing overlaps between a number of small restriction fragments. The most fruitful solution to the problem of obtaining sequence from the ends of the insert proved to be the forced

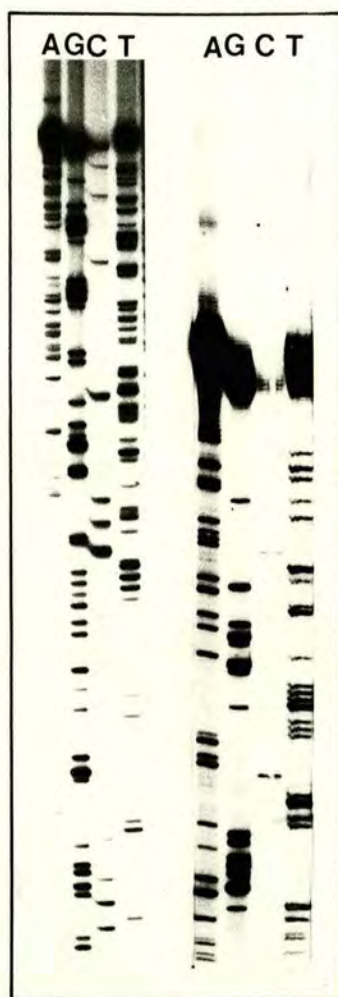


Fig.4.5 Premature termination and laddering caused by homopolymer tails  
Autoradiographs of 6%(w/v) polyacrylamide sequencing gels.

cloning of fragments generated by a double digest with PstI and another enzyme such that the PstI site and the homopolymer tails lay distal to the primer hybridisation site of M13, Fig.4.6. In this way the sequence could be determined up to the start of the homopolymer tails. The disadvantage of this method is that it only permits the sequence to be determined from one strand of the DNA. Although independently isolated clones generated with different enzymes were sequenced, ideally the sequence should be determined from both strands of the DNA throughout. In this case to do this would probably require the use of specific oligonucleotides as primers, or to use the chemical sequencing method of Maxam and Gilbert (1977). However more than 80% of the sequence was obtained from both strands, and the entire sequence was determined from multiple overlapping clones.

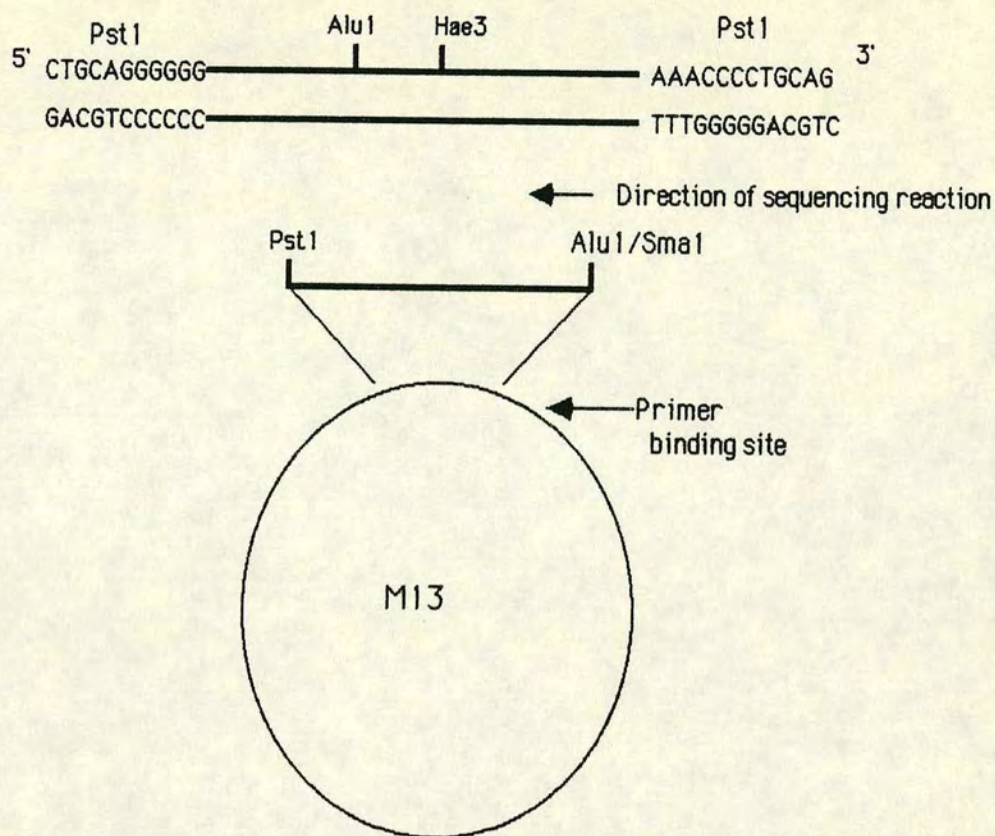
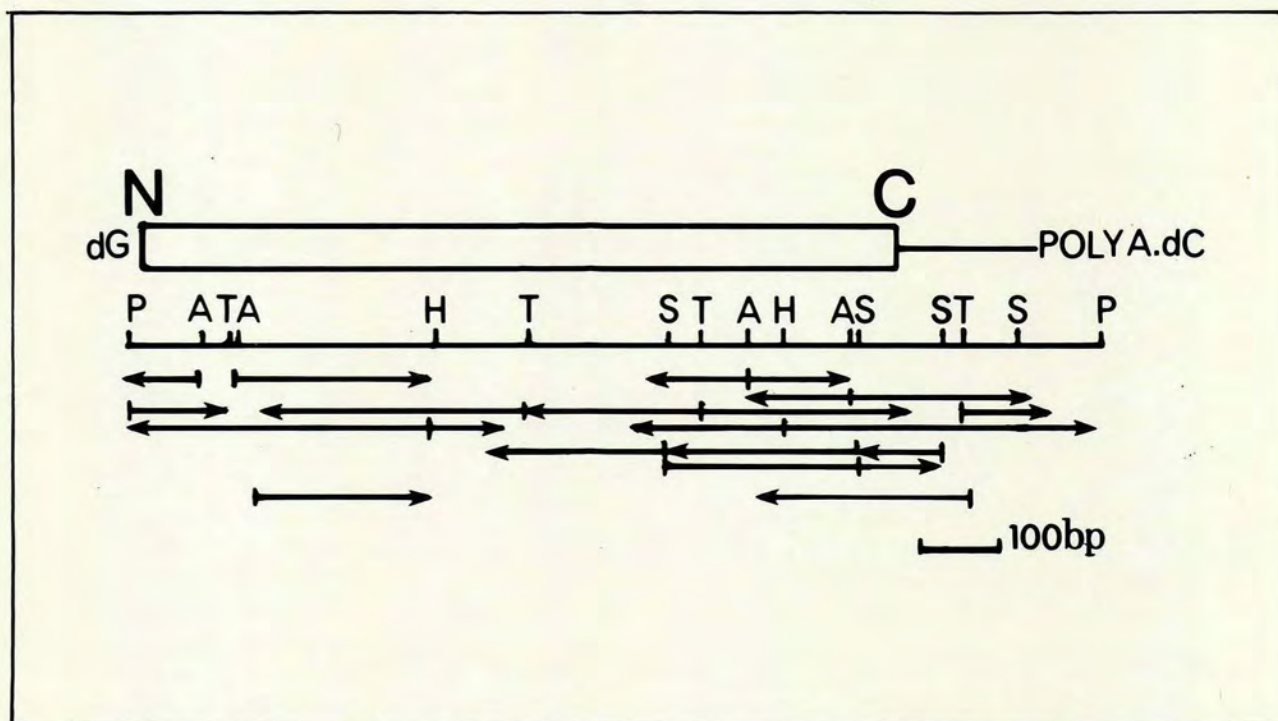


Fig.4.6. Diagram illustrating the principle of 'forced cloning'

#### 4.4 DISCUSSION

##### 4.4.1 Nucleotide sequence of pANT-1, a cDNA encoding the adenine nucleotide translocator of maize.

The nucleotide sequence of the entire 1.2kb insert of plasmid pANT-1 was determined from independent overlapping subclones covering the entire length. The sequencing strategy is illustrated in Fig.4.7 The cDNA contains an open reading frame of 954bp (318 amino acids) commencing with an ATG codon at position +3 and ending with the codon TAA at +957. There is 180bp of 3' untranslated sequence, and ca. 60 adenine nucleotides comprising the poly A<sup>+</sup>, Fig.4.8. The G+C content of the coding region is



**Fig.4.7** Restriction map of the cDNA insert of plasmid pANT-1 illustrating the strategy used for determining the nucleotide sequence.

The extent of the open reading frame (box), 3' untranslated region and poly A 'tail' are indicated in relation to the map. N and C denote the amino and carboxy termini of the polypeptide. P=PstI, A=AluI, T=TaqI, H=HaeIII, S=Sau3A.



Fig.4.8 Nucleotide sequence and predicted amino acid sequence of the maize adenine nucleotide translocator, derived from clone pANT-1.





48.9%, compared to 38.3% for the 3' untranslated region. The overall G+C content is 43.6%.

The 3' untranslated sequence does not contain an obvious polyadenylation signal conforming to the animal consensus sequence 5'-AATAAA-3', which is usually located 15-40 bases 5' to the point of polyA addition (Proudfoot and Brownlee 1976). There are two sequences of 5'-AATAAT-3', a variant of the consensus sequence found near to the point of polyA addition in some plant mRNA's (Messing et al., 1983). However, both of these sequences are rather remote from the point of polyA addition, 76 and 172 nucleotides 5' to the start of the polyA. A cDNA clone encoding malate synthase from cucumber also apparently lacks a polyadenylation signal (Smith and Leaver 1985), as do the genes encoding the small subunit of RuBP carboxylase from a variety of plant species, (Smith et al., 1983). It must be emphasised that none of these various sequences proposed as polyadenylation signals in higher plants have been shown to be sufficient or even necessary for correct polyadenylation. Several Saccharomyces cerevisiae mRNAs also lack 5'-AATAAA-3', and foreign genes do not require this sequence in order to produce polyadenylated transcripts when expressed in yeast, (Birnstiel et al., 1985)

#### 4.4.2 Amino acid sequence homology with translocator polypeptides from other organisms.

The deduced amino acid sequence of the maize ATP/ADP translocator exhibits a high degree of homology to previously published sequences for this protein from Neurospora crassa, Bos taurus and Saccharomyces cerevisiae (Arrends and Sebald 1984), Aquila et al., (1982) Adrian et al., in preparation.). When the sequences are aligned to maximise the homology between them, maize is 74.8% homologous to Neurospora 64.7% homologous

to yeast and 50.3% homologous to beef. In addition, many of the substitutions are conservative in nature, as demonstrated by the remarkable similarity of the hydropathy profiles of the proteins and the distribution of charged residues along the polypeptide chain. This is illustrated for the maize and N.crassa polypeptides in Fig.4.9.

When the deduced amino acid sequences are aligned as in Fig.4.10, the open reading frame encoding the maize polypeptide extends 5 amino acids beyond the amino terminus of the Neurospora protein, where a methionine occurs in the correct reading frame. It is not possible to determine whether this is the amino terminal methionine of the maize polypeptide as the cDNA clone terminates 2 nucleotides 5' to this ATG, and the homology between the three proteins falls off essentially to zero at the amino terminus. The maize translocator protein is acetylated at the N terminus (protein sequencing unit, University of Aberdeen), so it has not proved possible to determine the amino terminal sequence from the purified protein. The location of the amino terminal sequence of the protein and its implications in mitochondrial targeting and import will be discussed in relation to the genomic clone sequence in Chapter 5.

#### 4.4.3 Nucleotide sequences of other maize ANT cDNA clones.

Partial sequence analysis was carried out on cDNA clones pANT-2 and pANT-3 (section 3.8), using the same shotgun cloning strategy, in order to determine whether either of them contained additional sequences 5' to those present in pANT-1, and in addition to determine if all 3 clones represent transcripts of the same gene. The clone pANT-3 is identical in the region sequenced to pANT-1, but terminates at nucleotide +49 relative to the pANT-1 sequence. The remaining clone pANT-2 has a much smaller insert. It was not possible to calculate the exact size due to the loss of



Fig.4.9 Hydropathy plot of the maize and Neurospora ANT polypeptides calculated using the indices of Kyte and Doolittle (1982) for an 11 amino acid window. The positions of positively and negatively charged residues are indicated, as are the six potential membrane spanning domains.

Fig 4.9

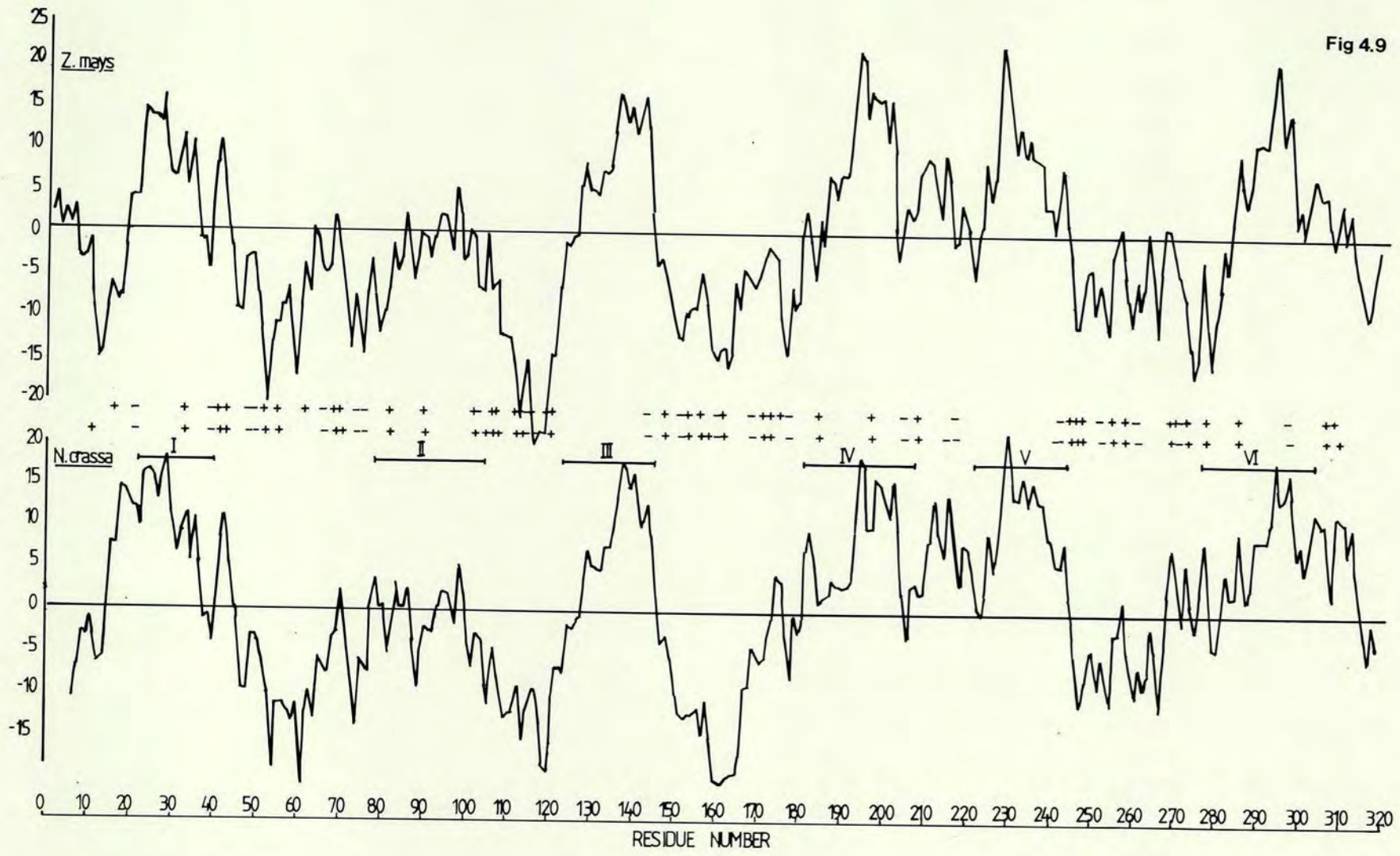


Fig 4.10

AMINO ACID SEQUENCES OF THE ADENINE NUCLEOTIDE TRANSLOCATOR  
POLYPEPTIDES FROM BEEF, MAIZE, NEUROSPORA AND YEAST, ALIGNED TO  
MAXIMISE THE HOMOLOGY BETWEEN THEM.

```

BEEF          S D Q A L S F L K D F L A G G V A A A I S K T A
MAIZE  M Q T P L C A N A P A E K G G K N F M I D F M M G G V S A A V S K T A
NEUROSPORA  M A E Q Q K V L G M P P F V A D F L M G G V S A A V S K T A
YEAST      M S H T E T Q T Q Q S H F G V D F L M G G V S A A I A K T G

V A P I E R V K L L L Q V Q H A S - K Q I S A E K Q Y K G I I D C V V
A A P I E R V K L L I Q N Q D E M I K S G R L S E P Y K G I V D C F K
A A P I E R I K L L V Q N Q D E M I R A G R L D R R Y N G I I D C F K
A A P I E R V K L L M Q N Q E E M L K Q G S L D T R Y K G I L D C F K

R I P K E Q G F L S F W R G N L A N V I R Y F P T Q A L N F A F K D K
R T I K D E G F S S L W R G Y T A N V I R Y F P T Q A L N F A F K D Y
R T T A D E G V M A L W R G N T A N V I R Y F P T Q A L N F A F R D K
R T A T H E G I V S F W R G N T A N V L R Y F P T Q A L N F A F K D K

Y K Q I F L G G V D R H K Q F W R Y F A G N L A S G G A A G A T S L C
F K R L F N F K K D R D G - Y W K W F A G N L A S G G A A G A S S L F
F K K M F G Y K K D V D G - Y W K W M A G N L A S G G A A G A T S L L
I K S L L S Y D R E R D G - Y A K W F A G N L F S G G A A G G L S L L

F V Y P L D F A R T R L A A D V - - G K G A A Q R E F T G L G N C I T
F V Y S L D Y A R T R L A N D A K A A K G G G E R Q F N G L V D V Y R
F V Y S L D Y A R T R L A N D A K S A K R G G E R Q F N G L V D V Y R
F V Y S L D Y A R T R L A A D A R G S K S T S Q R Q F N G L L D V Y K

K I F K S D G L R G L Y Q G F N V S V Q G I I I Y R A A Y F G V Y D T
K T L K S D G I A G L Y R G F N I S C V G I I V Y R G L Y F G L Y D S
K T I A S D G I A G L Y R G F G P S V A G I V V Y R G L Y F G L Y D S
K T L K T D G L L G L Y R G F V P S V L G I I V Y R G L Y F G L Y D S

A K G M L P D P K K V H I - I V S W M I A Q T V T A V A G L V S Y P F
I K P V V L T G N L Q D N F F A S F A L G W L I T N G A G L A S Y P I
I K P V L L V G D L K N N F L A S F A L G W C V T T A A G I A S Y P E
F K P V L L T G A L E G S F V A S F L L G W V I T M G A S T A S Y P L

D T V R R R R M M M Q S G R K G A D I M Y T G T V D C W R K I A K D E G
D T V R R R R M M M T S G - - - E A V K Y K S S L D A F Q Q I L K K E G
D T I R R R R M M M T S G - - - E A V K Y K S S F D A A S Q I V A K E G
D T V R R R R M M M T S G - - - Q T I K Y D G A L D C L R K I V Q K E G

P K A F F K G A W S N V L R G M G G A F V L V L Y D E I - - - - - K
P K S L F K G A G A N I L R A I A G A G V L S G Y D Q L Q I L F F G K
V K S L F K G A G A N I L R G V A G A G V L S I Y D Q L Q V L L F G K
A Y S L F K G C G A N I F R G V A A A G V I S L Y D Q L Q L I M F G K

K F V *
K Y G S G G A *
A F K G G S G *
K F K *

```

one of the flanking PstI sites. Restriction mapping data suggest that this is due to a deletion resulting in loss of some vector sequences adjacent to the inserted cDNA. The region of this clone sequenced corresponds to amino acids 200 to 310 of pANT-1 but does not include the deletion break point which occurs at the 5' end of the sequence, or the 3' untranslated region. The nucleotide sequence of pANT-2 thus far determined reveals a number of nucleotide substitutions relative to pANT-1, most of which occur in the 3rd position of the codon and do not change the amino acid specified by that codon. Some of these changes do however create restriction site polymorphisms, such as the loss of the ClaI site in pANT-2. This result indicates that there are at least two different genes encoding ATP/ADP translocator polypeptides expressed in young dark grown coleoptile tissue.

#### 4.4.4 RNA Transcripts of the ANT genes.

Samples of total cellular RNA and the poly A<sup>+</sup> RNA preparation used to construct the library were fractionated by electrophoresis in formaldehyde-agarose gels and transferred to nitrocellulose. When hybridised with either nick-translated pANT-1 plasmid DNA or with strand specific M13 subclones of the plasmid, a single transcript of ca. 1600 nucleotides was detected, Fig.4.11. The cloned cDNA is 1196 nucleotides long including the cloned portion of the poly A tail. Assuming the transcript is a mature message, this leaves an additional 400 nucleotides to be accounted for by additional lengths of polyA and the 5' untranslated sequence, which implies that the latter may be unusually long. The probes and hybridisation conditions used in this experiment would not discriminate between transcripts of the two different genes. The detection of a single band, even on over exposure of the autoradiograph



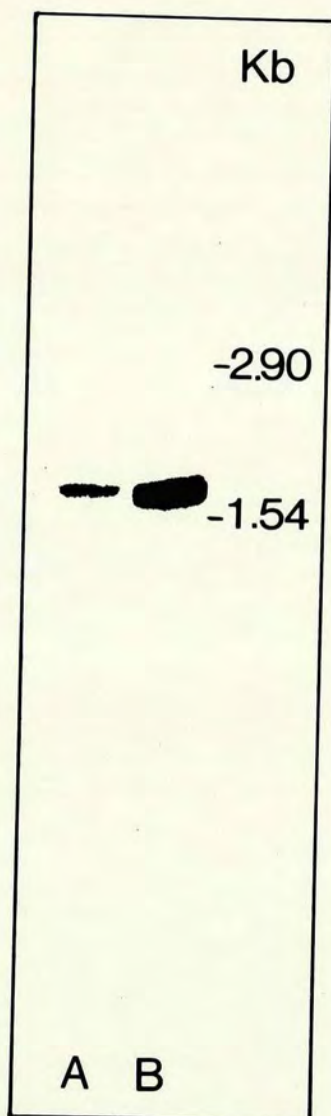


Fig.4.11 RNA transcripts of the maize adenine nucleotide translocator.

Panel A represents 15ug total RNA, Panel B, 2ug poly A<sup>+</sup> RNA which were fractionated by electrophoresis in the same 1%(w/v) formaldehyde agarose gel. RNA was transferred to nitrocellulose, probed with nick translated plasmid pANT-1 and autoradiographed. The size markers are E. coli rRNA.

(not shown) suggests that the transcripts of the two genes are sufficiently similar in size not to be resolved.

#### 4.4.5 Codon usage in pANT-1.

Despite the high degree of nucleotide sequence homology between the maize, yeast and Neurospora crassa adenine nucleotide translocator genes, the codon usage is quite different. Codon usage for the three genes is summarised in Table 4a. The maize gene uses 51 out of 61 possible amino acid specifying codons while the yeast gene uses 57 and the Neurospora gene uses only 36. In both maize and Neurospora TAA is used to specify stop while yeast uses TGA. Codons with A in the 3rd position are under represented in all three genes. In Neurospora, these occur <1% of the time, (one CAA codon specifying glutamine in a protein of 313 amino acids). In maize the figure is 8% or 25 out of 318 codons and in yeast 16% or 50 out of 309 codons. This is in contrast to the situation in several other maize nuclear genes, although maize and soybean actin show a similar bias (Shah et al., 1983) However, the number of sequences available is distorted by the plethora of storage protein gene sequences, so it remains to be seen whether the ATP/ADP translocator is a special case or whether this bias against the use of A in the wobble position is

- (1) representative of maize nuclear genes in general, or
- (2) representative of genes which are constitutively expressed, or
- (3) a feature of genes whose products are destined for the mitochondrion.

The only other higher plant nuclear gene encoding a mitochondrial protein to be sequenced to date, the beta subunit of F<sub>1</sub>ATPase from Nicotiana plumbaginifolia, (Boutry and Chua 1985) does not show this bias against A in the 3rd position.

codon	m	n	y	codon	m	n	y	codon	m	n	y	codon	m	n	y
GGG	3	0	5	GAG	5	7	4	GTG	3	1	6	GCG	3	2	7
GGA	8	0	5	GAA	3	0	5	GTA	0	0	2	GCA	2	0	3
GGT	19	30	15	GAT	9	6	7	GTT	7	1	5	GCT	25	12	10
GGC	8	7	8	GAC	7	11	8	GTC	5	21	4	GCC	5	21	4
AGG	7	0	5	AAG	24	23	13	ATG	8	10	9	ACG	0	0	6
AGA	2	0	7	AAA	2	0	8	ATA	0	0	2	ACA	3	0	4
AGT	1	0	1	AAT	6	0	5	ATT	11	3	8	ACT	5	2	5
AGC	1	0	0	AAC	8	11	2	ATC	7	12	3	ACC	2	10	4
TGG	4	4	3	TAG	0	0	0	TTG	6	3	16	TCG	0	1	7
TGA	0	0	1	TAA	1	1	0	TTA	0	0	2	TCA	3	0	4
TGT	2	0	2	TAT	4	0	5	TTT	8	0	9	TCT	8	5	6
TGC	1	2	1	TAC	12	14	9	TTC	15	18	11	TCC	8	12	4
CGG	0	0	4	CAG	9	8	10	CTG	7	1	8	CCG	2	1	0
CGA	0	0	0	CAA	1	1	3	CTA	1	0	4	CCA	1	0	0
CGT	6	14	1	CAT	0	0	1	CTT	8	3	5	CCT	2	2	1
CGC	2	5	1	CAC	0	0	2	CTC	6	21	2	CCC	3	4	4

Table 4a Codon usage for the translocator genes of maize (pANT-1) (m) Neurospora (n) and yeast (y)

In yeast and E.coli where numerous gene sequences are available and the relative amounts of the different tRNA species are known, it is clear that the codon usage relates to the level of expression of many genes (Bennetzen and Hall 1982). In general genes which are expressed at a high level contain many codons for the most abundant tRNA species. This avoids depleting the pool of charged tRNA's of the rarer species which might lead to an inhibition of protein synthesis or an increase in missense reading. Furthermore the use of 'rare' codons may represent a means of 'fine-tuning' expression of a gene at the level of translation. Codon usage will also reflect the overall base composition of the sequence. The high G+C content of the Neurospora adenine nucleotide translocator coding region (58.5%, as opposed to 48.9% for the maize translocator coding region), may reflect or be a consequence of, the overwhelming preference for C in the 3rd position (55% of codons). In contrast the most abundant codons used in the maize gene (38% of the total) are those ending in T.

#### 4.4.6 Predicted secondary and higher order structure of the ATP/ADP translocator protein.

The hydropathy profiles of the adenine nucleotide translocator polypeptides of Neurospora and maize are shown in Fig.4.9 The striking similarities between them, both with respect to the conservation of hydrophobic and hydrophilic domains and the distribution of charged residues, serves to further confirm the high degree of homology between them. There are six regions of sufficient length and overall hydrophobicity to traverse the membrane as alpha helices (Saraste and Walker 1982) and these are indicated on Fig.4.9. Some of the segments contain charged residues which would need to be neutralised to permit being buried in the membrane.

Another striking feature of the hydropathy profiles is the presence of a three fold repeat (Saraste and Walker 1982). This feature is apparent in all the translocator polypeptides at both the nucleotide and amino acid level, and is illustrated by the maize translocator sequence in Fig.4.12. As a result of analysis of the beef polypeptide sequence, Saraste and Walker (1982), proposed that the gene evolved as a result of two gene duplications. The occurrence of the repeated feature in all translocators sequenced to date would appear to bear this out, particularly as the repeat is also obvious at the nucleotide sequence level.

The topology of the translocator in the membrane has been probed using pyridoxal phosphate as an impermeant agent to modify lysine residues accessible in isolated mitochondria and sub mitochondrial particles from beef heart (Bogner *et al.*, 1983). These studies support the suggestion that segments I-VI span the membrane, and put the hydrophilic regions linking II and III, and IV and V on the cytoplasmic side of the bilayer, while the segments linking I and II, and II and IV are accesible from the matrix. The effects of binding CAT and BKA on the labelling pattern of the polypeptide were also studied. Different sets of lysine residues were labelled in the c and m states, giving further support to the idea that a conformational change takes place upon transition between the cytosol facing and matrix facing binding sites.

Binding of both ATR and CAT convert the translocator to the c state, but whereas CAT is irreversibly bound, ATR can be displaced. In ATR 'loaded' mitochondria, lys 22 and 146 were labelled to a greater extent than in CAT loaded mitochondria, suggesting that these residues are unmasked by ATR removal and hence may be at or close to the ligand binding site. Of the 14 lysines which could be labelled in these studies, only six are conserved between all four species, and these include lys 22

Fig. 4.12 The amino acid sequence of the maize Adenine Nucleotide Translocator derived from clone PANT-1, aligned after Saraste and Walker (1982), to display the internal repeat.

```

1  M 0 T P L C A N A P A E K G K N F M I D F M M G G V S A A V S K T A A A P
    123
    W F A G N L A S G G A A G A S S L F F V Y S
    224
    F F A S F A L G W L I T N G A G L A S Y P
39  I E - - R V K L L I O N O D E M I - K S G R L S E P Y K G I V D C F K R T I
    145
    L D Y A R T R L A N D A K - - - A A K G G E R Q - F N G L V D V Y R K T L
    246
    I D T V R R R M M M T S - - - - - G E A V K Y K S S L D A F Q Q I I
75  K D E G F S S L W R G Y T A N V I R Y F P T Q A L N F A F K D Y F K R L F N
    179
    K S D G I A G L Y R G F N I S C V G I I V Y R G L Y F G L Y D S I K F V V L
    274
    K K E G P K S L F K G A G A N I L R A I A G A G V L S G - Y D Q L Q I L F F
113  F K K D R D G Y W K
    217
    T G N L Q D N
    310
    G K K Y G S G G A

```

and 146.

It has been known for several years that N-ethyl maleimide inhibits the translocation of adenine nucleotides (Vignais and Vignais 1972). This reagent alkylates sulphhydryl groups and thus indicating the involvement of a cysteine in the translocation process. A recent publication (Boulay and Vignais 1984) demonstrated directly that the cysteine alkylated in the beef polypeptide is cys 56. This is supported by the sequence comparisons shown in Fig.4.10 as this is the only cysteine which is conserved in all the sequences.

#### 4.5 CONCLUSION.

The clone pANT-1 represents a cloned cDNA for the ATP/ADP translocator of maize. This assertion is supported by the high degree of both nucleotide and amino acid sequence homology between this clone and the adenine nucleotide translocators from beef heart, yeast and Neurospora crassa mitochondria. This homology implies a high degree of structural and functional similarity.

The nucleotide translocator of maize lacks a 5'-AATAAA-3' sequence, which supports the view that this sequence may not be essential for polyadenylation of higher plant mRNA's. The codon usage of the translocator gene differs from that of many other maize genes, and from the Neurospora translocator. The significance of this is uncertain.

The cDNA clone pANT-1 encodes 5 more amino acids than the Neurospora and 11 amino acids than the beef polypeptide. The size of the transcript, ca. 1600 nucleotides, suggests that either the coding sequence is much longer, or that the transcript has an abnormally long 5' end. In addition the recovery of 2 distinct cDNA's, pANT-1 and pANT-2 raises the possibility that the translocator is encoded by a small multigene family.

In order to address these questions, the construction of a maize nuclear DNA library in a bacteriophage lambda vector, and the isolation and characterisation of genomic clones for the adenine nucleotide translocator was undertaken.



## CHAPTER 5. ORGANISATION, STRUCTURE AND EXPRESSION OF THE NUCLEAR GENES ENCODING ADENINE TRANSLOCATOR POLYPEPTIDES IN ZEA MAYS L.

### 5.1 RATIONALE

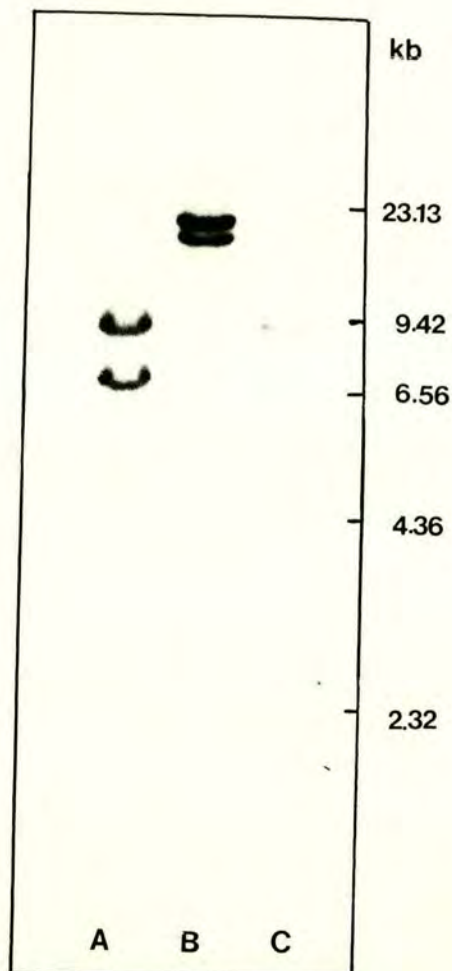
Characterisation of the cDNA clones for the ATP/ADP translocator (described in section 4.3), suggests that, in contrast to the situation in *Neurospora crassa* (Arrends and Sebald, 1984), there may be multiple genes encoding the polypeptide in maize. In order to confirm and extend these observations, the isolation and characterisation of the structural genes corresponding to the cDNA clones pANT-1' and pANT-2 was undertaken. The cDNA clones were used as hybridisation probes in Southern blots and to identify genomic clones from a maize nuclear genomic library.

### 5.2 GENOMIC ORGANISATION.

#### 5.2.1 Analysis of total and nuclear DNA by Southern blotting.

Ten microgramme aliquots of total DNA were digested separately with each of the restriction endonucleases EcoR1 and HindIII. Ten microgrammes of DNA purified from isolated nuclei ('nuclear DNA' expected to be largely free of contaminating chloroplast and mitochondrial DNA) was digested with EcoR1. The restriction fragments were separated by electrophoresis through a 0.8% (w/v) agarose gel. The DNA was transferred to nitrocellulose by capillary blotting and hybridised with the <sup>32</sup>P labelled 1.2kb cDNA insert from pANT-1, Fig 5.1.

The total and nuclear DNA digested with Eco R1 gave identical patterns of hybridisation. Three discrete DNA fragments with estimated sizes of 26.2, 8.2 and 6.4kb hybridised with the probe. Three HindIII generated fragments with estimated sizes of 18.5kb, 14.9kb and 8.4kb also



**Fig.5.1** Nuclear genomic Southern blot of total DNA and DNA from purified nuclei.

The 1.2kb cDNA insert from pANT-1 was labelled with  $^{32}\text{P}$  by the random primer method of Feinberg and Vogelstein (1984) to a specific activity of  $8.5 \times 10^8 \text{cpm/ug}$ . Hybridisation and washes were as described in the legend to Fig.3.6.

Lane A: 10ug total DNA digested with EcoR1.

Lane B: 10ug total DNA digested with HindIII.

Lane C: 10ug DNA from purified nuclei digested with EcoR1.

Markers are from a HindIII digest of  $\lambda\text{cl857}$ .

hybridised to the probe. The 8.4kb HindIII fragment reproducibly hybridises less intensely than the other two HindIII fragments. The fact that it is the smallest in size makes it unlikely that it is a partial digestion product. It may represent a sequence which shares some homology to the probe but less than the homology between the probe and the other two HindIII fragments, or it may represent a restriction fragment which has only a small overlap with the probe sequence. The ca. 26.2kb Eco R1 fragment appears in variable stoichiometry in different blots (compare Figs.5.1 and 5.3 ). This may be a partial digestion product, or it may reflect the relatively inefficient transfer of large DNA fragments to nitrocellulose. The latter argument is supported by a recent observation that a longer de-purination treatment of the gel, (15min. in 0.25N HCl as opposed to 5min.) and overnight transfer as opposed to 3-4h., increases the relative proportion of the 26.2kb fragment detected by hybridisation.

None of the cDNA clones isolated contain restriction sites for Eco R1 or HindIII. As the entire 1.2kb cDNA insert from pANT-1 was used as a probe in these experiments, the results presented here have two alternative explanations:

- (1) Each DNA fragment contains a complete copy of a gene encoding the translocator, with the restriction sites lying outside the coding sequence;
- (2) The coding region of one of the genes is split by an intron or introns containing both a HindIII and Eco R1 site, generating two hybridising fragments, each representing part of the gene.

Thus the minimum estimate for the number of translocator genes, based on Southern hybridisation and supported by recovery of cDNA clones, is two. It is not possible to distinguish on the basis of hybridisation to nuclear genomic DNA whether any of the fragments which hybridise to the

probe contain more than one copy of the gene.

### 5.2.2 Do chloroplast and mitochondrial DNA contain sequences homologous to the translocator?

In the light of recent reports concerning DNA transfer between organelles (van den Boogart *et al.*, 1982, Stern and Lonsdale 1982, Kemble *et al.*, 1983, Timmis and Scott, 1983, Wright and Cummings 1983), DNA from maize chloroplasts and mitochondria was examined for sequences homologous to the translocator. Ten microgrammes of nuclear DNA (extracted from purified nuclei) and 1 $\mu$ g each of maize chloroplast and mitochondrial DNA were digested with Eco R1 and separated by electrophoresis through a 0.8% (w/v) agarose gel. The DNA was transferred and hybridised with pANT-1 insert as before. The autoradiograph is shown in Fig.5.2. The 8.2 and 6.4kb Eco R1 fragments are clearly present in the nuclear DNA lane, although the 26.2kb fragment is only faintly visible. No hybridisation is observed in the mtDNA lane, while in the chloroplast DNA lane faintly hybridising fragments are visible co-migrating with the nuclear fragments. These are almost certainly attributable to nuclear contamination of the chloroplast DNA preparation. No unique fragments are observed. This result, and the observation that total and nuclear DNA give the same pattern of hybridisation when digested with Eco R1 shows that there are no sequences homologous to the translocator gene sequence in either the chloroplast or mitochondrial genomes of maize.

## 5.3 ISOLATION OF ANT GENOMIC CLONES.

### 5.3.1 Construction of a maize nuclear genomic library in $\lambda$ gtWES.

The bacteriophage lambda vector  $\lambda$ gtWES.  $\lambda$ B is a DNA replacement vector which can accommodate Eco R1 fragments of 1-14kb in length (Leder

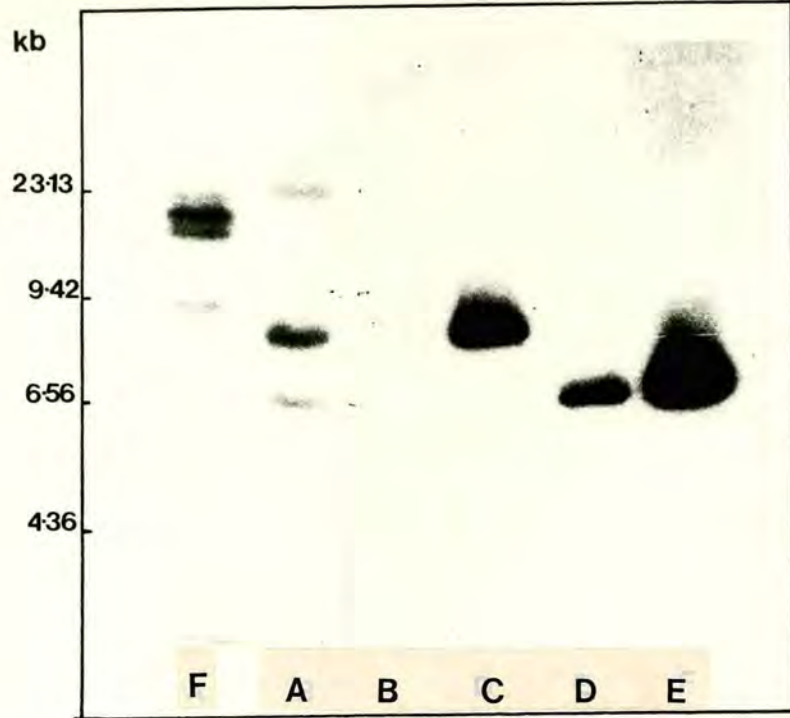


Fig.5.3

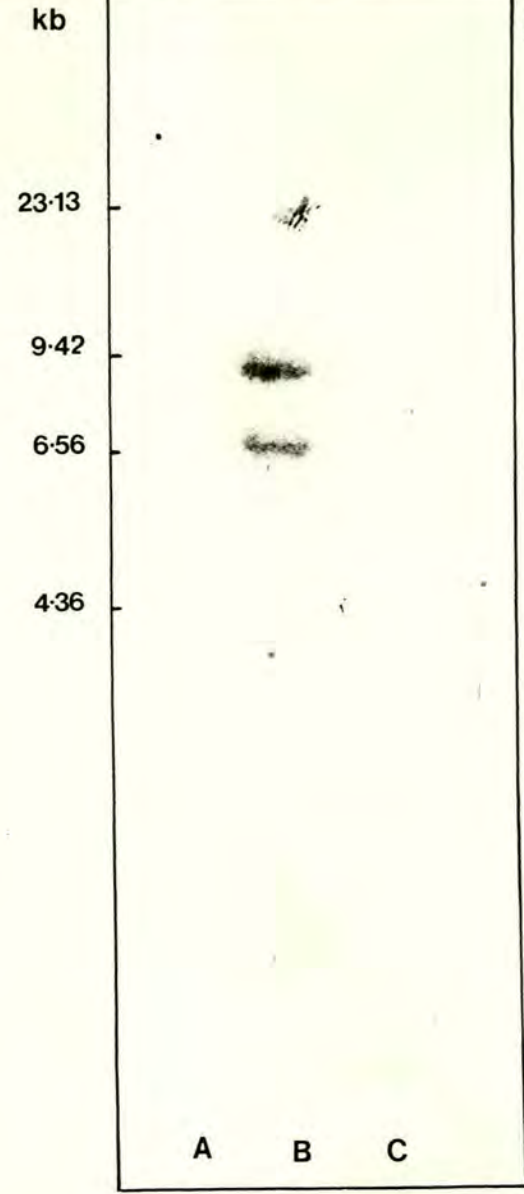


Fig.5.2

Fig.5.2 Maize chloroplast and mitochondrial genomes do not contain sequences homologous to the translocator.

Autoradiograph of a filter probed as in Fig.5.1. Lane A 1ug mtDNA, Lane B 10ug nuclear DNA, Lane C 1ug CtDNA.

Fig. 5.3 The clones pλ8 and pλ32 correspond to EcoRI fragments identified in Southern blots.

Autoradiograph of filter hybridised as described in Fig.5.1 legend.  
 Lane A total DNA digested with EcoRI  
 Lanes B and C, 1 and 10 gene copy equivalents of pλ8 digested with EcoRI  
 Lanes D and E, 1 and 10 gene copy equivalents of pλ32 digested with EcoRI

**F** total DNA digested with HindIII

et al., 1977). When the purified Eco R1 generated right and left arms (13.84 and 21.70kb respectively) are ligated together they are too small to be packaged efficiently, thus providing an effective selection against non-recombinants.

Total maize DNA was digested to completion with Eco R1, and ligated to lambda arms prepared by sucrose gradient centrifugation of Eco R1 digested  $\lambda$ gtWES DNA, (section 2.5.4) The optimum ratio of maize DNA to lambda arms was determined in a trial ligation and packaging experiment (Table 5a).

TABLE 5a

<u>Maize DNA (ug)</u>	<u>Arms (ug)</u>	<u>No. of plaques</u>	<u>Plaques per ug Maize DNA</u>
0	1.0	320	0
0.05	1.0	1400	$2.8 \times 10^4$
0.1	1.0	2200	$2.2 \times 10^4$
0.5	1.0	12000	$2.4 \times 10^4$
1.0	1.0	18000	$1.8 \times 10^4$

The bacterial host strain used was BHB2600, (Hohn 1979) an E. coli K803 derivative which is  $recA^+$ . K803 strains have been reported to give higher plating efficiencies with phage carrying maize DNA inserts (Fedoroff 1983).

To construct the library, 20ug Eco R1 digested DNA was ligated to 20ug of lambda arms, packaged and plated at a density of ca.  $10^4$  plaques per 90mm diameter plate. Approximately  $10^6$  phage plaques were obtained. Eco R1 recognises a hexanucleotide sequence, so cuts on average once every  $4^6$ bp or 4kb. In a library of  $10^6$  clones the probability of cloning any

given single copy sequence is 0.55. This is calculated from the formula of Clarke and Carbon (1976):

$$N = \frac{\ln(1-P)}{\ln(1-f)}$$

Where N is the number of clones

f the fraction of the genome present in a single recombinant

P the probability of the sequence being present in the library

However the probability of cloning the 6.4 and 8.2kb fragments will be much higher than this as a significant proportion of the genome will consist of Eco R1 fragments outside the cloning size range of 1-14kb for  $\lambda$ gtWES.

### 5.3.2 Screening the library for sequences homologous to pANT-1.

The library was screened initially with pANT-1 plasmid DNA labelled with  $^{32}\text{P}$  by nick-translation. Nitrocellulose filters were made according to the procedure of Benton and Davis (1977), and hybridised with the probe for 16h at 65 $^{\circ}\text{C}$  in 3xSSPE, 10x Denhardt's solution, 0.1%(w/v)SDS, 100ug/ml herring sperm DNA, and washed at 65 $^{\circ}\text{C}$  for 2x15min in 2xSSPE, 0.1%(w/v)SDS. On the first round of screening 35 hybridising plaques were obtained. These were picked and replated and screened as before. Nine clones were strongly positive on the second screen.

Attempts to plaque purify some of these phage were initially unsuccessful. Every time a single hybridising plaque was picked and replated, a high proportion of the plaques did not hybridise when rescreened with the probe. When DNA was made from single hybridising plaques, all the clones tested had identical restriction patterns and did not hybridise to the probe. These results would suggest that the clones are unstable and a high proportion tend to rearrange and delete part of their

inserts. Apparently this is not an uncommon problem with lambda clones of higher plant DNAs (D. Baulcombe pers. comm.). In this instance, the problem was overcome by changing host strains. Using E.coli strain ED8654 as a host it proved possible to plaque purify and make DNA from two clones  $\lambda 8$  and  $\lambda 32$ , which contained Eco R1 inserts of 8.0 and 6.4kb respectively, and hybridised intensely to the  $^{32}\text{P}$  labelled cDNA insert from pANT-1. The reason why ED8654 proved a more suitable host for recombinant phage than BHB2600 is unknown as both strains are recA+.

#### 5.4 CHARACTERISATION OF THE GENOMIC CLONES.

##### 5.4.1 Restriction endonuclease mapping.

The Eco R1 fragments of maize DNA cloned in  $\lambda 8$  and  $\lambda 32$  were re-cloned into the Eco R1 site of pUC9 to facilitate restriction mapping. The resulting plasmids were designated p $\lambda 8$  and p $\lambda 32$ . Because of the problems of instability associated with some of the lambda clones, it was important to demonstrate that these Eco R1 fragments did indeed correspond to those detected by Southern hybridisation and had not undergone rearrangement during the cloning procedure. Aliquots of Eco R1 digested p $\lambda 8$  and p $\lambda 32$ , calculated to be equivalent to one and ten copies of the gene per haploid genome, were fractionated by electrophoresis through a 0.8% (w/v) agarose gel alongside an Eco R1 digest of total maize DNA. The gel was transferred to nitrocellulose and probed with the cDNA insert of pANT-1 as before. The resulting autoradiograph is shown in Fig.5.3. Both cloned fragments co-migrated with the hybridising fragments in maize DNA, making it unlikely that any rearrangement had taken place during cloning.

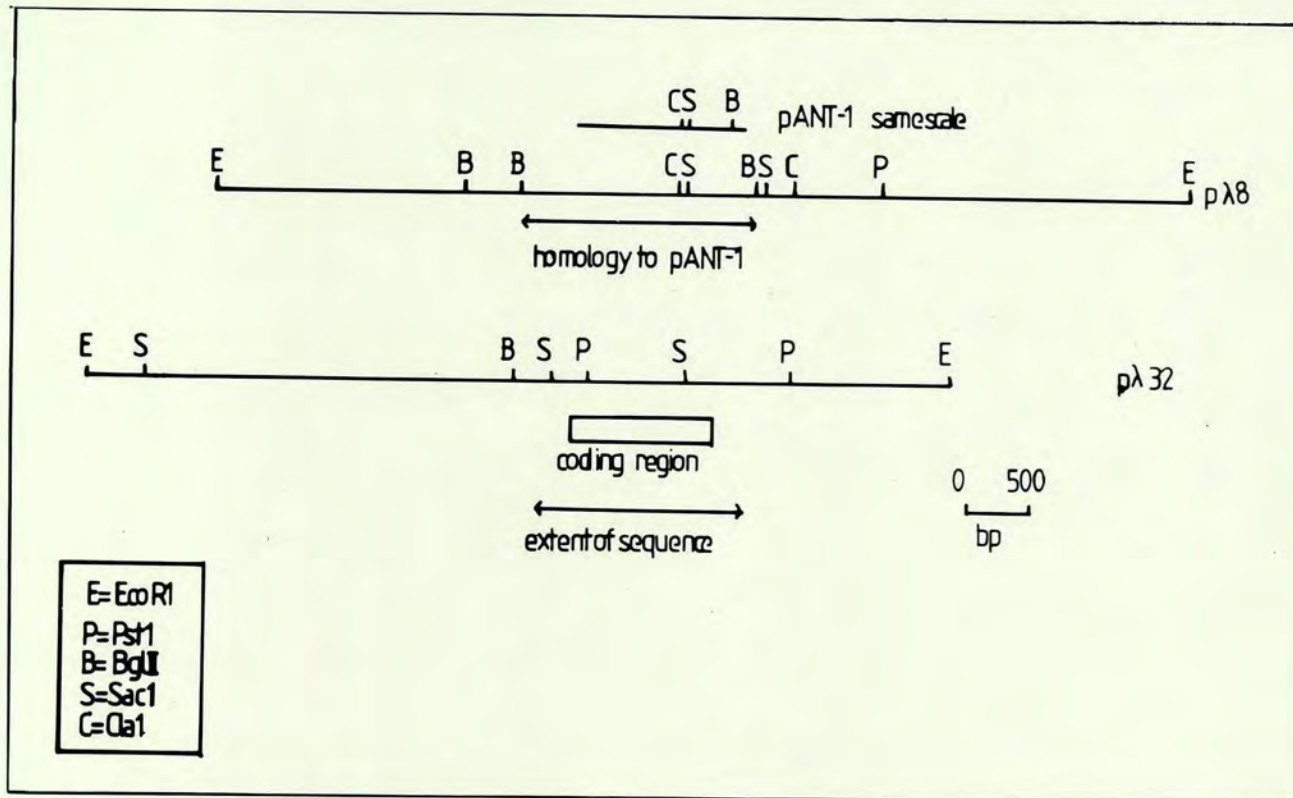
Restriction maps of the two clones were constructed by means of a series of single and double digests with enzymes recognising



hexanucleotide cutting sites in the DNA. The gene was located on the restriction map in each case by transferring the digested DNA to nitrocellulose and probing with the  $^{32}\text{P}$  labelled cDNA insert from pANT-1. The orientation of the coding sequence was determined by probing the mapping gels with a probe derived from the N-terminal region of the cDNA clone. The restriction maps of the two clones are shown in Fig.5.4.

From the restriction maps it appears that the insert of p $\lambda$ 8 carries the gene from which the cDNA clone pANT-1 was derived. The diagnostic restriction sites are the adjacent ClaI and SacI sites, which occur in the coding region towards the C terminus, and the BglIII site which, in the cDNA clone, occurs 26 nucleotides 5' to the point of polyA addition and so acts as a convenient marker for the 3' end of the mRNA. The insert in plasmid p $\lambda$ 32 lacks both the ClaI and BglIII sites, as does pANT-2.(section 4.3.3). The hypothesis that the ClaI site is absent from p $\lambda$ 32 as a result of an intervening sequence present in the genomic clone 5' to the SacI site can be discounted, as an N-terminal specific pANT-1 probe hybridises to the 1.1kb SacI fragment suggesting that the 5' end of the coding region lies within this restriction fragment(not shown). Furthermore sequence analysis of the cDNA clone pANT-2 indicated that the loss of the ClaI site was a consequence of a single base change from C to T changing the sequence 5'-ATCGAT-3' in pANT-1 to 5'-ATTGAT-3' in pANT-2. Thus the available evidence suggested that the gene from which the pANT-1 cDNA was derived is located on the 8.0kb Eco R1 fragment cloned in p $\lambda$ 8 and designated G1. The gene located on the 6.4kb Eco R1 fragment, designated G2, is consistent in its restriction map with the clone pANT-2.

Restriction mapping and hybridisation data also indicate that both fragments probably contain complete copies of the gene, although the presence of an intervening sequence in the extreme 5' or 3' ends of the



**Fig.5.4** Restriction maps of the nuclear DNA inserts in pλ8 and pλ32.

pANT-1 is shown to the same scale. The region sequenced from pλ32 is indicated as is the region of homology between pλ8 and pANT-1 as determined by Southern hybridisation.

transcription unit cannot be ruled out in the absence of sequence data. If this is the case and a portion of one of the genes lies on another restriction fragment e.g. the 26.2kb Eco R1 fragment, the intron must be several kilobases long. This is unlikely as the majority of plant introns are of the order of a few hundred base pairs.

#### 5.4.2 Nucleotide sequence analysis.

In order to confirm that the translocator gene G2 corresponded to the cDNA clone pANT-2, and to investigate the extent of sequence divergence between the two copies of the gene, the nucleotide sequence of G2 was determined.

Southern hybridisation had located the gene on the restriction map as indicated in Fig.5.4. Pst1, Sac1 and Pst1/Sac1 clones were constructed in the M13 vectors mp11 and mp19. Pst1 and Sac1 single digests of p  $\lambda$ 32 were cloned in mp11. Restriction fragments from double digests of Pst1 and Sac1 were cloned in mp11 and mp19 in order to obtain clones in both orientations relative to the M13 primer hybridisation site. White plaques from the resulting transformations were picked onto grids and nitrocellulose filters of the phage containing plates made by taking plaque 'lifts'. The filters were probed with the 1.2kb cDNA insert from pANT-1, labelled with  $^{32}$ P by nick-translation. Twenty clones were plaque purified and single and double stranded DNA prepared. The double stranded DNA was digested with the cloning enzyme(s) and electrophoresed on a 1% agarose gel, alongside the appropriate digestions of p  $\lambda$ 32. In this way it was possible to identify clones containing specific inserts, Fig.5.5. For example 18/PS2 carries the 0.4kb Sac1/Pst1 fragment, 11/S7 the 1.1kb Sac1 fragment, 11/PA2 the 1.5kb Pst1 fragment, and 11/S1 the 4.7kb Sac1 fragment.

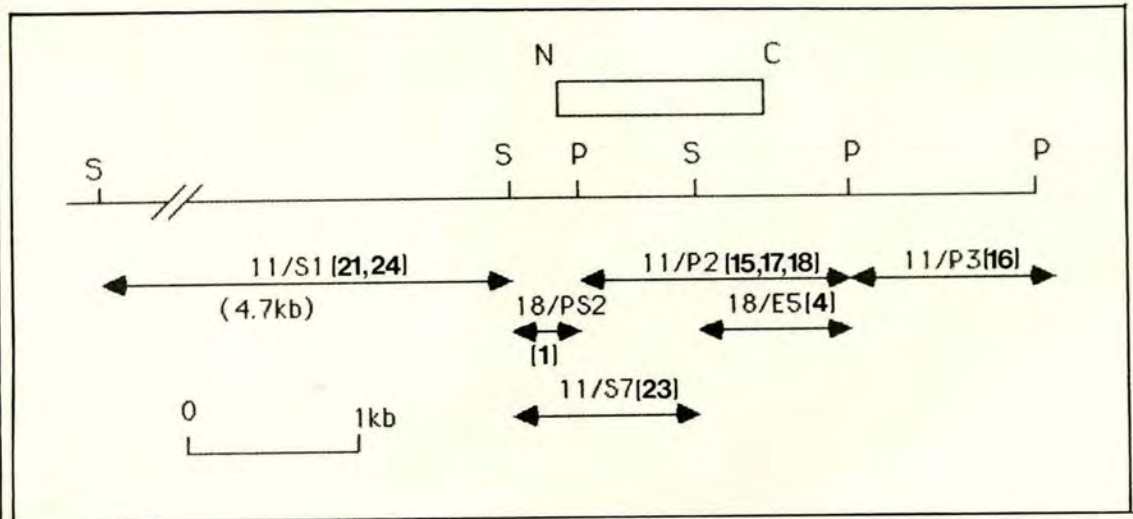
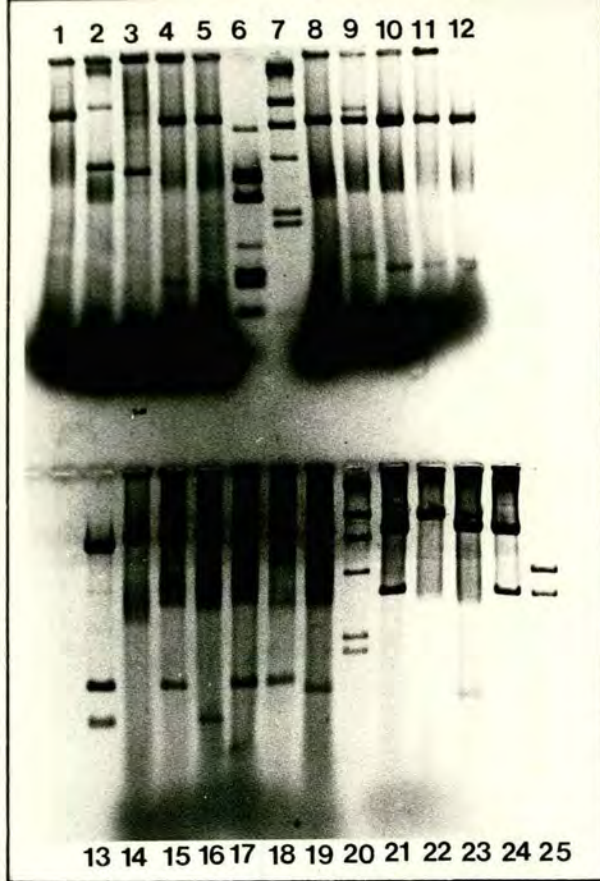


Fig. 5.5 Identification of PstI and SacI subclones of pλ32.

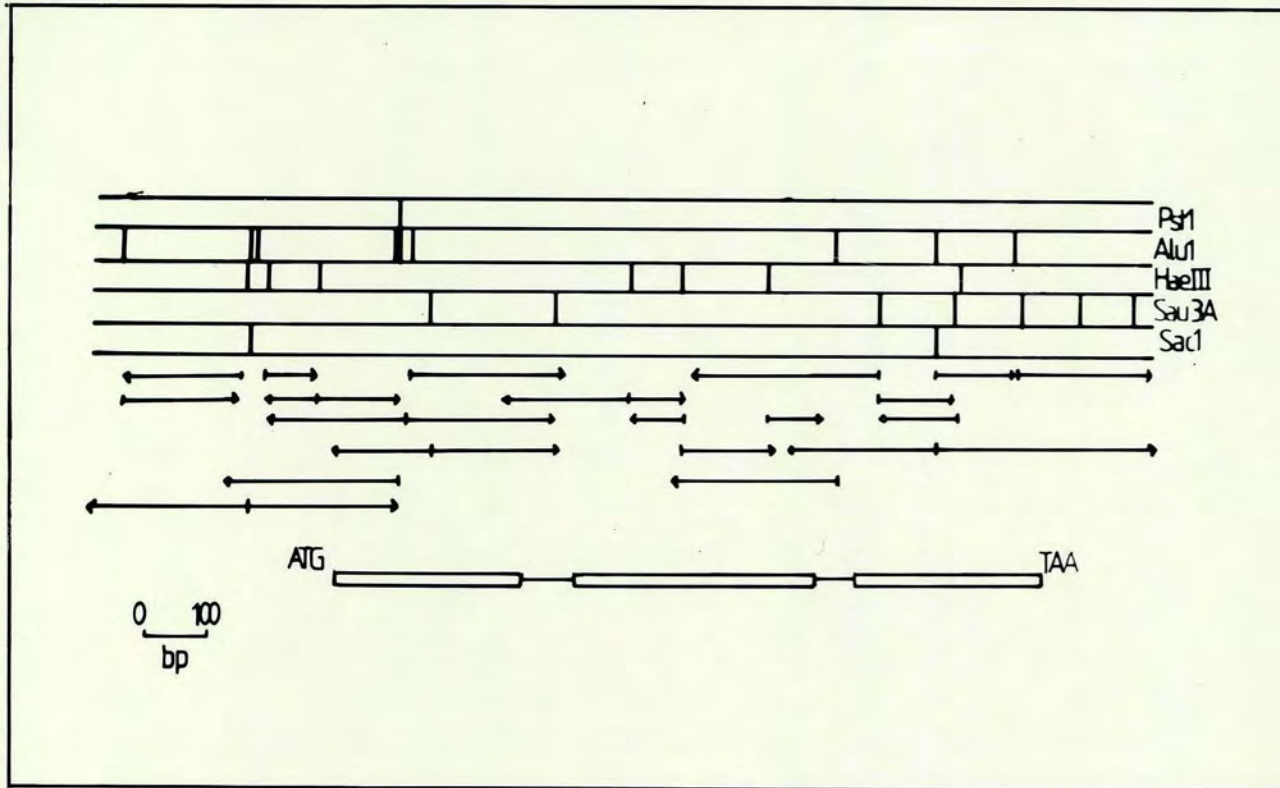
Left panel: Ethidium bromide stained gel showing PstI + SacI (lanes 1-5 and 8-12), PstI (lanes 14-19) and SacI (lanes 21-24) digests of recombinant M13 clones. PstI + SacI, PstI and SacI digests of pλ32 are included in lanes 6, 13 and 25. Lanes 7 and 20 are HindIII digests.

Right panel: Clones from above located on a restriction map of pλ32.

- Nucleotide sequence analysis of these clones confirmed that G2 is a translocator gene and differs from the pANT-1 sequence in a number of positions. The DNA sequence was completed by shotgun cloning AluI, HaeIII, and Sau3A fragments as before. Single stranded gaps were filled by cloning specific restriction fragments isolated from acrylamide gels, or by 'clone turnaround' (Messing 1983). The latter procedure can be adopted when a restriction fragment has been cloned in one orientation only, and sequence information is required from the distal cloning site. The double stranded RFDNA is digested with the cloning enzyme(s), and either religated or ligated with another appropriately digested M13 vector which will yield recombinants of the correct orientation. This approach was used to turn around clone 18/PS2, in order to obtain sequence from the SacI site. The RFDNA was digested with PstI and SacI to release the insert, and ligated with PstI/SacI digested mp19. As the SacI site in the mp19 polylinker is proximal to the primer hybridisation site, the clones obtained were in the desired orientation. Using a combination of these techniques a continuous sequence of 1.72kb was obtained, which included the entire gene sequence and a substantial amount of sequence 5' to the coding region. Of this 94% was confirmed from two or more independently isolated clones and 75% was sequenced from both strands of the DNA. All overlaps were verified from at least one other clone. The sequencing strategy is presented in Fig.5.6. Fine scale restriction maps of the G2 gene and the cDNA clones pANT-1 and pANT-2 are shown in Fig.5.7.

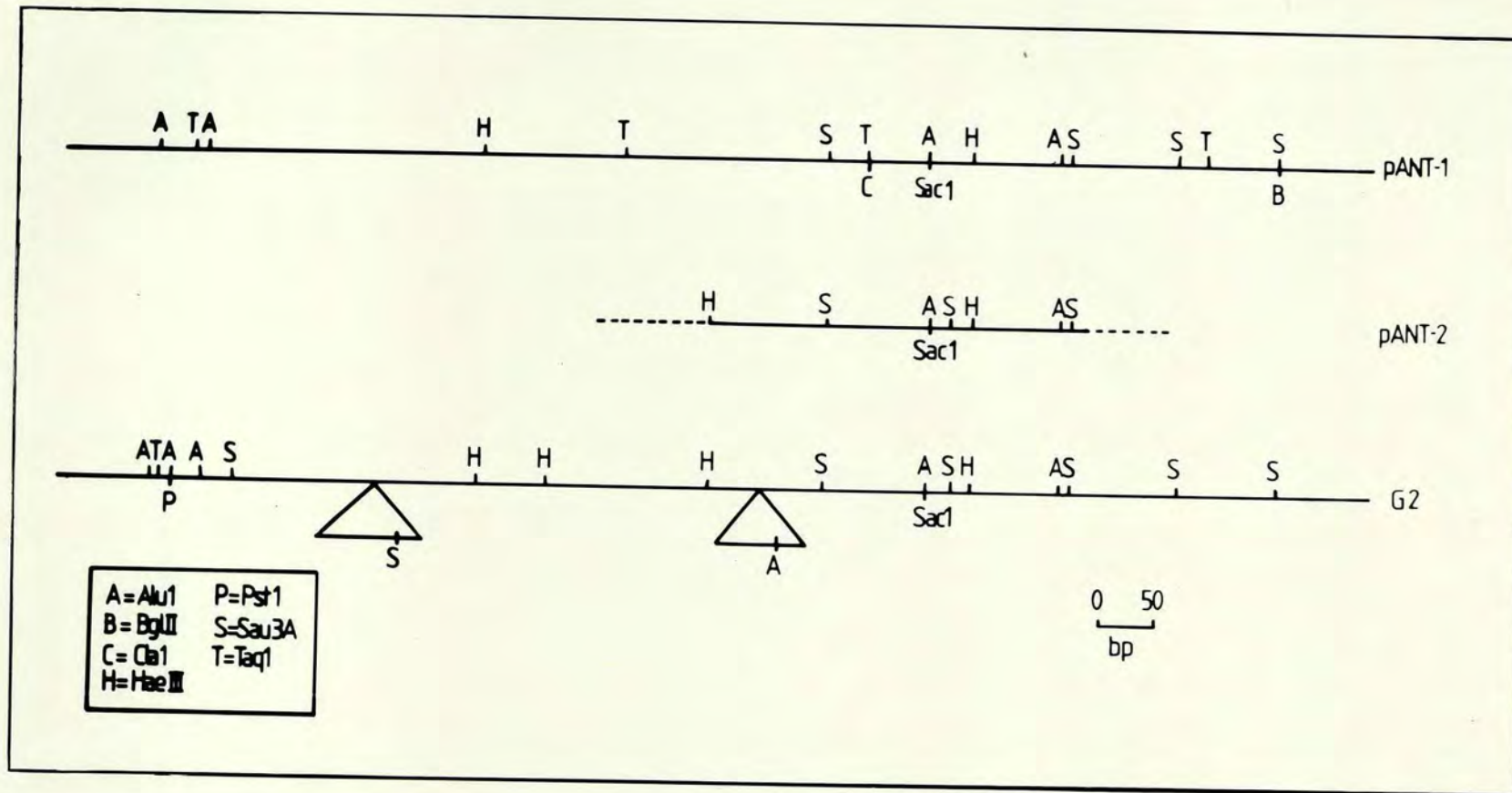
#### 5.4.3 Mapping the 5' point of divergence between the genomic clones.

In order to identify the region in which the sequences of the G1 and G2 genes diverge, M13 subclones from the 5' end of the coding sequence and 5' untranslated region of the G2 gene were used to probe filters to which DNA



**Fig.5.6** The sequencing strategy employed for the G2 gene.

The arrows indicate the extent and direction of the sequence determined. Exons are identified by boxes and the introns by thin lines.



**Fig. 5.7** Fine scale restriction map of the cDNA clones pANT-1 and pANT-2, and the genomic clone G2 aligned at the common SacI site.

The distances between the different restriction sites were determined by DNA sequence analysis. The extent of sequence determined is indicated by solid lines.

from p $\lambda$ 32 and p $\lambda$ 8 had been bound. The filters were hybridised in 3XSSPE, 10X Denhardt's solution, 0.1%(w/v)SDS at 65°C and washed in 2XSSPE, 0.1%(w/v)SDS at the same temperature. The results are presented in Table 5b

<u>PROBE</u>	<u>CO-ORDINATES</u> (Fig 5.8)	<u>HYBRID.</u> <u>TO G1</u>	<u>HYBRID.</u> <u>TO G2</u>
pANT-1	400-1700	+	+
18PS2	256-504	+	+
18A6	45-235	-	+
18H13	282-361	-	+

These results place the point of divergence between co-ordinates 361 and 504. From the nucleotide sequence of pANT-1 it is known that the two genes are highly homologous up to the point where the cDNA clone ends (base 400 in the G2 sequence). Hence the two genes probably diverge between 361 and 400. This is almost certainly within the transcription unit, and possibly between the two potential translational starts.

## 5.5 COMPARISON OF THE cDNA AND GENOMIC CLONES.

### 5.5.1 The G2 gene encodes a different polypeptide from that encoded by pANT-1.

The nucleotide sequence of the genomic clone G2 is presented in Fig.5.8. Inspection of this sequence reveals that it is very similar but not identical to that determined from pANT-1. The two nucleotide sequences are shown aligned in Fig.5.9. Most of the nucleotide substitutions occur in the third position of the codon (60 out of 68). Seven of these substitutions result in a change in the amino acid, at codons 2, 6, 66, 85, 137, 192, and 232 (taking 1 to be the first methionine encoded in pANT-1). The nucleotide and amino acid changes are summarised in Table 5c, as are the restriction site polymorphisms introduced by these substitutions. Of the seven amino acid changes, three are conservative and five are non -





Fig.5.8 Nucleotide and predicted amino acid sequence of the G2 gene.

The arrow indicates the 5' end of the cDNA clone pANT-1 in relation to the G2 sequence.





TABLE 5c

SUMMARY OF DIFFERENCES BETWEEN THE CODING SEQUENCES OF G2 AND pANT-1.

(1) Sequence divergence

NUCLEOTIDES

68 (= 7.1%)

AMINO ACIDS

8 (= 2.5%)

(2) Nucleotide substitutions

BASE IN CODON	EXPRESSED*	SILENT
1st	2	1
2nd	6	0
3rd	<u>1</u>	<u>58</u>
Total	9	59

\*One codon has changes in both 1st and 2nd bases.

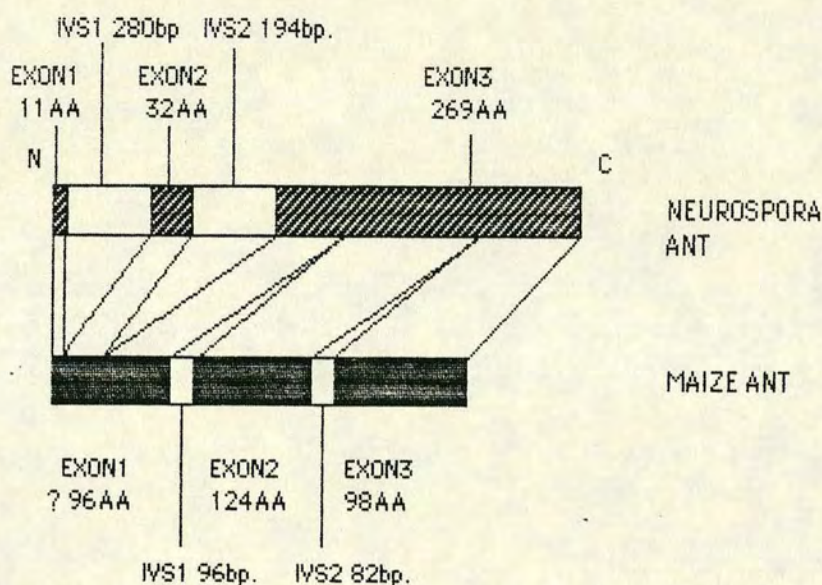
(3) Retriiction site polymorphisms. (co-ordinates taken from Fig.5.8)

G2 seq	pANT-1 seq	co-ordinate	site
CTGCAGCTG	CTGCTGCTG	504	Pst I/PvuII
ATTGAA	ATCGAG	518	Taq I
TGATCA	TGATTA	560	Sau3A/Bcl I
TGGCCA	TGGCTA	951	HaeIII/Bal I
GTGGAT	GTCGAT	1005	Taq I
ATTGAT	ATCGAT	1308	Cla I
AGATCC	AGATTC	1389	Sau3A
GGCCCC	GGGCCC	1403	Apa I
GGCGCT	GGTGCT	1458	Hha I
TGATCT	AGATCT	1673	BglII

conservative. None of these substitutions occur in regions of the polypeptide which are invariant in all translocators sequenced to date. The sequence of pANT-2 thus far determined corroborates these changes and supports the hypothesis that pANT-2 is derived from a transcript of G2. This provides the strongest evidence that the gene G2 is expressed.

### 5.5.2 Intervening sequences in the G2 gene.

When the sequences of pANT-1 and the G2 gene are aligned as in Fig. 5.9, it becomes apparent that the latter contains two intervening sequences which split the coding region into three exons of approximately equal length. Intron 1 is 95 nucleotides in length and interrupts the coding sequence between codons 96 and 97 (using the pANT-1 numbering). Intron 2 is 82 nucleotides long and interrupts the coding sequence between amino acids 220 and 221. In neither instance is a codon split. The yeast translocator gene does not contain introns (Adrian *et al.*, in preparation), while the Neurospora crassa gene, like that of maize, contains two introns of 280 and 194 bp, but inserted at different positions in the coding sequence, Fig. 5.10. There are a number of cases, e.g. actin genes, insulin genes where the position of introns within the gene is conserved even between widely different species (Breathnach and Chambon 1981). This does not appear to be the case for the nucleotide translocator. The asymmetric distribution of the introns in the Neurospora gene, and the occurrence of introns at different locations in the maize gene would suggest that these introns were acquired independently, after the postulated gene duplication which gave rise to the translocator (section 4.4.6, and Saraste and Walker 1982).

Fig.5.10 Introns in the maize and Neurospora ANT genes

INTRON-EXON BOUNDARIES.

		5'splice	3'splice
NEUROSPORA ANT	IVS1	TGCCG	GTAT.....ACAG
	IVS2	ACCAG	GTAC.....CTAG
MAIZE ANT	IVS1	CTCAG	GTAG.....CTAG
	IVS2	TCCAG	GTTT.....GCAG
ANIMAL CONSENSUS		A AG	GTPuA.....C AG G
		C	T G T

The sequences at the 5' and 3' splice junctions of the maize translocator introns conform well to the consensus sequence for eukaryotic nuclear genes (Mount 1982).

Recent publications on the mechanism of eukaryotic pre-mRNA splicing (reviewed by Keller 1984), have indicated that sequences contained within the intron also play a role in splicing. In yeast and higher eukaryotes, excision of introns proceeds via a 'lariat' intermediate, where the free G residue at the 5' end of the intron (generated by a cleavage step

at the 5' intron/exon boundary) becomes attached via a 2'3' phosphodiester bond to an internal A residue. Direct sequencing of these branch point intermediates shows that, although variable, they conform to a consensus 5'-pyXpyTpuApy-3', where A is the branch point. In Saccharomyces nuclear gene introns, the sequence requirement is much stricter and the branch point sequence has been named the TACTAAC box. These sequences occur 18-40 nucleotides from the 3' splice junction, and most introns in nuclear genes from a variety of organisms have a sequence which fits this consensus. Experimental evidence for the functional significance of these sequences is as follows:

(1) Lariat structures have been observed in the electron microscope for a number of genes spliced in vitro (Zeitlin and Efstatidis, 1984)

(2) If the branch point sequence is mutated, cryptic sites may be used (Reed and Maniatis 1985).

(3) Introduction of the beta globin branch point sequence into a synthetic intron greatly enhanced the efficiency of splicing in vitro (Rautman and Breathnach 1985).

A branch point consensus has been derived for plant genes by comparison of intron sequences (Keller and Noon 1985). The maize translocator gene introns contain sequences in the appropriate position which conform to both the plant and general consensus sequences.

Table 5c Potential branch point sequences from the G2 gene introns.

	<u>SEQUENCE</u>	<u>POSITION FROM 3' INTRON BOUNDARY</u>
IVS1	TCCTAA <u>A</u>	-16
IVS2	CTCTAA <u>T</u>	-50
GENERAL	pyXpyTpu <u>A</u> py	
PLANT	pyTpu <u>A</u> (T/C/A)	-18 to -40



Plant and S.cerevisiae genes differ from other eukaryotic nuclear genes in that their introns tend to be small, less than 500bp. In contrast many animal genes have introns several kilobases long. It has been shown that artificially increasing the length of S.cerevisiae introns leads to a reduction in the efficiency with which they are spliced (Klinz and Gallwitz 1985). While this has not been demonstrated directly for a plant gene, two 1.4kb Mu-1 insertions in the first intron of the Adh-1 gene of maize reduce the levels of mRNA produced to 40% and 13% of the wild type (Bennetzen et al., 1984). This could either be due to an effect on transcription or a reduced efficiency of splicing analogous to that seen in the yeast system. If the latter were the case, it would provide an explanation as to why plant introns tend to be small compared to their animal counterparts.

### 5.5.3 The ATP/ADP translocator polypeptide of maize may be significantly larger than those of beef, yeast and Neurospora.

Inspection of the nucleotide sequence of the genomic clone G2 indicates that the long open reading frame which encodes the translocator polypeptide extends beyond the methionine identified as a potential translational start from the pANT-1 sequence, to a second methionine 11 amino acids further upstream (Fig.5.8). This raises the question of which methionine is used for initiation of translation of the protein, and if it is indeed the upstream one, what the implications of the additional sequence are for targeting the polypeptide to the mitochondrion.

In the absence of a cDNA clone covering this region and the inability to determine directly the amino terminal sequence from the purified protein (the amino terminus is blocked) , the probability of each

methionine acting as the initiator codon has to be assessed by indirect, and therefore less reliable, means.

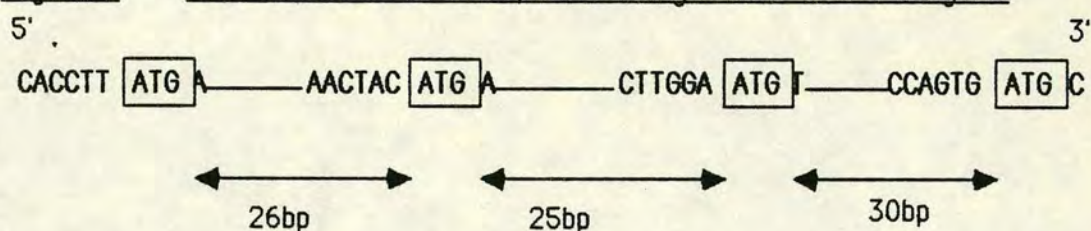
Hybridisation studies suggest that the two copies of the gene diverge within or just 5' to this additional 11 amino acids (section 5.4.3). Amino acid sequence homology between all four translocators sequenced to date falls to essentially zero at the amino terminus. The codon usage of the additional 11 amino acids of the G2 gene is consistent with the rest of the coding region and with pANT-1 (section 4.4.5), but does not constitute proof that this sequence is translated.

The amino terminal sequences of proteins imported into mitochondria, while not sharing any great homology, appear to have certain characteristics (Reid 1984). These are an excess of basic over acidic amino acids and a relatively high content of serine and threonine. Proteins which are synthesised as higher molecular weight precursors, and those imported without proteolytic cleavage both show these features, with the exception of all the adenine nucleotide translocator polypeptides examined to date. They are neither particularly basic, or with the exception of the yeast polypeptide, rich in serine and threonine. This raises the possibility that the translocator may be imported via a different pathway, perhaps interacting with a different kind of receptor, or that some other part of the polypeptide may be important in targeting to mitochondria. Thus there are no features of the deduced amino acid sequence following the two methionines which give a clue as to which is the genuine amino terminus of the polypeptide.

In eukaryotes translation begins at the first ATG codon in ca. 90% of cases (Kozak 1984). However there are a number of cases in the literature, including examples from higher plants, where an internal ATG appears to be used (Hoffman *et al.*, 1982, Hoffman and Donaldson 1985). The sequence

context of the ATG used for initiation appears to be important. Kozak (1984) has suggested 5'-CCpuCCATG-3' as a consensus sequence, with the most important feature being a purine at position -3 (+1 is the A of the ATG), which occurs in 97% of translational starts. The importance of the purine has been demonstrated by site directed mutagenesis (unpublished data cited in Kozak 1984b). In the case of non functional upstream ATG's usually a pyrimidine occurs at -3.

Fig.5.11 ATG codons 5' to the open reading frame in the G2 gene.



The sequences surrounding the two potential initiator codons in the maize ANT are 5'-TTGGAATG-3' and 5'-CAGTGATG-3'. Both have a purine (G) at -3 but neither show a good match to the CCACC consensus, 0/5 for the upstream ATG and 1/5 for the downstream ATG. In considering the merits of the potential starts in these terms it should be noted that two other ATG's occur in the sequence just 5' to the upstream ATG (Fig.5.11). One of these is out of frame with the coding sequence, whilst the other is in frame but could only encounter a 10 amino acid peptide before encountering a translational stop. There is no evidence from the nucleotide sequence of an intron in this region which could, by its excision, remove the stop codon from the mature message. However Kozak (1985) has demonstrated that, at least *in vitro*, ribosomes may start translating at the first ATG encountered and then re-initiate at an internal ATG if a stop codon occurs between the two, as in this case. The sequence surrounding

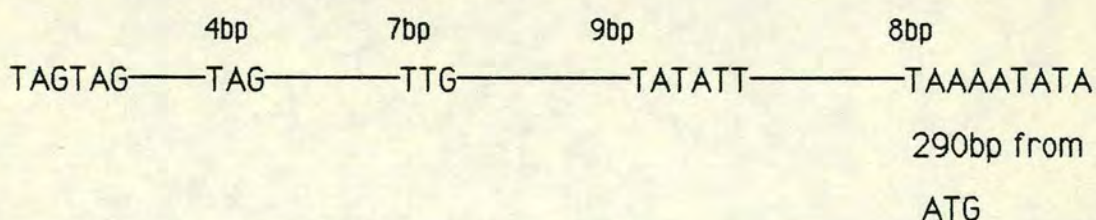
this most distal ATG is 5'-CACCTTATG-3', which has homology to the CCpuCCATG consensus but displaced two nucleotides 5' so that a pyrimidine (C) occupies the -3 position.

Hence consideration of the sequence context of possible initiation codons does not give a clear indication of which may be used in vivo, and indeed raises the possibility that even a third translational start exists.

#### 5.5.4 Potential transcriptional control sequences.

The sequence 5' to the gene was analysed for typical eukaryotic transcriptional control sequences (Breathnach and Chambon 1981). Two TATA like sequences are located at nucleotides 65 to 70, 5'-TATATT-3', and at 78 to 86 5'-TAAAATATA-3'. No CAAT or AGGA boxes, usually found 40-50 nucleotides 5' to the TATA (Messing et al., 1983), occur, but the trinucleotide TTG which is a feature of the -35 homology of prokaryotic promoters occurs at position 53 in the sequence. The -35 motif is located ca.25 bases 5' to the TATA in prokaryotic promoters (Rosenberg and Court 1979), and in this instance is 23 bases from the second TATA like sequence. Further upstream from the TTG, at positions 33-45, the trinucleotide TAG is repeated three times. There is no experimental evidence concerning which if either of these TATA sequences act as promoters in vivo.

Fig.5.12. Potential promoter sequences.



#### 5.5.6 The 5' untranslated region has some unusual features.

The cap site is usually located approximately 30 nucleotides 3' to the

TATA box. If this is the case for the G2 translocator gene, the 5' untranslated sequence is abnormally long, up to 300 nucleotides if translation commences at the methionine included in the cDNA clone pANT-1. This would be in reasonable agreement with the size of the transcript determined from 'Northern blotting', which is ca. 1600 nucleotides, i.e. 300 nucleotides of 5' nontranslated sequence, 954 nucleotides coding region, 180 nucleotides 3' non translated region and 60+ poly A residues. This adds up to 1494+ nucleotides.

In order to determine the transcriptional start, S1 nuclease analysis and primer extension were carried out. These results are preliminary and inconclusive, but are presented here for discussion as they may have some relevance to future work. An example of an S1 mapping experiment is shown in Fig.5.13.

Interpretation of S1 mapping data in the case of the maize adenine nucleotide translocator is complicated due to the fact that the RNA preparation used contains a heterogeneous population of translocator mRNA's derived from at least two different structural genes. Further more the exact extent of sequence homology between the two genes at the 5' end is unknown as yet. An S1 protected fragment will be generated wherever there is a discontinuity between the labelled single stranded DNA used as a probe and the RNA molecule to which it is partially complementary. Therefore an S1 site may represent:

(1) the 5' end of an RNA molecule

(2) a point where the sequence of the probe and mRNA diverge, as might happen if the probe is derived from one copy of the gene and the mRNA from another

(3) a 3' intron/exon boundary, or

(4) a point where the mRNA forms a self complimentary structure in

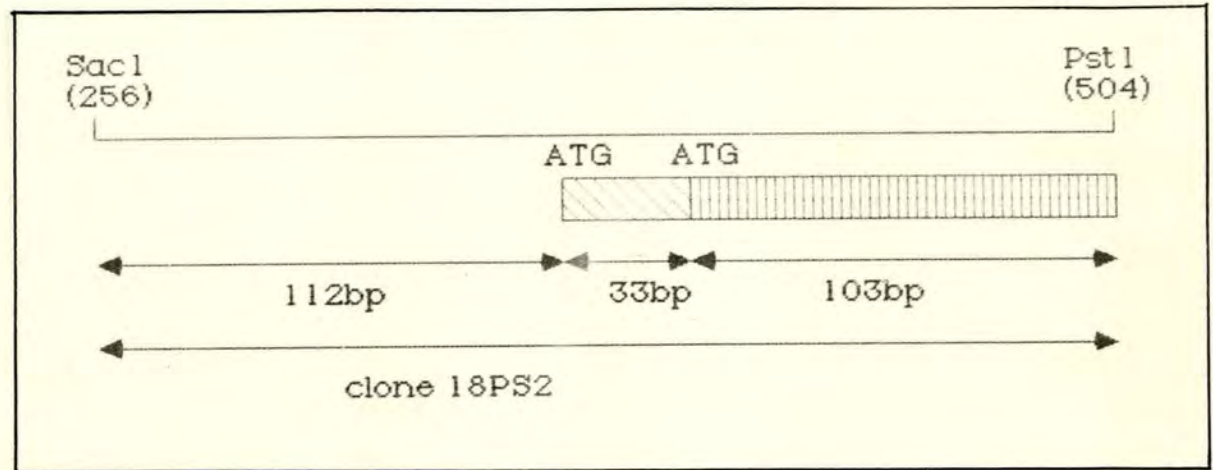
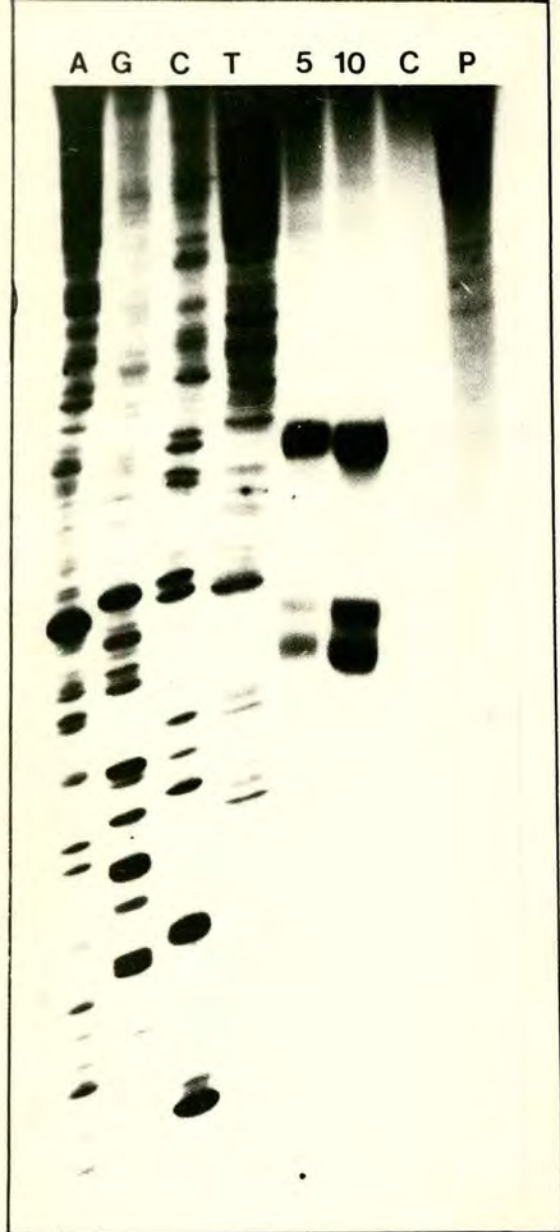


Fig.5.13. S1 nuclease mapping the transcripts of the G2 gene using clone 18/PS2.

Left panel: Autoradiograph of S1 mapping gel. Lanes A,G,C,T = sequencing reaction of clone 18/PS2; 5, 10, = probe incubated with 5 and 10 ug polyA<sup>+</sup> RNA, C = probe + tRNA control, P = 18/PS2 probe alone (no RNA).  
 Right panel: Diagram showing clone 18/PS2 in relation to the G2 sequence.

preference to hybridising with the probe.

For these reasons, primer extension provides useful corroboration of S1 data. The length of the extended primer should correspond to the distance from the primer to the 5' end of the mRNA, unless a strong stop sequence for reverse transcriptase is encountered. So if an S1 site and a site mapped by primer extension co-incide, it is probable that this represents a genuine 5' end of an RNA molecule.

Figure 5.14 indicates the S1 and primer extension sites mapped with three different clones and two different primers. As can be seen from the figure, the RNA has the potential to form considerable secondary structure, although whether any of these stem loop structures are stable enough to form *in vitro* is uncertain. The S1 sites mapped are scattered, whether as a result of the potential secondary structure of the region, or due to inaccuracies in sizing, is unclear. It must be emphasised that these experiments were performed once and have not yet been repeated.

Figure 5.14 serves to emphasise the inverted repeats clustered in the region 236-365, where five pairs of repeats occur. The sequence from 290-322 is of particular interest as it displays 72% homology to a sequence which occurs in the 5' untranslated region of the *Neurospora crassa* ANT transcript. This level of nucleotide sequence homology is actually higher than that observed between the coding regions of these two genes (69%). Like its maize counterpart, this *Neurospora* sequence has the potential to form a stem loop structure. Both these sequences have some homology to the SV40 enhancer core sequence (Weiher *et al.*, 1983).

<b>SV40</b>	<b>TGGAAAG</b>
<b><u>N.crassa</u></b>	<b>GGAAAG</b>
<b><u>Z.mays.</u></b>	<b>TGAAAG</b>





The palindrome is also similar to a sequence which occurs 5' to the yeast CYC7 gene and which has been implicated in the regulation of gene expression (Iborra et al., 1985 and section 1.5.4). Whether these sequences have any functional significance remains purely speculative, but it is tempting to think that some factor involved in modulating gene expression may bind to this region. Sequences with homology to the viral enhancer have been observed upstream of ribulose biphosphate carboxylase small subunit genes in a number of dicotyledenous plants, and in Petunia their presence appears to be correlated with the level of expression of the different genes (Dean et al., 1985 submitted). This sequence also occurs 5' to the sucrose synthase gene from maize which may account for up to 3% of the total protein in endosperm (Werr et al., 1985). While it is tempting to look for 'conserved' sequence motifs between different genes and although this may give some indication of common regulatory mechanisms, caution must be exercised in extrapolating from relatively well characterised systems such as SV40 to such diverse organisms as higher plants where as yet comparatively little is known about the regulation of gene expression.

#### 5.5.6 Is the leader sequence spliced?

The G2 translocator sequence was compared to the only other known gene sequence of a nuclearly encoded mitochondrial protein from a higher plant, the beta subunit of mitochondrial  $F_1$ ATPase from Nicotiana plumbagnifolia (Boutry and Chua, 1985, submitted). The computer programme WORDSEARCH, part of the Wisconsin package on the VAX/VMS was used to search for regions of homology between the two sequences in the hope that some common sequence element which might be related to the expression of the two genes could be identified. Common sequence

elements have been found 5' to genes induced by heat shock in both plants and animals (Czarnecka et al., 1985, Pelham 1982), and possibly in the case of enzymes induced by anaerobiosis (Dennis et al., 1985), which are co-ordinately expressed (Hake et al., 1985). Interestingly the computer identified an unsuspected homology of a different kind. The sequence between 114 and 190 in the G2 sequence is 57% homologous to intron V of the beta gene. The alignment between the two is shown in Fig.5.15. The extent of the homology almost precisely delimits the intron in Nicotiana, and the homologous sequence in maize is flanked by perfect intron/exon boundaries. This is surprising, given that intron sequences are not apparently subjected to a great deal of selective pressure, and tend to diverge rapidly, although often the position of the insertion is conserved. For example maize alcohol dehydrogenase (Adh) genes 1 and 2 are probably derived from a common ancestral gene. Both have nine introns in identical positions which differ in both length and nucleotide sequence. (Dennis et al., 1985).

An alternative explanation is that the sequence common to the Nicotiana beta gene and the maize translocator is some kind of potentially mobile genetic element. However as it is only 76bp long it is clearly too small to be autonomous, and defective transposable elements such as dissociation (Ds) tend to diverge in sequence, sometimes remaining homologous only at the terminal inverted repeats, which are thought to be important in transposition, (Doring and Starlinger 1984). The maize element is not flanked by any sequences which suggest that it might be mobile, the only repeated sequences being the direct repeat AGGT which is homologous with intron/ exon boundaries. Experiments are under way to determine whether this sequence is found elsewhere in the maize genome.

In order to find out whether this sequence is expressed in translocator

115 TCTTATTTGTTTCTCTTTATGTAATG.....TCCTAAGTTTGTGTGT GTTGACATTATGTTGCAGCCAACAGTTTTGAG 189 **maize**  
 || | || | || ||||| || | ||| || | ||| || | || | || | |||  
 2663 TCAAAATTATTTGCTCTTTATCTCATAATGCTATTTCCTTTGTGTTGCTGG CCTGACATATTGTGACCTCAACCATTTCAGAG 2744 **nicotiana**

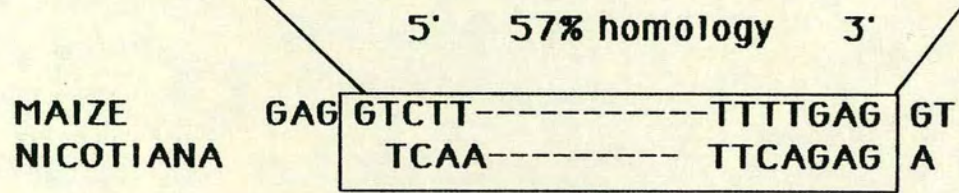


Fig.5.15 Alignment between intron V of the beta subunit of *Nicotiana plumbagnifolium* and the sequence 5' to the G2 maize translocator gene. The numbers are taken from the sequence of Boutry and Chua 1985, and from Fig 5.8.

mRNA, clone 19/A6 was used as a probe in an S1 mapping experiment. A major cluster of S1 protected fragments of 39, 40, 41 and 42 bases were observed, which correspond exactly to the four bases adjacent to the putative intron/exon boundary at 190 Fig.5.8. This very preliminary data suggests that the element may be spliced out of the transcript.

There are several precedents for introns in the 5' untranslated sequences of mRNA's. Chicken, human and rat preproinsulin genes are interrupted in this fashion (quoted in Breathnach and Chambon 1981) as are the genes for Troponin T in rat, (Brietbart et al 1985), alcohol dehydrogenase in Drosophila (Benyajati et al, 1983), alpha amylase in mouse (reviewed by Flavell 1981), and sucrose synthase in maize (Werr et al, 1985). In the case of alpha amylase and alcohol dehydrogenase, differential splicing of the 5' leader sequence occurs in different tissues or in different developmental stages, and may play a role in modulating expression of the gene.

#### 5.5.7 The 3' untranslated region.

The 3' untranslated region of the clones pANT-1 and G2 are clearly related in sequence, but show a greater degree of divergence than the coding sequence. The nucleotide sequence homology between them is 79% compared to 93% for the coding region, and they differ in a number of places by the insertion or deletion of small blocks of nucleotides in one sequence relative to the other. As is the case with pANT-1, the G2 clone lacks the sequence 5'-AATAAA-3', which has been shown to be important for polyadenylation in animal and viral systems (reviewed by Proudfoot 1984). It is unclear whether the 3' ends of plant and yeast mRNA's are formed by an identical mechanism to that operating in higher eukaryotes, where transcription termination is followed by endonucleolytic cleavage

at the selected site and immediate polyadenylation. In the case of yeast genes which, like plant genes, frequently lack AATAAA, the addition of poly A to the transcript appears to be coupled to transcript termination rather than processing (Birnstiel et al., 1985). Furthermore a functional AATAAA sequence is not required for efficient polyadenylation of animal genes expressed in yeast, (Birnstiel et al., 1985) Even in animal cells sequences other than AATAAA are needed to direct selection of the correct sites for cleavage. In particular the consensus sequence pyGTGTTpypy (Mc Lauchlan et al., 1985) which is located 3' to the polyA site of many genes has been implicated by deletion studies to be important in producing wild type levels of mRNA. Other sequences and small nuclear RNA's may also be involved in generating correct 3' ends (Berget 1984). Termination of transcription and polyadenylation in eukaryotes is clearly a complex process and the sequence requirements for this remain largely speculative, particularly in higher plants.

## CHAPTER 6    GENERAL DISCUSSION

### 6.1 AIMS.

The intention of this Chapter is to summarise the major results presented in preceding chapters and to discuss some remaining unresolved questions in the wider context of higher plant genes and genomes. Future areas for research will also be discussed.

### 6.2 SUMMARY OF RESULTS.

The principal achievements of the research presented in this thesis are outlined below:

(1) The construction of cDNA and nuclear genomic libraries from the inbred maize line B37. These libraries have been demonstrated to contain sequences representing nuclear genes, and can be used for the future isolation of additional nuclear gene sequences.

(2) The isolation and characterisation by hybridisation and nucleotide sequence analysis of two cloned cDNAs and their corresponding structural genes encoding the ATP/ADP translocator of maize. This line of investigation has shown that:

(a) The primary sequence and probably the higher order structures of the adenine nucleotide translocators of a variety of organisms, Zea mays, Saccharomyces cerevisiae, Neurospora crassa and Bos taurus are highly homologous.

(b) In contrast to the situation in S. cerevisiae and N. crassa, two separate genes for the adenine nucleotide translocator are transcribed in maize. A third copy of the gene or a related sequence may also be present in the nuclear genome of maize.

(c) DNA sequence analysis suggests that the maize ANT polypeptide may

contain an additional 16 amino acids at the amino terminus compared with the Neurospora polypeptide. This raises the possibility that the ANT polypeptide of maize may undergo processing at the amino terminus upon, or after, import into mitochondria.

(d) The gene G2, which encodes an ANT polypeptide, is split by two introns. These are in different locations compared to the introns in the homologous gene from Neurospora (Arrends and Sebald 1984). In contrast, the Saccharomyces cerevisiae gene is uninterrupted, (Adrian et al., in preparation). The introns contain the normal consensus sequences required for the splicing of eukaryotic pre mRNAs (Mount 1982, Keller 1984).

(e) The 5' untranslated region of the G2 translocator gene is unusually long and may be spliced. This region contains a number of inverted repeat sequences. Their significance, if any, in the regulation of expression of the gene is unknown.

(f) The maize translocator sequences support the contention that the sequence motif 5'-AATAAA-3' is not an essential requirement for the addition of polyA to higher plant mRNAs

(g) It was shown that maize mitochondrial and chloroplast DNA do not contain sequences homologous to the adenine nucleotide translocator.

### 6.3 UNRESOLVED QUESTIONS.

#### 6.3.1 Why are there multiple copies of the translocator gene in maize?

In higher plants, many polypeptides appear to be encoded by multiple genes. These fall into two categories, either quite large gene families which can be further sub divided into groups on the basis of sequence homology, or much smaller families which consist of only 2-3 members. Examples of the former include seed storage proteins,(Sorenson 1984), actins,(Hightower and Meagher 1985) heat-shock proteins,(Key et al.,

1985), the small subunit of RUBP carboxylase, (Dunsmuir et al., 1983) and the chlorophyll a/b binding proteins, (Dunsmuir 1985). The latter class include alcohol dehydrogenase (Dennis et al., 1985), catalase (Sorenson 1982), sucrose synthase (Werr et al., 1985) the ANT from maize, and the beta subunit of mitochondrial ATPase from Nicotiana plumbaginifolia (Boutry and Chua 1985).

While tissue or developmental specific expression of several individual genes within a family have been reported, the biological significance in many cases remains unclear. The small subunit genes of Petunia, where the levels of mRNA derived from individual genes may vary 25 fold in green leaf tissue, but the amino acid sequences of the encoded polypeptides are virtually identical (Dean et al., 1985) are one example. In other cases such as sucrose synthase and alcohol dehydrogenase of maize, the nucleotide sequences of the different genes have diverged sufficiently for them to encode polypeptides with slightly different catalytic or regulatory properties, although it is probable that they originated from a single ancestral gene.

The coding sequences of the two copies of the maize ANT gene have diverged to a greater extent than those of the small subunit genes of Petunia. (Dunsmuir et al., 1983), but less than those of the ADH genes of maize (Dennis et al., 1985), or the chlorophyll a/b binding proteins of Petunia (Dunsmuir 1985). It seems unlikely on the basis of the limited sequence divergence that the properties of the two maize ANT polypeptides are radically different. At present there is no data concerning the individual transcription of the two ANT genes, so whether both are transcribed to the same extent in the same tissue is unknown. The translocator polypeptides purified from beef heart, liver and kidney mitochondria have some different antigenic determinants, (Schultheiss



and Klingenberg, 1984) which implies that in mammals there may also be multiple genes for the translocator which exhibit tissue specific regulation. In the case of cytochrome c oxidase from mammals there are a number of tissue specific variants of the smaller nuclear encoded subunits which might confer slightly different kinetic or regulatory properties upon the complex, (Kadenbach et al., 1982, Kuhn-Nentwig and Kadenbach 1985). In yeast there are two non identical genes for cytochrome oxidase V, both of which are expressed (Cumsky et al., 1985). The COXVa gene encodes the polypeptide normally found in cytochrome oxidase preparations, but COXVb is expressed in vegetatively growing cells and can restore respiration to coxVa structural gene mutants. In Neurospora crassa and Aspergillus nidulans the gene for subunit 9 of the mitochondrial ATPase is encoded in the nucleus. However a copy is present in the mtDNA which, although exhibiting sequence divergence from the nuclear copy, remains a potentially functional gene although its expression has not yet been demonstrated (van den Boogaart et al., 1982, Brown et al., 1985). The significance of tissue and developmental expression is not clear, but it is tempting to believe that the different forms of these subunits may play a role in the modulation of enzyme activity in response to changing cellular demands.

If both genes are transcribed in the same tissue, what is the relationship between the two polypeptides? The functional ANT is a dimer. In maize do homologous or heterologous subunits preferentially associate?

Now that sequence data are available for the two copies of the translocator gene, it will be possible to construct DNA probes specific for each mRNA and so quantify the expression of each in different tissues or at different developmental stages. These experiments may lead to some

insight into the function of the two gene copies.

### 6.3.2 Is the maize translocator made as a larger precursor?

The nucleotide sequence of the genomic clone G2 suggests that the open reading frame encoding the translocator is significantly larger than that encoding the equivalent polypeptide from fungi. It is not clear whether this additional sequence constitutes a pre-piece which functions as a targeting signal and which is proteolytically removed. Formal proof requires the determination of the amino terminal sequence of the purified protein. This has not proved possible as the amino terminal residue is acetylated, rendering the polypeptide intractable to analysis by Edman degradation. It may however prove possible to address this problem by alternative means, as outlined below.

#### 6.3.2.1 Homology between the G1 and G2 genes.

Hybridisation studies undertaken with probes originating from the 5' untranslated sequence and the start of the G2 open reading frame indicate that the nucleotide sequences of the G1 and G2 genes diverge within or just 5' to the first few amino acids of the G2 open reading frame, (section 5.4.3). Isolation and DNA sequence analysis of this region from the G1 gene will identify the point of divergence and determine whether the extended open reading frame present in G2 is common to both copies of the gene. If it is, this would support the hypothesis that translation commences at the more 5' of the two ATGs.

#### 6.3.2.2 Comparison of the in vitro synthesised product and the mature protein.

Higher molecular weight precursors of mitochondrial proteins are

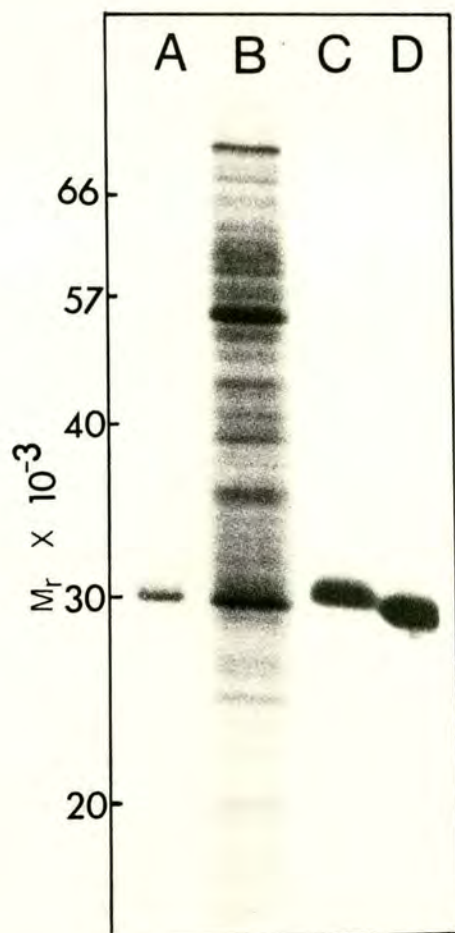
usually identified by virtue of their slower electrophoretic mobility on SDS PAGE. Immunoprecipitation of the in vitro synthesised precursor, followed by SDS PAGE alongside the purified mature translocator protein, would reveal whether the precursor and mature proteins behave differently in polyacrylamide gels. However, even if this were the case, it would not prove that this was due to the presence of an additional sequence, as the precursor and mature proteins are known to have different conformations and may bind SDS to differing extents. The purified translocator protein is known to migrate anomalously in SDS gels, running at ca. 30kD instead of the 40.52kD predicted from sequence analysis, and migrates slightly faster than the 'native' translocator from a solubilised membrane extract, Fig.6.1 .

### 6.3.3 Where does transcription commence?

The availability of a complete set of strand specific M13 subclones covering the entire region from the presumptive TATA box to the coding region will allow identification of the subclone within which transcription initiates. Single strand specific nuclease mapping and primer extension can then be repeated using more specific probes covering this region.

In order to determine whether a splicing event occurs in the 5' untranslated sequence it may be necessary to make a 5' extension cDNA library, as R loop analysis is not sufficiently sensitive to detect intervening sequences of less than 100bp. This would involve the following steps:

- (1) Purification of a single stranded restriction fragment derived from the 5' end of the coding sequence and complementary to the mRNA
- (2) Annealing of this primer fragment to mRNA and extension of the



**FIG.6.1** SDS polyacrylamide gel of the maize ANT.

Lanes A,C purified translocator protein; B,D total mitochondrial protein. A,B stained with Coomassie Blue; C,D replicate gel blotted to nitrocellulose, probed with an antibody to the maize translocator and visualised by autoradiography after incubation with iodinated protein A. (Figure courtesy of Dr. M. Hawkesford)

primer by reverse transcriptase or the Klenow fragment of E. coli polymerase I

(3) Cloning and DNA sequence analysis of the extended products.

#### 6.4 FUTURE DIRECTIONS

The research described in this thesis is part of a larger programme aimed at increasing the understanding of the regulation of mitochondrial biogenesis in plants. The isolation and characterisation of nuclear and mitochondrially encoded genes is an essential first step towards this goal. We have already isolated and analysed a number of genes encoded in mtDNA, and begun to use them to characterise the organisation and expression of mtDNA in fertile and male sterile lines of maize and in the developing wheat leaf (Fox and Leaver 1981, Dawson et al., 1984, Isaac et al., 1985a, Isaac et al., 1985b, J. Topping unpublished data). The isolation and sequence analysis of a cDNA clone for the ANT (Baker and Leaver 1985) is the first report of a cloned nuclear gene encoding a mitochondrial polypeptide from a higher plant. These clones will enable us to complement our studies on mitochondrial gene expression by parallel studies on nuclear genes encoding mitochondrial polypeptides. We have also purified the translocator polypeptide from maize mitochondria and raised an antibody to it (M. Hawkesford, unpublished data), and are in the process of isolating both genes and polypeptides of the respiratory complexes of the mitochondrial inner membrane.

The combination of gene probes and antibodies will allow us to extend our studies to a variety of developmental systems where significant changes in respiration, and presumably also in mitochondrial biogenesis, occur. These include germination, greening of etiolated tissue, fruit ripening and pollen formation. Eventually it should prove possible to

correlate molecular, biochemical and ultrastructural information to provide a complete picture of the role of mitochondria in these important developmental events.

A second area of research is to use cloned nuclear genes to reconstitute protein import in vitro, and to begin to characterise the nature of the 'signals' which allow proteins to partition to the correct subcellular organelle. As discussed in section 1.6, many features of protein import by (fungal) mitochondria and (plant) chloroplasts appear to be similar, but to make valid comparisons between the two will require the use of homologous systems.

Our long term aim is to understand the molecular basis and mechanisms of mitochondrial biogenesis, and how this influences and is influenced by the development of the plant as a whole. The availability of cloned nuclear genes for mitochondrial proteins makes possible both in vitro and in vivo experiments which should increase the understanding of this complex and fascinating process.

BIBLIOGRAPHY

- Ades, I. and Butow, R. (1980) *J. Biol. Chem.* 255 9918-9924.
- Adrian, G., McCammon, M., Montgomery, D. and Douglas, M.G. (1985) in preparation.
- Aquila, H., Eiermann, W., Babel, W., and Klingenberg, M. (1978) *Eur. J. Biochem.* 85 549-560.
- Aquila, H., Misra, D., Eulitz, M. and Klingenberg, M. (1982) *Hoppe-Seyler's Z. Physiol. Chem.* 363 345-349.
- Arrends, H. and Sebald, W. (1984) *EMBO. J.* 3 377-382.
- Austin, J. and Apprille J.R. (1984) *J. Biol. Chem.* 259 154-
- Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. (USA)* 69 1408-1412.
- Baggetto, L., Gaulkeron, D.C. and Godinot, C. (1984) *Arch. Biochem. Biophys.* 232 670-678.
- Bailey-Serres, J.N., Hanson, D.K., Fox, T.D. and Leaver C.J. (in preparation)
- Baker, A., and Leaver, C.J. (1985) *Nucleic Acids Res.* 13 5857-5869.
- Barnes, W.M. (1977) *Science (Wash)* 195 393-394.
- Bedbrook, J., Smith, S.M. and Ellis, R.J. (1980) *Nature* 287 692-697.
- Beggs, J.D. (1978) *Nature* 275 104-108.
- Bennetzen, J.L. and Hall, B.D. (1982) *J. Biol. Chem.* 257 3026-3031.
- Bennetzen, J.L., Swanson, J., Taylor, W.C. and Freeling, M. (1984) *Proc. Natl. Acad. Sci. (USA)* 81 4125-4128.
- Benton, W.D. and Davis, R.W. (1977) *Science (Wash)* 196 180-182
- Benyajati, C., Spoerel, N., Haymerle, H. and Ashburner, M. (1983) *Cell* 33 125-133.

- Berry-Lowe, S.L., McKnight, T.D, Shah, D.M. and Meagher, R.B. (1982) *J. Mol. Appl. Genet.* 1 483-498.
- Berget, S.M. (1984) *Nature* 309 179-181.
- Bewley, J.D. and Black, M. (1985) *Seeds, Physiology of Development and Germination.* pp139-166. Plenum Press, New York.
- Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.* 7 1513-1523.
- Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) *Cell* 41 349-359.
- Block, M.R., Lauquin, G.J.M. and Vignais P.V. (1981) *Biochem.* 20 2692-2699.
- Block, M.R., Lauquin, G.J.M. and Vignais, P.V. (1979) *FEBS Letts.* 104 425-430.
- Boer, P.H., McIntosh, J.E., Gray, M.W. and Bonen, L. (1985) *Nucleic Acids Res.* 13 2281-2292.
- Bogner, W., Aquila, H. and Klingenberg, M. (1983) In 'Structure and Function of Membrane Proteins' (Eds. Quagliariello, E. and Palmieri, F.) pp145-156. Elsevier Science publishers, B.V. Amsterdam.
- Bohni, P., Gasser, S., Leaver, C. and Schatz, G. (1980) In 'Organisation and Expression of the Mitochondrial Genome' (Ed. Kroon, A.M. and Saccone, C.) pp423-433 Elsevier North Holland Biomedical Press.
- Bonen, L., Boer, P. and Gray, M.W. (1984) *EMBO J.* 3 2531-2536.
- Bonner, T.I., Brenner, D.J., Neufeld, B.R. and Britten, R.J. (1973) *J. Mol. Biol.* 81 123-135.
- Boulay, F. and Vignais, P.V. (1984) *Biochemistry* 23 4807-4812.
- Boutry, M. and Chua, N.H. (1985) *EMBO. J.* 4 2159-2168.
- Breathnach, R. and Chambon, P. (1981) *Ann. Rev. Biochem.* 50 349-384.
- Brietbart, R.E., Nguyen, H.T., Medford, R.M., Destree, A.T., Mahdavi, V and Nadal- Ginard, B (1985) *Cell* 41 67-82.



- Brogie, R., Coruzzi, G., Lamppa, G., Keith, B and Chua, N.H. (1983) *Biotechnology* 1 55-61.
- Broome, S. and Gilbert, W. (1978) *Proc. Natl. Acad. Sci. (USA)* 75 2746-2749.
- Brown, T.A., Waring, R.B., Scazzocchio, C. and Davies, R.W. (1985) *Curr. Genet.* 9 113-117.
- Buchanan, B.B., Eiermann, W., Riccio, P., Aquila, H. and Klingenberg, M. (1976) *Proc. Natl. Acad. Sci. (USA)* 73 2280-2284.
- Buell, G.N., Wickens, M.P., Payvar, F., and Schimke, R.T. (1978) *J. Biol. Chem.* 253 2471-2482.
- Carrasco, A.E., McGinnis, W., Gehring, W.J. and De Robertis, E.M. (1984) *Cell* 37 409-414.
- Cashmore, A., Szabo, L., Timko, M., Kausch, A., van den Broeck, G., Schreier, P., Bohnert, H., Herrera- Estrella, L., van Montagu, M. and Schell, J. (1985) *Biotechnology* 3 803-807.
- Chappell, J.B. and Crofts, A.R. (1965) *Biochem. J.* 95 707-716.
- Chirgwin, J.M., Przybyla, A.E., MacDonald, J. and Rutter, W.J. (1979) *Biochemistry* 18 5294-5299.
- Chomyn, A., Mariottini, P., Cleeter, M.W.J., Ragan, C.I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R.F., and Attardi, G. (1985) *Nature* 314 592-597.
- Clark, L. and Carbon, J. (1976) *Cell* 9 91-99.
- Clowes, F.A.L. and Juniper, B.E. (1968) 'Plant Cells' Botanical Monographs Vol. 8 (Ed. Burnett, J.H.) Blackwell, Oxford.
- Contamine, V. and Bolotin-Fukuhara, M. (1984) *Mol. Gen. Genet.* 193 280-287.
- Cooper, T.G. and Beevers, H. (1969) *J. Biol. Chem.* 244 3507-3513.
- Costanzo and Fox (1985) International symposium on Achievements and Perspectives in Mitochondrial Research (Org. Quagliariello, E., Slater, E.C., Palmieri, F., Saccone, C. and Kroon, A.M.) p258. Adriatica Editrice.

- Cumsky, M.G., Ko, C., Trueblood, C.E. and Poyton, R.O. (1985) Proc. Natl. Acad. Sci. (USA) 82 2235-2240.
- Czarnecka, E., Gurley, W.B., Nagao, R.T., Mosquera, L.A. and Key, J.L. (1985) Proc. Natl. Acad. Sci. (USA) 82 3726-3730.
- Dagert, M and Ehrlich, S.D. (1979) Gene 6 23-28.
- Dawson A.J. (1983) PhD. Thesis, University of Edinburgh.
- Dawson, A.J., Hodge, T.P., Isaac, P.G., Leaver, C.J. and Lonsdale, D.M. in preparation.
- Dawson, A.J., Jones, V.P. and Leaver, C.J. (1984) EMBO J. 3 2107-2113.
- Day, D.A. and Wiskich, J.T. (1978) Biochim. Biophys. Acta 501 396-404.
- Dean, C., van den Elzen, P., Tamaki, S., Dunsmuir, P. and Bedbrook, J. (1985) Submitted, EMBO J.
- de Haan, M., van Loon, A.P.G.M., Kriek, J., Vaessen, R.T.M.J. and Grivell, L.A. (1984) Eur. J. Biochem. 138 169-177.
- Deng, G-r., and Wu, R. (1981) Nucleic acids Res. 9 4173-4188.
- Dennis, E.S., Sachs, M.M., Gerlach, W.L., Finnegan, E.J. and Peacock W.J. (1985) Nucleic Acids Res. 13 727-744.
- Dewey, R.E., Schuster, A.M., Levings, C.S. III and Timothy, D.H. (1985) Proc. Natl. Acad. Sci.(USA) 82 1015-1019.
- Dieckmann, C.L., Pape, C.K. and Tzagoloff, A. (1982) Proc. Natl. Acad. Sci. (USA) 79 1805-1809.
- Dingman, C.W. and Peacock, A.C. (1968) Biochemistry 7 659-668.
- Dixon, L.K. and Leaver, C.J. (1982) Plant Mol. Biol. 1 89-102.
- Douce, R., Manella, C.A. and Bonner, W.D.Jr. (1973) Biochim. Biophys. Acta 292 105-116.

- Douglas, M.G., Geller, B.L. and Emr, S.D. (1984) *Proc. Natl. Acad. Sci* 81 3983-3987.
- Doring, H-P., and Starlinger, P. (1984) *Cell* 39 253-259.
- Dujardin, G., Labouesse, M., Netter, P. and Slonimski, P.P. (1983) In 'Mitochondria 1983: Nucleo-mitochondrial Interactions' (Ed. Schweyen, R.J., Wolf, K. and Kaudewitz, F.) pp233-250. deGruyter, Berlin/New York.
- Dujon, B. (1983) In 'Mitochondria 1983: Nucleo-mitochondrial Interactions. (Ed. Schweyen, R.J. Wolf, K. and Kaudewitz, F.) pp1-24. de Gruyter, Berlin/New York.
- Dunsmuir, P. (1985) *Nucleic Acids Res.* 13 2503-2518
- Dunsmuir, P., Smith, S.M. and Bedbrook, J. (1983) *J. Mol. Appl. Genet.* 2 285-300.
- Dunsmuir, P., Smith, S.M. and Bedbrook, J. (1983) *Nucleic Acids Res.* 11 4177-4183.
- Duszynski, J., Groen, A.K., Wanders, R.J.A. Vervoorn, R.C. and Tager, J.M. (1982) *FEBS. Letts.* 146 262-265.
- Efstradiatis, A., Kafatos, F.C., Maxam, A.M. and Maniatis T. (1976) *Cell* 7 279-289
- Ellis, R.J. (1981) *Ann. Rev. Plant Physiol.* 32 111-137.
- Fagan, J.B., Pastan, I. and de Crombrughe, B. (1980) *Nucleic Acids Res.* 8 3055-3064.
- Farrelly, F. and Butow, R.A. (1983) *Nature* 301 296-301.
- Faye, G. and Simon, M. (1983) *Cell* 32 77-87.
- Feinberg, A.P. and Vogelstein, B. (1984) *Anal. Biochem.* 137 266-267.
- Firgaira, F.A., Hendrick, J.P., Kalousek, F., Kraus, J.P. and Rosenberg, L.E. (1984) *Science (Wash)* 226 1319-1322.
- Flavell, R.A. (1981) *Nature* 290 541-542.

- Fox, T.D. and Leaver, C.J. (1981) *Cell* 26 315-323.
- Gasser, S.M., Daum, G. and Schatz, G. (1982a) *J. Biol. Chem.* 257 13034-13041.
- Gasser, S.M., Ohashi, A., Daum, G. Bohni, P.C., Gibson, J., Reid, G.A., Yonetani, T., and Schatz, G. (1982b) *Proc. Natl. Acad. Sci. (USA)* 79 267-271.
- Gasser, S.M. and Schatz, G. (1983) *J. Biol. Chem.* 258 3427-3430.
- Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) *Nucleic Acids Res.* 7 2116-2136
- Gietl, C. and Hock, B. (1982) *Plant Physiol.* 70 483-487.
- Graham, D. (1980) In 'The Biochemistry of Plants, A Comprehensive Treatise' Vol 2 (Davis, D.D. Ed.) pp526-580. Academic Press, New York.
- Groen, A.K., Wanders, R.J.A., Westerhoff, H.V., van der Meer, R. and Tager, J.M. (1982) *J. Biol. Chem.* 257 2754-2757.
- Gronenborn, B. and Messing, J. (1978) *Nature* 272 375-377.
- Groot, G.S.P., Out, T. and Souverijn, J.H.M. (1975) *FEBS Letts.* 49 314-316.
- Grossman, A.R., Bartlett, S.G. and Chua, N.H. (1980) *Nature* 285 625-628.
- Grunstein, M. and Hogness, D. (1975) *Proc. Natl. Acad. Sci. (USA)* 72 3961-3965.
- Guarente, L. (1984) *Cell* 36 799-800
- Guarente, L. and Hoar, E (1984) *Proc. Natl. Acad. Sci. (USA)* 81 7860-7864.
- Guarente, L., LaLonde, B., Gifford, P., and Alani, E. (1984) *Cell* 36 503-511.
- Gubler, U. and Hoffman, B.J. (1983) *Gene* 25 263-269.
- Haffner, M.H., Chin, M.B. and Lane, B.G. (1978) *Can. J. Biochem.* 56 729-733.
- Hake, S., Kelley, P.M., Taylor, W.C. and Freeling, M. (1985) *J. Biol. Chem.* 260 5050-5054.

- Hampp, R., Goller, M., and Ziegler, H. (1982) *Plant Physiol.* 69 448-455.
- Hanson, J.B. and Day, D.A. (1980) In *Biochemistry of Plants, A Comprehensive Treatise*, Vol 1 (Ed. Conn and Stumpf) pp 315-357  
Acaademic Press, London/New York.
- Hase, T., Riezman, H., Suda, K. and Schatz, G. (1983) *EMBO J.* 2 2169-2172.
- Hase, T., Muller, U., Riezman, H. and Schatz, G. (1984) *EMBO J.* 3 3157-3164.
- Hay, R., Bohni, P. and Gasser, S. (1984) *Biochim. Biophys. Acta.* 779 65-87.
- Heber, U. (1974) *Ann. Rev. Plant Physiol.* 25 393-421.
- Heldt, H.W. (1969) *FEBS Letts.* 5 11-14.
- Henderson, P.J.F. and Lardy, H.A. (1970) *J. Biol. Chem.* 245 1319-1326.
- Hennig, B., Koehler, H., and Neupert, W. (1983) *Proc. Natl. Acad. Sci. (USA)*  
80 4963-4967.
- Hiesel, R. and Brennicke, A. (1983) *EMBO J.* 2 2173-2178.
- Highfield, P.E. and Elis, R.J. (1978) *Nature* 271 420-424.
- Hightower, R.C. and Meagher, R.B. (1985) *EMBO J.* 4 1-8.
- Hoffman, L.M., Ma, Y., and Barker, R.F. (1982) *Nucleic Acids Res.* 10  
7819-7828.
- Hoffman, L.M., and Donaldson, D.D. (1985) *EMBO J.* 4 883-889.
- Honda, S.I., Hongladarom, T. and Laties, G.G. (1966) *J. Exp. Bot.* 17 460-472.
- Hohn, B. (1979) *Methods in Enzymology* 68 299-309.
- Horwich, A.L., Fenton, W.A., Williams, K.R., Kalousek, F., Kraus, J.P.,  
Doolittle, R.F., Konigsberg, W. and Rosenberg, L.E. (1984) *Science (Wash.)*  
224 1068-1074.
- Horwich, A.L., Kalousek, F., Mellman, I. and Rosenberg, L.E. (1985) *EMBO J.* 4  
1129-1135.

- Howley, P.M., Israel, M.A., Law, M-F. and Martin, M.A. (1979) *J. Biol. Chem.* 254 4876-4883.
- Hu, N.T. and Messing, J. (1982) *Gene* 17 271-277.
- Hurt, E.C., Pesold-Hurt, B. and Schatz, G. (1984) *FEBS Letts.* 178 306-310.
- Hurt, E.C., Pesold-Hurt, B., Suda, K., Oppliger, W. and Schatz, G. (1985) *EMBO. J.* 4 2061-2068.
- Iborra, F., Francingues, M-C. and Guerineau, M. (1985) *Mol. Gen. Genet.* 199 117-122.
- Isaac, P.G., Brennicke, A., Dunbar, S.M. and Leaver, C.J. (1985) *Curr. Genet.* in press.
- Isaac, P.G., Jones, V.P. and Leaver, C.J. (1985a) *EMBO J.* 4 1617-1623.
- Ise, W., Haiker, H., and Weiss, H. (1985) *EMBO. J.* 4 2075-2080.
- Jackson, D.A., Symons, R.H. and Berg, P. (1972) *Proc. Natl. Acad. Sci. (USA)* 69 2904-2909.
- Kacser, H. and Burns, J.A. (1973) *In Rate Control of Biological Processes* (Davies, D.D. Ed.) pp 65-104 Cambridge University Press, London and New York.
- Kadenbach, B., Hartmann, R., Glanville, R. and Buse, G. (1982) *FEBS Letts.* 38 236-238.
- Kalderon, D., Roberts, B.L., Richardson, W.D. and Smith, A.E. (1984) *Cell* 39 499-509.
- Kao, T-h., Moon, E., and Wu, R. (1984) *Nucleic Acids Res.* 12 7305-7315.
- Kaput, J., Goltz, S. and Blobel, G. (1982) *J. Biol. Chem.* 257 15054-15058.
- Katz, L., Kingsbury, D.T. and Helinski, D.R. (1973) *J. Bacteriol.* 114 577-591.
- Keller, E.B. and Noon W.A. (1985) *Proc. Natl. Acad. Sci. (USA)* 81 7417-7420.
- Keller, W. (1984) *Cell* 39 423-425.

- Kemble, R.J., Mans, R.J., Gabay-Laughnan, S. and Laughnan, J.R. (1983) *Nature* 304 744-747.
- Key, J.L., Gurley, W.B., Nagao, R.T., Czarnecka, W. and Mansfield, M.A. (1985) In 'Molecular Form and Function of the Plant Genome.' NATO ASI Series A, Life Sciences, Vol 83 (Eds. van Vloten-Doting, L., Groot, G.S.P. and Hall, T.C. ) pp81-100. Plenum Press, London/New York.
- Klingenberg, M (1985) In 'The Enzymes of Biological Membranes'. Vol 4, 2nd Ed. (Ed. Martonosi, A.M.) pp511-553. Plenum Press, London/ New York.
- Klingenberg, M. and Appel, M. (1980) *FEBS Letts.* 119 195-200.
- Klingenberg, M. and Heldt, H.W. (1982) In 'Metabolic Compartmentation' (Ed. Sies, H.) pp101-122. Academic Press, London/New York.
- Klingenberg, M., Mayer, I. and Dahms, A.S. (1984) *Biochemistry* 23 2442-2449.
- Klingenberg, M., Riccio, P. and Aquila, H. (1978) *Biochim. Biophys. Acta* 503 193-203.
- Klitz, F.J. and Gallwitz, D. (1985) *Nucleic Acids Res.* 13 3791-3804.
- Korb, H. and Neupert, W. (1978) *Eur. J. Biochem.* 91 609-620.
- Kovac, L., Lachowicz, T.M. and Slonimski, P.P. (1967) *Science (Wash)* 158 1564-1567.
- Kozak, M. (1984a) *Nucleic Acids Res.* 12 857-872.
- Kozak, M. (1984b) *Nucleic Acids Res.* 12 3873-3893.
- Kramer, R. and Klingenberg, M. (1977) *FEBS Letts.* 82 363-367.
- Kramer, R., Aquila, H. and Klingenberg, M. (1977) *Biochemistry* 16 4949-4953.
- Kramer, R., and Klingenberg, M. (1980) *Biochemistry* 19 556-560.
- Kuhn-Nentwig, L. and Kadenbach, B. (1985) *Eur. J. Biochem.* 149 147-158.
- Kyte J. and Doolittle R.F. (1982) *J. Mol. Biol.* 157 105-132.

- Land, H., Grez, M., Hauser, H., Lindenmaier, W and Schutz, G. (1981) *Nucleic Acids Res.* 9 2251-2256.
- Land, H., Grez, M., Hauser, H., Lindenmaier, W and Schutz, G. (1983) *Methods in Enzymology* 100 285-292.
- LaNoue, K.F., and Schoolwerth, A.C. (1979) *Ann. Rev. Biochem.* 48 871-922.
- Laskey, R.A. and Mills, A.D. (1977) *FEBS Letts.* 116 14-16.
- Leaver, C.J. and Gray, M.W. (1982) *Ann. Rev. Plant Physiol.* 33 373-402.
- Leaver, C.J. and Ingle, J. (1971) *Biochem. J.* 123 235-244.
- Leder, P., Tiemeier, D. and Enquist, L. (1977) *Science (Wash)* 196 175-177.
- Lehrach, H.D., Diamond, D., Wozney, J.M. and Boedtker, H. (1977) *Biochemistry* 16 4743-4751.
- Lingappa, V.R., Shields, S., Woo, S.L.C. and Blobel, G. (1978) *J. Cell Biol.* 79 567-572.
- Lingappa, V.R., Lingappa, J.R. and Blobel, G. (1979) *Nature* 281 117-121.
- Lingappa, V.R., Chaidez, J., Yost, C.S. and Hedgpeth, J. (1984) *Proc. Natl. Acad. Sci. (USA)* 81 456-460.
- Lonsdale, D.M., Hodge, T.P. and Fauron, C. M-R. (1984) *Nucleic Acids Res.* 12 9249-9261.
- Lück, D.J.L. and Reich, E. (1964) *Proc. Natl. Acad. Sci. (USA)* 52 931-938.
- Luthe, D.S. and Quatrano, R.S. (1980) *Plant Physiol.* 65 305-308.
- Maarse, A.C., van Loon, A.P.G.M., Riezman, H., Gregor, I., Schatz, G. and Grivell, L.A. (1984) *EMBO. J.* 3 2831-2837.
- Maccacchini, M-L., Rudin, Y. and Schatz, G. (1979) *J. Biol. Chem.* 254 7468-7471.
- Maniatis, T., Kee, S.G., Efstradiatis, A. and Kafatos, F.C. (1976) *Cell* 8 163-182.



- Maniatis, T. Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbour Laboratory.
- Marres, C.A.M., van Loon, A.P.G.M., Oudshoorn, P., van Steeg, H, Grivell, L.A. and Slater, E.C. (1985) *Eur. J. Biochem.* 147 153-161.
- Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. (USA)* 74 560-564.
- McAda, P.C. and Douglas, M.G. (1982) *J. Biol. Chem.* 257 3177-3182.
- McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A. and Gehring, W.J. (1984) *Cell* 37 403-408.
- McGraw, P. and Tzagoloff, A. (1983) *J. Biol. Chem.* 258 9459-9468.
- McLauchlan, J., Gaffney, D., Whitton, J.L. and Clemens, J.B. (1985) *Nucleic Acids Res.* 13 1347-1368.
- McMaster, G.K. and Carmichael, G.G. (1977) *Proc. Natl. Acad. Sci. (USA)* 74 4835-4838.
- Messing, J. (1979) *Recombinant DNA Technical Bulletin*, NIH publication No. 79-99, 2, No. 2, 43-48
- Messing, J. (1983) *Methods in Enzymology* 101 20-78.
- Messing, J., Geraghty, D., Hiedecker, G., Hu, N-T., Kridl, J. and Rubenstein, I. (1983) In '*Genetic Engineering of Plants*'. (Ed. Kosuge, T., Meredith, C.P. and Hollaender, A.) pp211-227. Plenum, New York.
- Messing, J., Groenenborn, B., Muller-Hill, B. and Hofschneider, P.H. (1977) *Proc. Natl. Acad. Sci. (USA)* 74 5463-5467.
- Messing, J., Crea, R., and Seeburg, P.H. (1981) *Nucleic Acids Res.* 9 309-321.
- Messing, J. and Vieira, J. (1982a) *Gene* 19 259-268.
- Messing, J. and Vieira, J. (1982b) *Gene* 19 269-276.
- Michelson and Orkin (1982) *J. Biol. Chem.* 257 14773-14782.
- Miura, S., Mori, M. and Tatibana, M. (1983) *J. Biol. Chem.* 258 6671-6674.

- Montgomery, D.M., Leung, D.W., Smith, M., Shalit, P., Faye, G., and Hall, B.D. (1980) Proc. Natl. Acad. Sci. (USA) 77 541-545.
- Moon, E., Kao, T-h. and Wu, R. (1985) Nucleic Acids Res. 13 3195-3212.
- Moore, A.L. and Bonner, W.D.Jr. (1981) Biochim. Biophys. Acta. 634 117-128.
- Moore, A.L., Jackson, C., Halliwell, B., Dench, J.E., and Hall, D.O. (1977) Biochem. Biophys. Res. Commun. 78 483-491.
- Morel, F., Lauquin, G., Lunardi, J., Duszynski, J. and Vignais, P.V. (1974) FEBS Letts. 39 133-138.
- Morohashi, K., Fuji-Kuriyama, Y., Okada, Y., Sogawa, K., Hirose, T. and Inayama, S (1984) Proc. Natl. Acad. Sci. (USA) 81 4647-4651.
- Morohashi, Y. and Bewley, J.D. (1980) Plant Physiol. 66 70-73.
- Morohashi, Y., Bewley, J.D., and Yeung, E.C. (1981) J. Exp. Bot. 32 605-613.
- Mount, S.M. (1982) Nucleic Acids Res. 10 459-472.
- Muller, P.P., Reif, M.K., Zonghou, S., Sengstag, C., Mason, T.L. and Fox, T.D. (1983) J. Mol. Biol. 175 431-452.
- Nagata, S., Tsunetsugu-Yokata, Y., Naito, A., and Kaziro, Y. (1983) Proc. Natl. Acad. Sci. (USA) 80 6192-6196.
- Nelson, T and Brutlag, D. (1979) Methods in Enzymology 68 41-50.
- Ohashi, A., Gibson, J., Gregor, I. and Schatz, G. (1982) J. Biol. Chem. 257 13042-13047.
- Ohta, S. and Schatz, G. (1984) EMBO J. 3 651-657.
- Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2 161-170.
- O'Malley, K., Pratt, P., Robertson, J., Lilly, M. and Douglas, M.G. (1982) J. Biol. Chem. 257 2097-2103.
- Palmer, J., Shields, C., Jorton, T. and Cohen, D. (1982) Nature 301 725-728.

- Palmer, J.D. and Shields, C.R. (1984) *Nature* 307 437-440.
- Palmer, J.M. (1976) *Ann. Rev. Plant Physiol.* 27 133-157.
- Peacock, S.L., McIver, C.M. and Monahan, J.J. (1981) *Biochim. Biophys. Acta.* 655 243-250.
- Pelham, H.R.B. (1982) *Cell* 30 517-528.
- Perlman, P.S. and Mahler, H.R. (1974) *Arch. Biochem. Biophys.* 162 248-271.
- Pfaff, E., Klingenberg, M. and Heldt, H.W. (1965) *Biochim. Biophys. Acta* 104 312-315.
- Pfaff, E., and Klingenberg, M. (1968) *Eur. J. Biochem.* 6 66-79.
- Pillar, T., Lang, B.F., Steinberger, I., Vogt, B. and Kaudewitz, F. (1983) *J. Biol. Chem* 258 7954-7959.
- Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E. and Surrey, S. (1982) *Proc. Natl. Acad. Sci. (USA)* 79 4298-4302.
- Pratje, E., Manhaupt, G., Michaelis, G. and Beyreuther, K. (1983) *EMBO J.* 2 1049-1054.
- Proudfoot, N.G. (1984) *Nature* 307 412-413.
- Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature* 263 211-214.
- Radloff, R., Bauer, W. and Vinograd, J. (1967) *Proc. Natl. Acad. Sci. (USA)* 57 1515-1521.
- Rautmann, G. and Breathnach, R. (1985) *Nature* 315 430-432.
- Reed, R., and Maniatis, T. (1985) *Cell* 41 95-105.
- Reid, G.A. (1984) In 'Current Topics in Membrane Transport' Vol 24 (Ed. Martonosi, A.) pp295-335. Academic Press, London/New York.
- Reid, G.A., Yonetani, T. and Schatz, G. (1982) *J. Biol. Chem.* 257 13068-13074.

- Ricciardi, R.P., Miller, J.S. and Roberts, B.E. (1979) Proc. Natl. Acad. Sci. (USA) 76 4927-4931.
- Riezman, H., Hay, R., Witte, C., Nelson, N. and Schatz, G. (1983) EMBO J. 2 1113-1118.
- Rigby,, P.W.J., Dieckmann, M., Rhoades, C. and Berg, P. (1977) J. Mol. Biol. 113 237-251.
- Robinson, C. and Ellis, R.J. (1984) Eur. J. Biochem. 142 337-342.
- Rosenberg, M. and Court, D. (1979) Ann. Rev. Genetics 13 319-353.
- Ruther, U. and Muller-Hill, B. (1982) EMBO J 2 1719-1794.
- Sadler, I., Suda, K., Schatz, G., Kaudewitz, F. and Haid, A. (1984) EMBO. J. 3 2137-2143
- Saltzgaber-Muller, J., Kunapuli, S.P. and Douglas, M.G. (1983) J. Biol. Chem. 258 11465-11470.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. (USA) 74 3643-3646.
- Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143 161-178.
- Saraste, M. and Walker, J.E. (1982) FEBS Letts. 144 250-254.
- Scarpulla, R.C., Agne, K.M., and Wu, R. (1981) J. Biol. Chem. 256 6480-6486.
- Scarpulla, R.C. and Wu, R (1982) Proc. Natl. Acad. Sci. (USA) 79 739-743.
- Schleyer, M., Schmidt, B. and Neupert, W (1982) Eur. J. Biochem. 125 109-116.
- Schmidt, B., Hennig, B., Kohler, H. and Neupert, W. (1983) J. Biol. Chem. 258 4687-4689.
- Schrier, P.H., Seftor, E.A., Schell, J. and Bohnert, H.J. (1985) EMBO J. 25-32.
- Schultheiss, H.P. and Klingenberg, M. (1984) Eur. J. Biochem. 143 599-605.

- Schuster, W. and Brennicke, A. (1985) *Curr. Genet.* 9 157-163.
- Shah, D.M., Hightower, R.C. and Meagher, R.B. (1983) *J. Mol. Appl. Genet.* 2 111-126.
- Sharp, P.A., Berk, A.J. and Berget, S.M. (1980) *Methods in Enzymology* 65 750-768.
- Shore, G.C., Power, F., Bendayan, M. and Carignan, P. (1981) *J. Biol. Chem.* 256 8761-8766.
- Singh, P. and Naik, M.S. (1984) *FEBS Letts.* 165 145-150.
- Smeekins, S., van Binsbergen, J and Weisbeck, P. (1985a) *Nucleic Acids Res.* 13 3179-3194.
- Smeekins, S., de Groot, M., van Binsbergen, J. and Weisbeck, P. (1985b) *Nature* 317 456-458.
- Smith, M., Leung, D. W., Gillams, S., Astell, C.R., Montgomery, D.M. and Hall, B.D. (1979) *Cell* 16 735-761.
- Smith, S.M., Bedbrook, J. and Speirs, J. (1983) *Nucleic Acids Res.* 11 8719-8734.
- Smith, S.M., and Leaver, C.J. (1985) Submitted to *Eur. J. Biochem.*
- Solomos, T. (1977) *Ann. Rev. Plant Physiol.* 28 279-297.
- Southern, E.M. (1975) *J. Mol. Biol.* 98 503-517.
- Sorenson, J.C. (1982) In 'Maize for Biological Research' (Ed. Sheridan, W.F.), pp135-138. Plant Molecular Biology Association.
- Sorenson, J.C. (1984) In 'Molecular Genetics of Plants' (Ed. Scandalios, J.G. and Caspari, E.W.) *Advances in Genetics*, Vol 22, Academic Press.
- Stern, D.B. and Lonsdale, D.M. (1982) *Nature* 299 698-702.
- Stern, D.B. and Palmer, J.D. (1984) *Proc. Natl. Acad. Sci. (USA)* 81 1946-1950.
- Suissa, M. and Schatz, G. (1982) *J. Biol. Chem.* 257 13048-13055.

- Szekely, E. and Montgomery, D. (1984) *Mol. Cell. Biol.* 4 939-946.
- Tager, J.M., Wanders, R.J.A., Groen, A.K., Kuntz, W., Bohnensack, R., Kuster, U., Letko, G., Bonme, G., Duszynski, J. and Wojtczack, L. (1983) *FEBS Letts.* 151 1-9.
- Timmis, J.N. and Scott, N.S. (1983) *Nature* 305 65-67.
- Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. (USA)* 77 5201-5205.
- Twigg, A.J. and Sherratt, D. (1980) *Nature* 283 216-218.
- Tzagoloff, A. (1982) *Mitochondria*. Plenum Press London/New York.
- Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M. (1977) *Science (Wash)* 196 1313-1319
- van den Boogart, P., Samallo, J. and Asgsteribb, E. (1982) *Nature* 298 187-189.
- van den Broeck, G., Timko, M.P., Kausch, A.P., Cashmore, A.R., van Montagu, M. and Herrera-Estrella, L. (1985) *Nature* 313 358-363.
- van Loon, A.P.G.M., (1984) PhD. Thesis, University of Amerstam.
- van Loon, A.P.G.M., de Groot, R., van Eyk, E., van der Horst, G.T.J. and Grivell, L.A. (1982) *Gene* 20 323-337.
- van Loon, A.P.G.M., van Eijk, E. and Grivell, L.A. (1983) *EMBO. J.* 2 1765-1770.
- van Loon, A.P.G.M., Maarse, A.C., Riezman, H. and Grivell, L.A. (1984) *Gene* 26 261-272.
- Viebrock, A., Perz, A., and Sebald, W. (1982) *EMBO. J.* 1 565-571.
- Vignais, P.V. (1976) *Biochim. Biophys. Acta* 456 1-38.
- Vignais, P.V. and Vignais, P.M. (1972) *FEBS Letts.* 26 27-31.
- Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) *EMBO J.* 1 945-951.

- Walter, P., Gilmore, R. and Blobel, G. (1984) *Cell* 38 5-8.
- Ward, B.L., Anderson, R.S. and Bendich, A. (1981) *Cell* 25 793-803.
- Weiher, H., Konig, M. and Gruss, P. (1983) *Science* 219 626-631.
- Werr, W., Frommer, W-B., Maas, C. and Starlinger, P. (1985) *EMBO J.* 4 1373-1380.
- Wheeler, A.M., and Hartley, M.R. (1975) *Nature* 257 66-67
- Wienand, U. and Feix, G. (1980) *FEBS Letts.* 116 14-16.
- Wiskich, J.T. (1977) *Ann. Rev. Plant Physiol.* 28 45-69.
- Wright, R.M. and Cummings, D.J. (1983) *Nature* 302 86-88
- Wright, R.M., Ko, C., Cumsky, M.G. and Poyton, R.O. (1984) *J. Biol. Chem.* 259 15401-15407.
- Yaffe, M.P., Ohta, S. and Schatz, G. (1985) *EMBO. J.* 4 2069-2074.
- Yamamoto, K.R., Alberts, B.M., Benzinger, R., Lawhorne, L. and Treiber, G. (1970) *Virology* 40 734-744.
- Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* 33 103-119.
- Young, R.A. and Davis, R.W. (1983) *Proc. Natl. Acad. Sci. (USA)* 80 1194-1198.
- Zeitlin, S. and Efstratiadis, A (1984) *Cell* 39 589-602.
- Zimmerman, R., Paluch, U., Sprinzel, M. and Neupert, W. (1979) *Eur. J. Biochem.* 99 247-252.
- Zimmerman, R. and Neupert, W (1980) *Eur. J. Biochem.* 109 217-229.
- Zitomer, R.S., Montgomery, D.M., Nichols, D.L. and Hall, B.D. (1979) *Proc. Natl. Acad. Sci. (USA)* 76 3627-3631.
- Zwizinski, C. and Neupert, W. (1983) *J. Biol. Chem.* 258 13340-13346.

---

**Isolation and sequence analysis of a cDNA encoding the ATP/ADP translocator of *Zea mays* L.**

---

Alison Baker and Christopher J. Leaver\*

---

Department of Botany, The King's Buildings, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JH, UK

---

Received 1 July 1985; Revised and Accepted 5 August 1985

---

**ABSTRACT**

A cDNA complementary to the mRNA for the ATP/ADP translocator of maize (*Zea mays* L.) has been identified by virtue of hybridisation with the homologous gene from yeast. The cloned cDNA has been shown by DNA sequence analysis to contain an open reading frame of 954bp., which encodes a polypeptide of molecular weight 40,519. This polypeptide exhibits a high degree of homology to the translocator polypeptides of beef heart and *Neurospora crassa* mitochondria.

**INTRODUCTION**

An essential prerequisite for growth and differentiation of higher plants is the ability of the mitochondrion and chloroplast to meet the changing energy requirements of the plant cell. Many developmental transitions are associated with, or dependent upon, marked changes in mitochondrial number, structure and metabolism. As part of a programme directed towards an understanding of how genetic, environmental and physiological factors regulate mitochondrial biogenesis and function, we are analysing the structure and expression of genes encoding components of the mitochondrial inner membrane.

The mitochondrial genome encodes less than 10% of the total mitochondrial proteins (ca. 20 polypeptides), most of which are polypeptides of the major enzyme complexes of the inner membrane. We have isolated and analysed mitochondrial genes for subunits of the cytochrome c oxidase (1,2) and bc<sub>1</sub>, respiratory complexes (3), and the F<sub>1</sub> ATP-synthase complex (4) of maize. However to date no report has been published of the isolation of a nuclear gene encoding a mitochondrial protein from a higher plant. As a first step towards the study of nuclear genes involved in mitochondrial biogenesis, we report here the identification and sequence analysis of a cDNA clone for the ATP/ADP translocator from *Zea mays* L.

The functional adenine nucleotide translocator is a homodimer with the individual subunits having an estimated molecular weight (by SDS PAGE) of



30,000 (bovine) to 33,000 (Neurospora) (5,6). It is the major polypeptide of the mitochondrial inner membrane, comprising about 10% of the total mitochondrial protein in maize. It shows a high degree of specificity for ATP and ADP, and mediates the transfer of energy generated by oxidative phosphorylation to the energy consuming processes of the cytosol (7). As a consequence it forms an important interface between mitochondrial and cytosolic metabolic pathways, and under some conditions exerts a significant control over the rate of oxidative phosphorylation by regulating the supply of ADP to the matrix as a substrate for the ATP synthase (8,9). The translocator may thus have a pivotal role in plants where both developmental and diurnal changes occur in the relative contributions from oxidative phosphorylation and photophosphorylation to cellular ATP levels.

### MATERIALS

Maize seed with a B37 nuclear background and the N (fertile) mitochondrial genotype, was supplied by Pioneer Hi-Bred International, Des Moines, Iowa  $\alpha$ [<sup>32</sup>P]dCTP, 410Ci/mmol, triethylammonium salt, was purchased from Amersham International. Enzymes were purchased from BRL, Amersham, BCL, NBL, and P&S Biochemicals. X-ray film was Dupont Cronex 4 or FujiRX (sequencing gels).

### EXPERIMENTAL PROCEDURES

#### Construction of a cDNA Library from Zea mays L.

Total nucleic acids were prepared (10) from coleoptiles of maize seedlings germinated for 60h at 28-29°C in total darkness. Poly(A) RNA was prepared by two cycles of elution from oligo dT cellulose (11). cDNA synthesis was performed essentially as described (12,13). First strand reaction contained 4 $\mu$ g poly(A) RNA per 50 $\mu$ l incubation mix (50mM Tris-HCl pH 8.3, 20mM KCl, 8mM MgCl<sub>2</sub>, 10mM dithiothreitol, 1mM each of dATP, dGTP, dCTP and dTTP (dNTP's), 150ng oligo dT) plus 30U placental RNase inhibitor and 30U reverse transcriptase. Incubation was for 10 min on ice followed by 60 min at 42°C. After boiling followed by rapid cooling on ice, 50 $\mu$ l of 2x second strand buffer (200mM HEPES-KOH pH 7, 2mM each dNTPs, 6mM KCl) and 15U of the Klenow fragment of DNA polymerase 1 were added and the reaction incubated at 15°C for 5h. Unincorporated nucleotides were removed by gel filtration through Sephadex G50. The double stranded cDNA was treated with S1 nuclease and size selected on sucrose gradients. The fraction having a mean size around 1200 nucleotides was inserted into the Pst 1 site of

plasmid pAT153 by G-C homopolymer tailing, used to transform *E. coli* strain 294, and a library of approximately 2,500 tetracycline resistant colonies obtained. These clones were stored frozen in microtitre plates at  $-80^{\circ}\text{C}$ . Filter replicas of the clone bank were prepared on Whatmann 541 paper and used in hybridisation studies (14).

#### Identification of a cDNA clone encoding the ATP/ADP translocator

The library was screened by colony hybridisation using a Bam H1 fragment from plasmid pYe OP1 (2.6), containing the entire yeast translocator gene (15). The Bam fragment was labelled with  $^{32}\text{P}$  by nick translation (16) to a specific activity of  $10^6$ - $10^7$  cpm per  $\mu\text{g}$ . Filters (7cm x 5cm) were prehybridised in 50% (v/v) formamide, 5xSSC pH 7, 150 $\mu\text{g}/\text{ml}$  denatured herring sperm DNA, in a total volume of 4ml per filter at  $25^{\circ}\text{C}$ , for 2-16h. Hybridisation was for 48h in the same buffer, but including the probe. Washes were 4x30 min in 2xSSC at room temperature. Filters were air dried and exposed to preflashed X-ray film, at  $-80^{\circ}\text{C}$  with intensifying screens for up to three weeks.

Potential positive colonies were lysed with hot sodium dodecyl sulphate as described (17), and the crude nucleic acid extracts separated by electrophoresis on 1% (w/v) agarose gels, transferred to nitrocellulose (18) and rescreened with the yeast gene probe as above, except that 10x Denhardt's solution and 0.1% (w/v) SDS were included in both pre-hybridisation and hybridisation buffers. Small amounts of plasmid DNA were prepared by the method of Birnboim and Doly (19).

Colony hybridisations using maize translocator probes were performed in 4xSSC, 10x Denhardt's solution, 0.1% (w/v) SDS, 150 $\mu\text{g}/\text{ml}$  denatured herring sperm DNA at  $65^{\circ}\text{C}$ .

#### DNA sequence analysis of the translocator clone pANT-1

The clone pANT-1 was sequenced using the M13/dideoxy chain terminator method (20,21). Fragments were derived from plasmid pANT-1 using restriction endonucleases recognising tetranucleotide sequences within the cDNA insert, and ligated into the appropriate site(s) of M13mp8 or mp11 replicative form (RF) DNA. Clones were also generated by Bal31 exonuclease digestion of the Pst1-digested plasmid, followed by blunt ended ligation to Sma1-cut M13 RFDNA.

Single stranded M13 DNA was purified as described (22), and sequenced according to Sanger et al., (21) except that the synthetic M13 primer was used to prime the sequencing reaction (23).

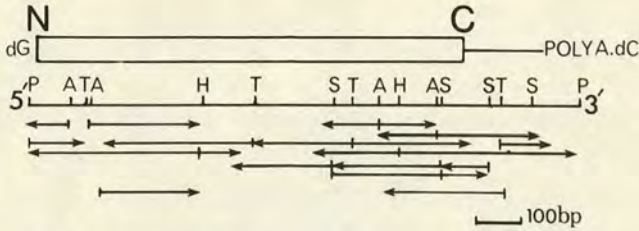
### Transcript analysis

Total cytosolic and poly(A) RNA were fractionated by electrophoresis in 1.25% (w/v) formaldehyde-agarose gels (24), transferred to nitrocellulose (25) and hybridised with either nick-translated plasmid DNA from pANT-1 or with labelled strand-specific M13 subclones (26). Prehybridisation was in 50% (w/v) formamide, 5xSET, 50mM NaPO<sub>4</sub> pH 7, 0.1% (w/v) sodium pyrophosphate, 0.1% (w/v) SDS, 10xDenhardt's solution, 100µg/ml herring sperm DNA, at 42°C for 2-6h (20xSET is 3M NaCl, 0.6M Tris-HCl pH 8, 20mM Na<sub>2</sub>EDTA pH 8). Hybridisation was overnight at 42°C in the same buffer, except that the concentration of Denhardt's solution was reduced ten-fold. Washes were performed in 2xSSC, 0.1% (w/v) SDS, 0.1% (w/v) sodium pyrophosphate at 50-55°C.

### RESULTS AND DISCUSSION

#### Identification of translocator clones

cDNA clones containing the maize translocator sequence were tentatively identified by colony hybridisation with a probe containing the entire yeast ATP/ADP translocator gene (15). In order to confirm their identity, candidate clones were subjected to a second round of screening. Crude nucleic acid extracts of these colonies were prepared (17), the recombinant plasmids separated from *E. coli* chromosomal DNA by agarose gel electrophoresis, transferred to nitrocellulose and hybridised under conditions of reduced stringency, with the yeast gene probe. One clone exhibited intense hybridisation to the probe, and subsequent restriction analysis of purified plasmid DNA from this colony revealed that this was due to homology with the ca. 1200 nucleotide cDNA insert present in this plasmid. This clone was designated pANT-1. When the cDNA library was re-examined using an M13 subclone of pANT-1 as a homologous probe under stringent hybridisation conditions, an additional two clones were identified. One clone contained a cDNA insert smaller than pANT-1 and had lost one of the flanking Pst 1 sites, while the other contained an insert of similar size to pANT-1 as judged by agarose gel electrophoresis. This represents a frequency of ca. 0.2% for translocator clones in the library, and is comparable with that observed for cDNA clones of the translocator and also the nuclear encoded proteolipid subunit of the F<sub>0</sub> portion of the ATP synthase in a cDNA library from *Neurospora* (27). The translocator is one of the most abundant inner membrane polypeptides, so in order to obtain clones of less abundant proteins a larger library may be required.



**Figure 1.** Restriction map of the cDNA insert of plasmid pANT-1 illustrating the strategy used for determining the DNA sequence. The extent of the open reading frame (box), 3' untranslated region and poly(A) are indicated in relation to the map. N and C denote the amino and carboxy termini of the protein. P = PstI A = AluI T = TaqI H = HaeIII S = Sau3A.

#### DNA sequence analysis

The DNA sequence of the entire cDNA insert of pANT-1 was determined from independent overlapping subclones covering the entire length (Fig. 1). Although in two regions, from the Taq I site at 1063 to the poly A and from the Taq I site at 468 to the Sau 3A site at 655, the sequence was determined from one strand only. Considerable difficulty was encountered in obtaining sequence from the ends of the insert which contained the homopolymer tails, as these fragments are recovered as subclones rather less frequently than expected. Furthermore, it only proved possible to sequence these clones in one orientation, from sites within the insert outwards towards the ends, because as soon as the homopolymer tails were reached the sequencing reaction would either terminate or generate spurious bands ('laddering'). The most fruitful approach to solving this problem proved to be the construction of clones from double digests of Pst I with Hae 111, Alu 1, or Taq 1, and ligating these fragments into the vector such that the sequences of interest lay adjacent to the primer hybridisation site of M13.

DNA sequence analysis of the maize cDNA clone pANT-1 reveals an open reading frame of 954 nucleotides (318 amino acids), 180 nucleotides of 3' untranslated sequence and ca. 60 nucleotides of the poly(A) tail (Fig. 2). The deduced amino acid sequence of the encoded polypeptide is highly homologous to previously published sequences for the ATP/ADP translocator of both Neurospora crassa and Bos taurus (28,29) (Fig. 3). The amino acid sequences from all three organisms are identical in 142 positions, while maize and Neurospora, and maize and beef, respectively have 238 (74.8%) and 160 (50.3%) amino acids in common. When the sequences are aligned, the maize open reading frame continues five amino acids beyond the amino





vs 33,000, (estimated) for Neurospora.

The 3' untranslated region does not contain an obvious polyadenylation signal conforming to the animal consensus AATAAA (30). This observation supports the suggestion that many plant genes differ from animal genes in this respect. A variation of this sequence, GATAA<sub>1-3</sub>, is usually located 15-40 bases from 5' to the site of polyadenylation (31). The maize adenine nucleotide translocator gene lacks the AATAAA motif, but two sequences of AATAAT do occur, one 8 nucleotides 3' to the termination codon and one 76 nucleotides 5' to the point of poly(A) addition. The sequence AATAAT is found in the zein subfamily B49 of maize (31), and in maize alcohol dehydrogenase-2 (32), which both lack the canonical AATAAA motif. However, none of these sequences have been shown to be sufficient or even necessary for correct polyadenylation of higher plant messenger RNA's.

### Codon usage

Despite the high overall nucleotide sequence homology, 69% in the coding region, codon usage between the Neurospora and maize translocator genes is quite different. The maize gene exhibits a marked bias towards T in the 3rd position of the codon, where it occurs 38% of the time. In the Neurospora gene 55% of the codons end with C, compared with 28% for maize. Both genes show a bias against codons ending in A, 8% for maize and less than 1% for Neurospora. This codon usage is in marked contrast to that observed for several other maize nuclear genes.

In yeast and E. coli it is clear that codon usage bias relates to the level of expression of many genes, the higher the level of expression, the greater the degree of bias towards codons corresponding to the major isoaccepting tRNA species (33). It has been suggested that this avoids depletion of the pool of charged tRNA's of rarer species, which might lead to an inhibition of translation or an increase in missense reading. In comparison to yeast and E. coli, relatively few plant genes have been sequenced, so it remains to be seen whether similar rules of codon bias apply and the significance, if any, of the observed differences in codon usage between the maize translocator and some other maize genes.

### RNA transcripts

Samples of total cellular RNA and the poly(A) RNA preparation used to construct the library were fractionated by electrophoresis in formaldehyde-agarose gels and transferred to nitrocellulose. When hybridised with either nick-translated pANT-1 plasmid DNA or with strand specific M13 subclones of the plasmid, a single transcript of ca. 1600 nucleotides was detected

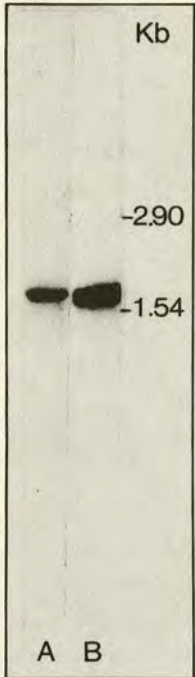


Figure 4. RNA transcripts of the maize adenine nucleotide translocator. Panel (A) represents 15 $\mu$ g of total RNA, panel (B) 2 $\mu$ g of poly(A) RNA, from 60h old dark grown maize coleoptiles, which were fractionated by electrophoresis in the same gel. RNA was transferred to nitrocellulose, probed with nick translated plasmid pANT-1, and autoradiographed. Size markers are *E. coli* rRNA.

(Fig. 4). The cloned cDNA is 1196 nucleotides long including the cloned portion of the poly(A) tail. This means that an additional 400 nucleotides remains to be accounted for. Assuming that the transcript observed represents a mature message and that 50-200 nucleotides is a reasonable estimate of the length of the poly(A) tail, then the 5' untranslated sequence may be in excess of 250 nucleotides. This would be unusual as several plant messengers have been shown to possess quite short leader sequences of less than 100 nucleotides (34).

Predicted higher order structure and arrangement of the translocator polypeptide in the mitochondrial inner membrane

The hydropathy profiles of the beef heart, *Neurospora* and maize translocator polypeptides are very similar and can virtually be superimposed one upon another, which reflects the remarkable degree of conservation seen in the primary sequence (data not shown). Six hydrophobic domains are apparent which may be membrane traversing alpha-helices (35). These segments are separated by regions of a more hydrophilic nature. Saraste and Walker (35) noted that the bovine polypeptide contains a marked internal repeat, and proposed that the gene evolved by triplication of an ancestral gene.



This repeated feature is also striking in the maize sequence at both the amino acid and nucleotide level, and supports the idea that a gene triplication took place before the progenitors of mammals, fungi and higher plants diverged.

Data derived from pyridoxal phosphate labelling studies of the beef translocator complexes with the inhibitors carboxyatractyloside (CAT) and atractyloside (ATR) have been cited to propose that lys-22 and lys-146 of the bovine polypeptide are involved in, or close to, the binding centre of the nucleotide translocator (36). The binding site for ATR is thought to be closely related to the binding site for adenine nucleotides but different to the site for bongkrekic acid, another inhibitor of nucleotide transport (37). Out of 14 lysine residues which could be labelled in these studies, only 6 are conserved between all three translocator proteins and these include the equivalents of lys-22 and lys-146 in the bovine protein (Fig. 3).

The ATP/ADP translocator is inhibited by N-ethyl maleimide, implicating a cysteine residue at or near the active site (38). A recent publication demonstrated directly that the sulphhydryl group alkylated in the bovine heart mitochondrial carrier is cys-56 (39). This is supported by the sequence comparisons shown in Fig. 3, as cys-56 is the only cysteine residue conserved between all three species.

### ACKNOWLEDGEMENTS

We would like to thank Pioneer Hi-Bred International for providing us with the seed used in these experiments, Mike Douglas (Texas A&M, San Antonio, Texas) for the gift of the yeast translocator clone pYeOPI (2.6), and David Baulcombe of the P.B.I., Cambridge, England for practical advice on cDNA cloning. AB is an S.E.R.C. supported postgraduate student, and this work was supported by an S.E.R.C. research grant to C.J.L.

\*To whom correspondence should be addressed

### REFERENCES

1. Fox, T.D. and Leaver, C.J. (1981) *Cell* 26, 315-323.
2. Isaac, P.G., Jones, V.P. and Leaver, C.J. (1985) In press, *EMBO J.*
3. Dawson, A.J., Jones, V.P. and Leaver, C.J. (1984) *EMBO J.* 3, 2107-2113.
4. Isaac, P.G., Brennicke, A., Dunbar, S.M. and Leaver, C.J. (1985) In preparation.
5. Klingenberg, M. (1979) *Trends Biochem. Sci.* 4, 249-252.
6. Zimmerman, R., Paluch, U., Sprinzl, M. and Neupert, W. (1979) *Eur. J. Biochem.* 99, 247-252.

7. Klingenberg, M. (1985) in 'The Enzymes of Biological Membranes' 2nd ed. Vol. 4. Martonosi, A.N. Ed., pp. 511-553. Plenum Press, New York and London.
8. Groen, A.K., Wanders, R.J.A., Westerhoff, H.V., van der Meer, R. and Tager, J.M. (1982) *J. Biol. Chem.* 257, 2754-2757.
9. Tager, J.M., Wanders, R.J.A., Groen, A.K., Kuntz, W., Bohnensack, R., Küster, U., Letko, G., Bohme, G., Duszynski, J. and Wojtczack, L. (1983) *FEBS Letters* 151, 1-9.
10. Leaver, C.J. and Ingle, J. (1971) *Biochem. J.* 123, 235-244.
11. Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. USA.* 69, 1408-1412.
12. Efstradiatis, A., Kafatos, F.C., Maxam, A.M. and Maniatis, T. (1976) *Cell* 7, 279-289.
13. Buell, G.N., Wickens, M.P., Payvar, F. and Schimke, R.T. (1978) *J. Biol. Chem.* 253, 2471-2482.
14. Gergan, J.P., Stern, R.H. and Wensink, P.C. (1979) *Nucleic Acids Res.* 7, 2116-2136.
15. O'Malley, K., Pratt, P., Robertson, J., Lilly, M. and Douglas, M.G. (1982) *J. Biol. Chem.* 257, 2097-2103.
16. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251.
17. Barnes, W.M. (1977) *Science* 195, 393-394.
18. Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517.
19. Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.* 7, 1513-1523.
20. Messing, J. and Vieira, J. (1982) *Gene* 19, 269-276.
21. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) *J. Mol. Biol.* 143, 161-178.
22. Messing, J. (1983) *Methods in Enzymology* 101, 20-78.
23. Messing, J., Crea, R. and Seeburg, P.H. (1981) *Nucleic Acids Res.* 9, 309-321.
24. Lehrach, H.D., Diamond, D., Wozney, J.M. and Boedtker, H. (1977) *Biochemistry* 16, 4743-4751.
25. Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA.* 77, 5201-5205.
26. Hu, N.-T. and Messing, J. (1982) *Gene* 17, 271-277.
27. Viebrock, A., Perz, A. and Sebald, W. (1982) *EMBO J.* 1, 565-571.
28. Arrends, H. and Sebald, W. (1984) *EMBO J.* 3, 377-382.
29. Aquila, H., Misra, D., Eulitz, M. and Klingenberg, M. (1982) *Hoppe-Seylers Z. Physiol. Chem.* 363, 345-349.
30. Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature* 263, 211-214.
31. Messing, J., Geraghty, D., Hiedecker, G., Hu, N.-T., Kridl, J. and Rubenstein, I. (1983) in 'Genetic Engineering of Plants' Kosuge, T., Meredith, C.P. and Hollaender, A., Eds. pp. 211-227. Plenum, New York.
32. Dennis, E.S., Sachs, M.M., Gerlach, W.L., Finnegan, E.J. and Peacock, W.J. (1985) *Nucleic Acids Res.* 13, 727-744.
33. Bennetzen, J.L. and Hall, B.D. (1982) *J. Biol. Chem.* 257, 3026-3031.
34. Slightom, J.L., Sun, S.M. and Hall, T.C. (1983) *Proc. Natl. Acad. Sci. USA.* 80, 1897-1908.
35. Saraste, M. and Walker, J.E. (1982) *FEBS Letters* 144, 250-254.
36. Bogner, W., Aquila, H. and Klingenberg, M. (1983) in 'Structure and Function of Membrane Proteins' Quagliariello, E. and Palmieri, F. Eds. pp. 145-156, Elsevier Science Publishers B.V., Amsterdam.
37. Block, M.R., Lauquin, G.J.M. and Vignais, P.V. (1979) *FEBS Letters* 104, 425-430.
38. Vignais, P.V. (1976) *Biochim. Biophys. Acta* 456, 1-38.
39. Boulay, F. and Vignais, P.V. (1984) *Biochemistry* 23, 4807-4812.

