MADS-BOX PROTEIN INTERACTIONS IN *ARABIDOPSIS THALIANA*

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

In the Department of Biology

University of Saskatchewan

Saskatoon

Laura Jeanne Voll Reader

January 1998

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree that the libraries of this university may make it freely available for inspection. I further agree that permission for copying of this thesis in manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the head of the department or the dean of the college in which my thesis work as done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Biology

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5E2

ABSTRACT

Arabidopsis thaliana, a member of the Cruciferae family, is widely used as a model system of plant development. In this study the structure - function relationships between *AG* and *AGL2* (***AG****AMOUS* **L**ike) and *AG* and *AGL4* were analyzed for their interaction capabilities at the protein level using the yeast two-hybrid system.

Five *agl2* loss-of-interaction mutants were identified that had lost the ability to interact with AG. All five mutations mapped to the K-box region of the deduced protein sequence of AGL2, and resulted in a substitution of a proline for either a leucine or serine. These mutants had also lost the ability to activate transcription in the two-hybrid system when cloned as *GAL4* **D**NA **B**inding (DB) protein fusions. The screens for ag mutants inactivated for AGL2 interaction recovered uninformative frameshifts, nonsense mutations, and silent mutations within the first one-third of the coding region. The screen for agl4 mutants that had gained the ability to interact with AG yielded phenotypic gain-of-interaction mutants, but no mutation events could be defined.

The position and nature of the agl2 loss-of-interaction mutations suggested that disruption of the coiled-coil structure of the K-box domain was deleterious for interaction between the AGL2 and AG proteins. The results

further suggest that the C terminus of the AGL2 protein contains a domain(s) responsible for the transcription activation properties exhibited by wild type AGL2 expression constructs in the yeast two-hybrid system.

ACKNOWLEDGMENTS

I would like to thank all those who answered my questions and gave me guidance throughout this project. This includes my committee members and supervisor, Dr. W. L. Crosby, but I would also like to thank Susanne Kohalmi for introducing me to the lab and the two-hybrid system and Jacek Nowak for his guidance inside and outside of the lab.

Funding was provided by a Post Graduate Scholarship A from the Natural Sciences and Engineering Research Council of Canada, number 189404.

DEDICATION

I dedicate this thesis to my husband, Sean, who has kept me laughing throughout this project and to my parents, for everything....

TABLE OF CONTENTS

PERMISSION TO USE ii
ABSTRACT iii
ACKNOWLEDGMENTS v
DEDICATION vi
TABLE OF CONTENTS vii
LIST OF FIGURES x
LIST OF TABLES xi
LIST OF APPENDICES xii
LIST OF ABBREVIATIONS xiii

1. GENERAL INTRODUCTION 1

2. LITERATURE REVIEW 3
    2.1 Floral Development 3
    2.2 *AGAMOUS* 5
    2.3 The *AGL* Gene Family 9
    2.4 The MADS-box Domain 9
    2.5 The K-box Domain 10
    2.6 Yeast Two-hybrid System 12
        2.6.1 Strains and vectors 13
        2.6.2 Advantages and disadvantages of the yeast two-hybrid system 19
    2.7 Two-hybrid Variations 21
        2.7.1 'Reverse' two-hybrid system 21
        2.7.2 One-hybrid systems 22

3. METHODS AND MATERIALS 25
    3.1 Generation of Point Mutation Libraries 25
        3.1.1 Heat shock transformation of mutator strain KD1122pCJ 25
        3.1.2 Introduction and assay of point mutations 26
        3.1.3 Mutated insert isolation and ligation into the non-mutant vector 27
        3.1.4 Electroporation into DH10B 27
        3.1.5 Calculations of library size 28
        3.1.6 Storage of library and calculation of aliquot titre 28
        3.1.7 Calculation of percent recombinants in the library 28
        3.1.8 Preparation of library DNA 29

3.2 High Efficiency Lithium Acetate Transformation of Yeast 29
3.2.1 Preparation of carrier DNA 29
3.2.2 Preparation of competent yeast cells 29
3.2.3 Transformation of yeast 31
3.2.4 Calculation of transformation efficiency and number of mutants screened 32
3.3 Colony Characterization 32
3.3.1 Initial characterizations 32
3.3.2 X-gal filter assay 32
3.3.3 Secondary characterization 33
3.4 Isolation of Yeast Plasmid DNA 33
3.5 Electroporation of Yeast DNA Samples 34
3.5.1 *E. coli* competent cells 34
3.5.2 Electroporation 35
3.6 PCR from *E. coli* colonies 35
3.7 Isolation of plasmid DNA from *E. coli* 36
3.8 Primers 37
3.9 Sequencing of DNA samples 37
3.10 Site-directed mutagenesis 37
3.10.1 Primer design 39
3.10.2 PCR 39
3.10.3 Confirmation of mutation by restriction analysis 41
3.10.4 Confirmation of interaction capabilities 41

4. RESULTS 42
4.1 Generation of Non-Directed Mutant Libraries 42
4.1.1 Reversion frequency calculations and the induction of point mutations 42
4.1.2 Library size and recombination frequencies 44
4.2 *Agl4* Mutants 44
4.3 *Agl2* Mutants 46
4.4 Sequence Analysis 46
4.4.1 Site-directed mutagenesis 46
4.4.2 *Agl2* mutants and the self-activation of markers 48
4.5 *Ag* Mutants 48
4.5.1 *Ag* mutant library in the 'prey' vector 48
4.5.2 *Ag* mutant library in the 'bait' vector 49
4.5.3 *Ag* mutant library in mutant 'bait' vector 50

5. DISCUSSION 52
5.1 The MADS-Box Gene Family 52
5.2 Protein-Protein Interaction and DNA Binding Activities Among MADS-Box Proteins 53
5.3 The Two-hybrid System and Point Mutation Analysis 55
5.3.1 *Agl4* mutants 55
5.3.2 *Agl2* mutants 56
5.3.3 Agl2 mutants, the K-box, and loss of transcription activation 57
5.3.4 *Ag* mutants 58
5.3.4.1 Aspects of the agl2Δ2 - AG interaction 58
5.3.4.2 Screening for *ag* mutants 59
5.3.4.3 Screening the AG mutant library in a mutated vector backbone 59
5.3.5 False-positives 60
5.4 Specificity of Function in the AGL Gene Family 61
5.5 Further Studies 64
5.5.1 Confirmation of two-hybrid data *in planta* 64
5.5.2 MADS-box genes as transcription factors 65
5.6 Conclusion 66

6. REFERENCES 67

7. APPENDICES 76

LIST OF FIGURES

2.1. Schematic of Wild Type *A. thaliana* Flower 4
2.2. The ABC Model of *A. thaliana* Floral Development 4
2.3. The ABC Model of *A. thaliana* Floral Development in an *ag* Mutant 6
2.4. Schematic of *ag A. thaliana* Flower. 6
2.5. Alignments of the AG, AGL2, and AGL4 Amino Acid Sequences 8
2.6. The GAL4 Based Yeast Two-Hybrid System 14
2.7a. Yeast Two-Hybrid System Vectors pBI770 and pBI880 17
2.7b. Yeast Two-Hybrid System Vectors pBI771 and pBI881. 18
4.1. Alignment of K-Boxes of Wild Type AGL2 and agl2-Pro Mutants 47

LIST OF TABLES

2.1. Protein Interactions Characterized by the Yeast Two-Hybrid System ... 15
2.2. Yeast, Bacteria, and Vectors used in this Study ... 15
2.3. Conventional versus 'Reverse' Two-Hybrid Selection ... 23
3.1. Primers Used For PCR and Sequencing ... 38
3.2. Primers Used For Site-Directed Mutagenesis ... 40
4.1. Point Mutation Library Reversion Frequencies ... 43
4.2. Point Mutation Library Sizes and Recombination Frequencies ... 45
4.3. Viable Cells and Number of Transformed Cells per Library Screen ... 45
4.4. *Ag* Point Mutations ... 51
7.1. Amino Acid Drop Out Powder for SD Media ... 79
7.2. Yeast Nitrogen Base Without Amino Acids ... 80

LIST OF APPENDICES

7.1 Media 76

LIST OF ABBREVIATIONS

3-AT = 3-amino-1,2-triazole
5-FOA = 5-fluoroorotic acid
aa = amino acid
AG = AGAMOUS*
AGL = AGAMOUS like
amp100 = ampicillin at a concentration of 100 mg/ml
AP1 = APETALA1
AP3 = APETALA3
ARS = autonomous replicating sequence
ARC = ARS consensus sequence
DB = DNA binding
DEF = DEFICIENS
GLO = GLOBOSA
h = hours
min = minutes
MRE = metal response element
OIG = organ identity gene
PCR = polymerase chain reaction
PI = PISTILLATA
s = seconds
SRB = suppressor of RNA polymerase B (polymerase II)
TA = transcriptional activation
TAF = TBP associated factor
TBP = TATA binding protein
U = enzyme units
UAS = upstream activation sequence

*The nomenclature adopted by the *Arabidopsis* community has been used throughout to designate gene names (*italics*) and protein names (standard type) and whether or not the gene or protein is wild type (UPPERCASE LETTERS) or mutant (lowercase letters).

1. GENERAL INTRODUCTION

Arabidopsis thaliana, commonly known as 'Mouse Ear Thale Cress', is a member of the Cruciferae family. The plant is used as a plant model system due to its high seed set, short life cycle, small genome size, small physical size and facile seed mutagenesis. As well, this plant is the object of a global sequencing project that is currently underway. The formation of *A. thaliana* flowers has been studied extensively and is a paradigm for studying the molecular basis of plant development. Three genes were the focus of this project; *AGAMOUS* (*AG*), *AGL2* (***AG****AMOUS* Like), and *AGL4*. The *AG* gene product is expressed in stamens and carpels while the *AGL2* and *AGL4* gene products are expressed early in the floral meristem, and subsequently throughout all four floral organ primordia. Through the use of the yeast two-hybrid system, a protein-protein interaction was identified between AG and AGL2. However AGL4, which is 93% identical to AGL2 at the deduced amino acid (aa) level, displayed a very weak interaction in this *in vivo* system (31). In addition, AGL2 was found to activate two-hybrid marker genes when present as a DNA binding construct, indicating this protein functions as a transcription activator in the yeast two-hybrid system. Despite the similarities in deduced primary structure between AGL2 and AGL4, this study has revealed significant differences in their function.

The aim of this study was to explore the structural basis for the functional differences exhibited by the closely related products of *AGL2* and *AGL4*. Point mutations were introduced into *AGL2*, *AGL4*, and *AG* using non-directed mutagenesis to identify AG-AGL2 loss-of-interaction mutants and AG-AGL4 gain-of-interaction mutants,. In the yeast two-hybrid system ag mutants were assayed for loss of interaction with AGL2, agl2 mutants were assayed for loss of interaction with AG, and agl4 mutants were assayed for gain of interaction with AG. Identified mutations were confirmed by analysis of mutants produced by site-directed mutagenesis. In addition, the ability of *agl2* mutants, present as 'bait' vectors, to independently activate markers in the two-hybrid system was determined. The positions of mutations that resulted in either gain or loss of interaction would lead to information about the domains important for the AG-AGL2 protein-protein interaction.

2. LITERATURE REVIEW

2.1 Floral Development

The wild type *A. thaliana* flower consists of four concentric whorls of organs; four sepals in whorl one, four petals in whorl two, six stamens in the third whorl, and two fused carpels in the fourth whorl (Figure 2.1). Others (5, 20,71) have proposed the ABC model of floral development to explain how the four distinct organs of the flower are formed by the action of **O**rgan **I**dentity **G**enes (OIGs; 18; Figure 2.2). In this model, the expression of class A genes (e.g. *APETALA 1* (*AP1*)) is required for the formation of sepals while the co-expression of class A and B genes (e.g. *APETALA 3* (*AP3*) and *PISTILLATA* (*PI*)) results in petals in whorl 2. In whorl 3, overlapping class B and C (e.g. *AG*) gene expression results in the formation of stamens, while class C gene expression alone results in formation of two fused carpels comprising the pistil with the ovary in whorl 4. Van Tunen isolated MADS-box genes from *Petunia hybrida,* including *FBP11,* that were expressed in the fourth whorl but only in the ovaries (66). In the *fbp11* mutant, carpels were formed in place of the ovules and in ectopic expression experiments, naked ovules were formed on the sepals and petals. The authors therefore proposed that a 'D' function be included in the 'ABCD' model to indicate the formation of the ovules as separate from that of the carpels.

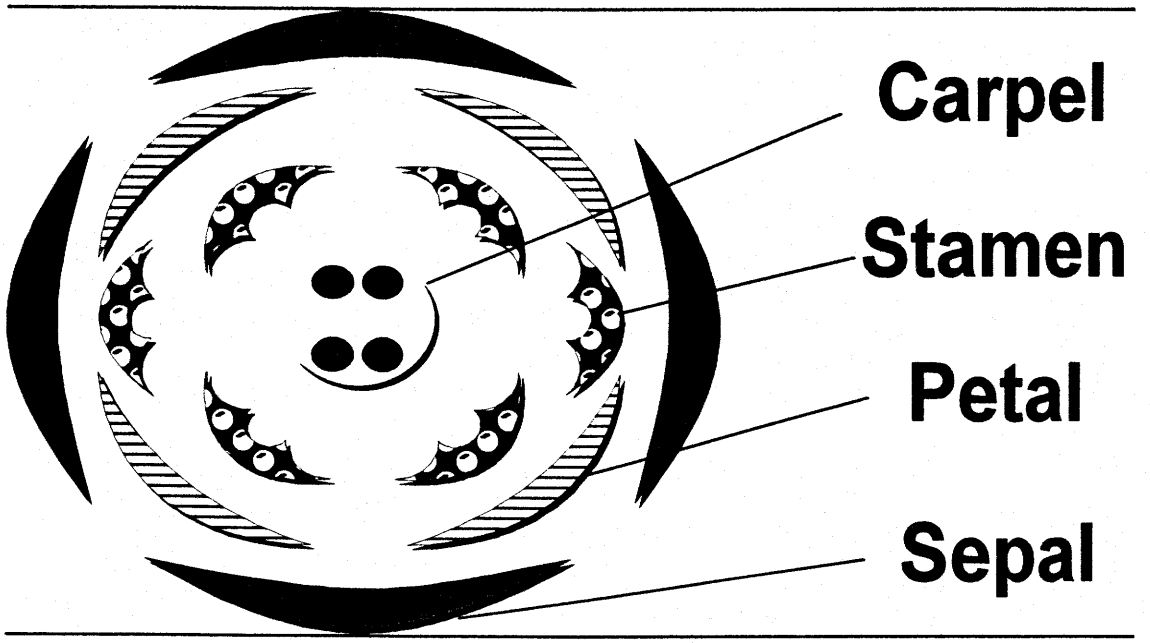


Figure 2.1. Schematic of Wild Type *A. thaliana* Flower
Floral organs are arranged in 4 concentric whorls or rings with the outer-most whorl denoted as whorl 1.

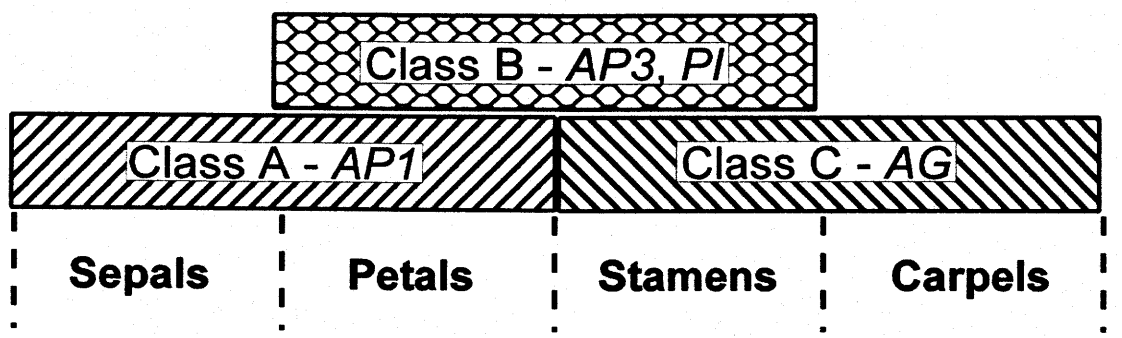


Figure 2.2. The ABC Model of *A. thaliana* Floral Development.
Organ identity genes have been grouped into 3 overlapping expression domains. Expression of class A genes results in sepal formation in whorl 1. Co-expression of A and B genes results in petals in whorl 2. Co-expression of class B and C genes results in stamens in whorl 3 and expression of class C genes in whorl 4 results in carpel formation.

A second group of genes involved in floral development are the cadastral genes, responsible for where in the flower the OIGs are expressed. Class A and C genes, as described above, can be both OIGs and cadastral genes (e.g. *AG*). In a wild type flower, the class A and C genes act to mutually restrict their expression to separate domains. In a class A or C mutant, the expression of the remaining class extends beyond the boundary between the domains and is expressed in all four whorls (see Figure 2.3).

Other genes, such as *SUPERMAN* (4) and *LEUNIG* (34), provide a mostly cadastral function. *LEUNIG* is a negative regulator of the class C gene *AG,* thereby preventing *AG* expression in whorl 1 and 2, but has no apparent organ identity properties (34). Alternatively *SUPERMAN* may also repress class B expression in whorl 4 (4).

2.2 *AGAMOUS*

The *AG* gene of *A. thaliana* is a homeotic class C OIG. Loss-of-function mutations at the *AG* locus result in the formation of six petals in the third whorl instead of stamens, and a new flower begins in the fourth whorl in place of the carpels (Figure 2.4). The *ag* flower has a repetitive structure that may contain up to 70 floral organs, indicating that AG plays a role in floral determinacy as well as organ identity. The deduced AG protein contains two highly conserved domains - the N-terminal MADS-box - an acronym for four sequence-related transcription factors, *MCM1*, *AG*, *DEFICIENSa* (*DEFa*), and *SRF* from yeast,

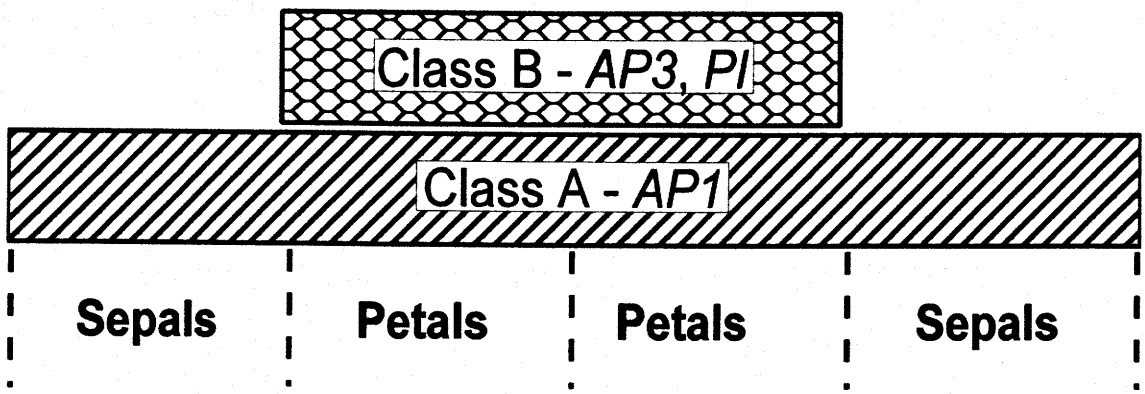


Figure 2.3. The ABC Model of *A. thaliana* Floral Development in an *ag* Mutant.
In a class C mutant, the expression of class A is no longer restricted to whorls 1 and 2. As a result, petals and sepals replace stamens and carpels in whorls 3 and 4.

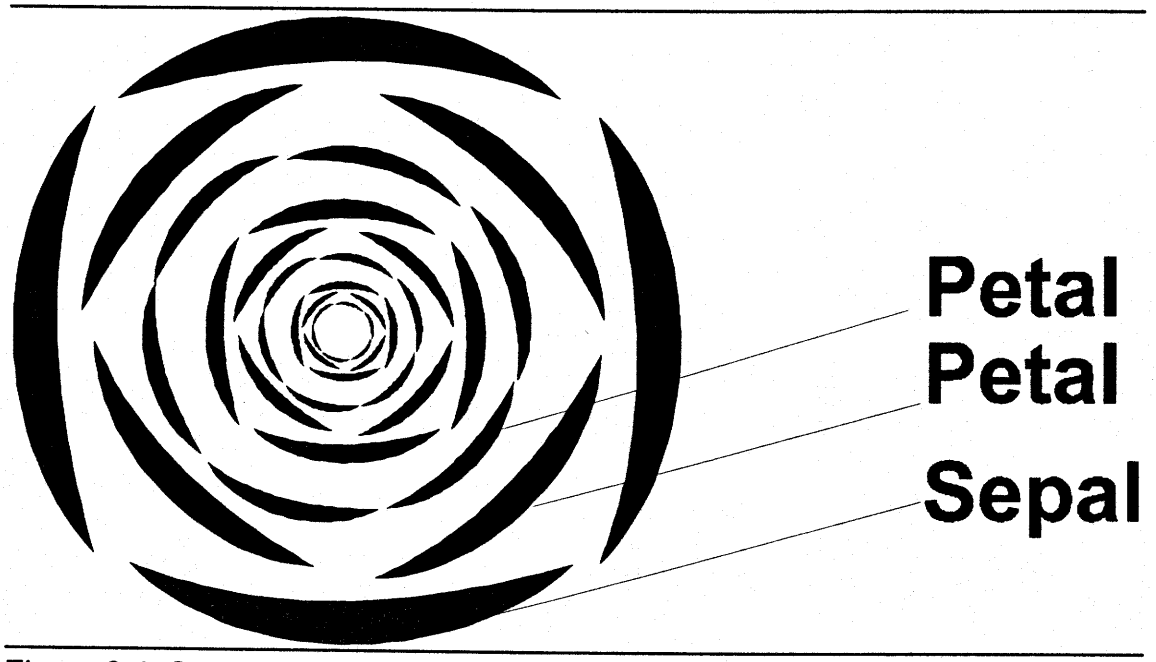


Figure 2.4. Schematic of *ag A. thaliana* Flower.
Note the absence of stamens and carpels and the repetition of the sepal-petal-petal pattern implying that AG plays a role in organ identity and determinancy.

A. thaliana, *Antirrhinum majus*, *Homo sapiens* respectively (59) and a C-terminal motif, the K-box, so named because of its similarity to coiled-coil domains found in keratin (Figure 2.5; 57). MCM1 and SRF are known transcription factors involved in the regulation of mating type-specific gene expression and regulation of the proto-oncogene *c-fos*, respectively (59). The presence of the MADS-box and K-box as well as the regulatory roles of the *AG* and *DEF1* genes in controlling floral organ identity have lead to the suggestion that they too may be transcription factors.

Extensive studies have been done to elucidate the functions attributable to each of the deduced protein domains (N terminus, MADS-box, intervening region, K-box, and C-terminus) of the *AG* gene product. AG deletion constructs used in gel mobility shift experiments determined that both the MADS-box and intervening regions were required for the AG protein to bind with a consensus DNA binding site known as the CarG box (39). However, dimerization of the *AG* gene product did not require the N-terminal one-quarter of the MADS box (40).

Using the yeast two-hybrid system members of the Crosby lab have found that the *AG* gene product interacts with AGL2, AGL9, AGL15, and AGL24, but only very weakly with AGL4 (31). Since AGL2 and AGL4 are 93% identical at the deduced amino acid level, these two gene products present an attractive model system for comparison of MADS-box protein function (37).

Figure 2.5. Alignments of the AG, AGL2, and AGL4 Amino Acid Sequences. Bold letters denote the MADS-Box and underlined letters denote the K-Box

2.3 The *AGL* Gene Family

The first six genes (*AGL1-AGL6*) in the *AGL* gene family were described on the basis of DNA sequence similarity to *AG*, rather than any other properties of expression pattern or function (37). In the subsequent six years some 21 additional members have been identified as structurally similar to *AG*, and all of the genes were found to contain MADS-box and K-box domains (15, 14, 22, 23, 40, 41, 31). Expression patterns revealed that the first 10 members of the family are expressed in flowers as are the floral homeotic genes *AP1*, *AP3*, *CAULIFLOWER (CAL)*, and *PI* (40, 52). *AGL11* and *AGL13* were expressed in ovules, *while AGL12*, *AGL14*, and *AGL17* were expressed in roots and *AGL15* was expressed in the embryo, thus extending the expression domain of selected *AGL* genes into non-floral tissues (47).

2.4 The MADS-box Domain

The MADS-box domain is a 56 aa sequence (Figure 2.5) that has been identified in genes from a number of plant species including rice (9, 47, 22), tomato (49, 43, 51, 44), *Brassica napus* (16), tobacco (42, 24), petunia (68, 2-4, 10, 18, 34, 64, 65), alfalfa (15), sorrel (1), *Asparagus officinalis L.* (39), white campion (14), orchid (35) and maize (44, 12). Putative glycosylation sites (NXT/S) and calmodulin-dependent phosphorylation sites (RXXS/T) were found in all of the *AGL* sequences (36). Since transcription factors such as the yeast heat shock factor, MCM1, and the GAL4 protein are regulated by

phosphorylation (63) and the SRF protein is regulated by glycosylation (63), it may be that plant MADS-box gene products are similarly post-translationally modified since they contain putative glycosylation and phosphorylation sites. In many cases these proteins bind as homo- or hetero- dimers to a consensus DNA recognition sequence known as the CArG box (nucleotide sequence $CC(A/T)_6GG$; 59). The notion that the *AGL* sequences encode transcription factors is based principally on the basis of sequence similarity, and must be further substantiated using alternate criteria based on their biological function.

Unpublished deletion studies from the Crosby lab of the AG-AGL2 protein-protein interaction have shown that the MADS-box is not essential for interactions identified in the two-hybrid system. The interacting cDNA partner may lack the MADS-box motif but an intact K-box domain is retained (31). These results suggested that the K-box plays a major role in mediating protein-protein interaction for AG-AGL2. However, the possibility that the MADS-box domain of AG plays a role in other protein interactions has not been experimentally excluded.

2.5 The K-box Domain

The K-box is a 70 aa coiled-coil motif located C-terminal to the MADS-box and intervening regions and is thought to play a role in protein-protein interactions (Figure 2.5; 46, 9). The domain, reminiscent of keratin, is a heptapeptide repeat structure with the first and fourth positions generally occupied by apolar residues and the remaining positions generally

being polar or charged residues (62). Deletion studies have implicated the K-box in mediating protein-protein interactions between MADS-box genes, and may also play a role in the stabilization of protein-DNA interaction and dimerization (49, 59).

Although no definitive functional role has been assigned to the K-box, it has been demonstrated that K-box deletions resulted in a reduction of the severity of the *ag* phenotype, since transgenic plants carrying this deletion construct exhibited correctly formed stamens and carpels in their inner whorls (40). Studies of the *ag-4* mutation carrying a deletion at the C-terminus of the K-box resulted in flowers with a reduced number of stamens in the third whorl, and internal flowers in the fourth (60). To confirm that K-box disruptions were resulting in these alterations to floral determinacy, a mutant was constructed in which the arginine in position 205 was changed to methionine. When present in an *ag-4* mutant background, this mutant, resulted in increased carpel numbers and additional whorls of stamens and carpels. A careful study of the K-box structure shows that AG-Met205 disrupts the hydrophobic residues located on the face of the coiled-coil K-box and could thereby prevent its interaction with other factors involved in carpel specificity and determinacy (60). Some evidence exists to suggest that the K-box has a role in protein dimerization. Both the *defA-101* allele of *A. majus*, defined by a single aa deletion in the N-terminus of the K-box, and the *ap3-1* allele of *A. thaliana*, with a point mutation in the K-box, result in temperature sensitive homeotic phenotypes, thus implicating the region in protein-protein interaction (21).

2.6 Yeast Two-hybrid System

To better understand the structure - function basis of MADS-box protein-protein interactions, the yeast two-hybrid system was utilized (11) . The system exploits the properties of the *Saccharomyces cerevisiae* GAL4 protein required to express genes that play a role in galactose utilization (45). Two of it's functional domains are of particular interest - the N-terminal **DNA B**inding domain (DB; residues 1-147) and the C-terminal **T**ranscription **A**ctivation domain (TA; residues 768-881; 46). Transcription activation is achieved by the binding of the DB domain to the **U**pstream **A**ctivation **S**equences (UAS) of the *GAL* genes and interaction with the RNA II polymerase complex bound TA domain. In the yeast two-hybrid system the GAL4 DB and TA domains are expressed as fusion proteins from separate plasmid expression constructs. In order to determine if two proteins interact, the corresponding genes are separately cloned and expressed as independent GAL4 DB- and TA- fusion proteins. If the proteins interact *in vivo* in the two-hybrid system, the DB- and TA- domains of the fusion proteins form a protein complex that mimics the properties of GAL4. In the yeast two-hybrid system, specialized two-hybrid host strains have been developed in which marker genes such as *lacZ* and *HIS3* have been placed under *GAL4* - UAS regulation. The activation of these two marker genes by a protein-protein interaction permits either genetic selection or biochemical marker assessment. The *HIS3* gene product confers resistance to 3-amino-1,2-triazole (3-AT) and allows those cells to grow only if they contain a protein-protein

interaction resulting in *HIS3* transcription. These putative interactors are then assayed for the second marker gene activity in which the colonies turn blue in the presence of X-gal when the *lacZ* marker is activated (Figure 2.6). The genetic selection (resistance to 3-AT) afforded by *HIS3* activation greatly reduces the number of colonies that have to be further tested for activation of the biochemical marker. Both of the markers incorporate different promoter contexts of the *GAL4 - UAS*, thus minimizing the identification of promoter-specific activating proteins. This system has been utilized to identify a large and growing list of biologically relevant *bona fide* protein-protein interactions (Table 2.1)

2.6.1 Strains and vectors

In this study, two-hybrid interactions were assessed in the yeast host strain YPB2 (Table 2.2). This strain is deleted for the *GAL4* and *GAL80* genes, which negatively regulate *GAL4* activity, and contains *lacZ* and *HIS3* marker genes under the control of the *GAL4 - UAS* sequences. The DB and TA fusion expression vectors used were derived from pPC62 and pPC86 respectively, with pBI880 and pBI770 containing the DB domain and pBI881 and pBI771 containing the TA domain (Figure 2.7a, b; 30). The pBI88x vectors included an additional nucleotide that resulted in a different reading frame than the pBI77x vectors. All inserted genes were cloned as *SalI- NotI* fragments into either vector set allowing for easy exchange and subcloning.

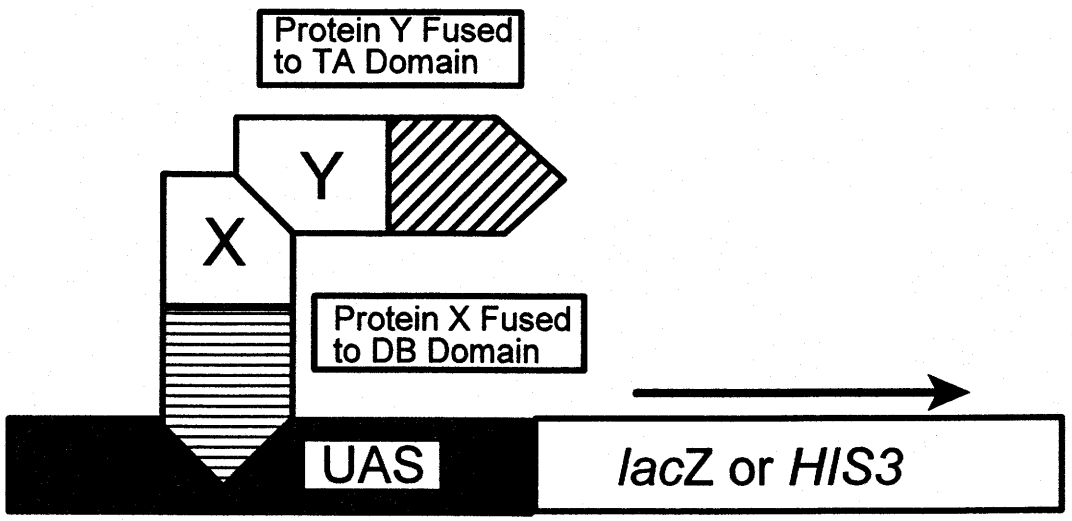


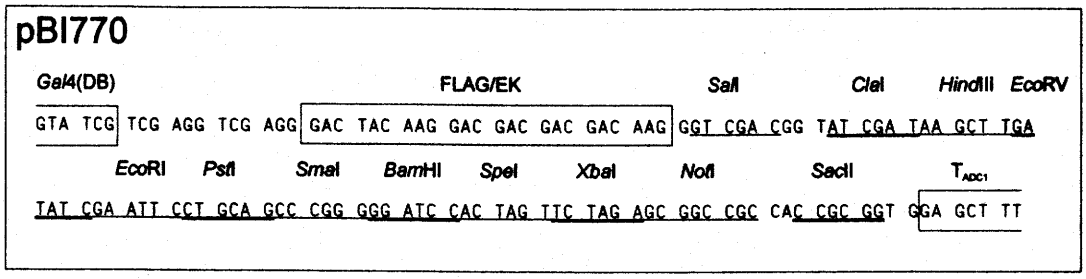
Figure 2.6. The GAL4 Based Yeast Two-Hybrid System.
Interaction between proteins X and Y is assayed for by constructing the two protein fusions, DB-X and TA-Y. The DB domain interacts with the UAS and the TA domain recruits RNA polymerase II. When a link is formed between DB and TA by the X-Y interaction, transcription of the marker genes occurs. Colonies that are capable of growth on 3-AT and turn blue in an X-gal assay are deemed positive for protein-protein interaction.

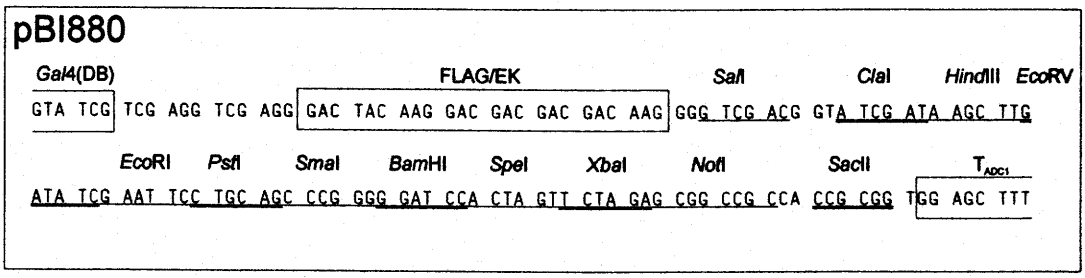
Table 2.1. Protein Interactions Characterized by the Yeast Two-Hybrid System

Organism	Protein Expressed by 'bait' Vector	Protein Expressed by 'prey' Vector	Reference
S. cerevisiae	SNF1	SNF4	(11)
Rat	Rab3a	rat brain cDNA library	(5)
Human	RhoA	Rho-GAP	(2)
Human	Ras	CDC25-GEF	(41)
Mouse	Ras	Raf	(68)
Human	Ras	NF1	(45)
E. coli	FliM	FliG, FliM, FliN	(38)
S. cerevisiae	PF1K	PF2K	(50)
HPV-16	E7	E7	(8)
BMV	BMV 1a	BMV 2a	(42)
TVMV	NIb	NIa and CP	(18)
Human	BTB/POZ	LAZ3/BCL6	(10)
HIV-1	NEF	CD4	(51)
Human	SHPTP2	Grb7	(23)
Human	dystrophin	syntrophin	(6)
Human	ER	ER	(70)

Table 2.2. Yeast, Bacteria, and Vectors used in this Study

Name	Description	Genotype	Reference
YPB2	*S. cerevisiae*	*MATa ura3-52 his3-200 ade2-101 lys2-801 trp1-901 leu2-3,112 can*^{R} *GAL4-542 gal80-538 LYS2::GAL1*$_{UAS}$*-LEU2-*$_{TATA}$*-HIS3* URA3::GAL4$_{17mers(x3)}$ *Cyc1*$_{TATA}$*-lacZ*	(3)
MaV95	*S. cerevisiae*	*MATa, ura3-52, leu2-3, 112 trp1-901, his3Δ200, ade2-101, gal4Δ, gal80Δ, can1*^{R}, *cyh2*^{R}, *GAL1::HIS3@LYS2, GAL1::lacZ*	(54)
KD1122pCJ	*E. coli*	*trp*A *leu arg*E *his thi thr spc*A *mut*D5 F'pCJ105(Cam^{R})	(17)
DH5α	*E. coli*	F^{-} ϕ80d*lacZ*ΔM15 Δ(*lac*ZYA-*arg*F) U169 *deo*R *rec*A1 *end*A1*hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 λ^{-} *th*i-1 *gyr*A96 *rel*A1	Gibco-BRL Burlington, ON
DH10B	*E. coli*	F^{-} *mrc*A Δ(*mrr-hsd*RMS-*mrc*BC) ϕ80d*lacZ*ΔM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*D139 Δ(*ara,leu*)7697 *gal*U *gal*K λ^{-} *rps*L *nup*G	Gibco-BRL Burlington, ON
Epicurian Coli® XL-2 Blue	*E. coli*	*recA1 endA1 gyrA96 thi-1 hdsR17 supE44 relA1 lac* [F' *pro*AB *lacI*^{q}Z*ΔM15* Tn*10* (Tet^{r}) Amy Cam^{r}]	Stratagene Aurora, ON





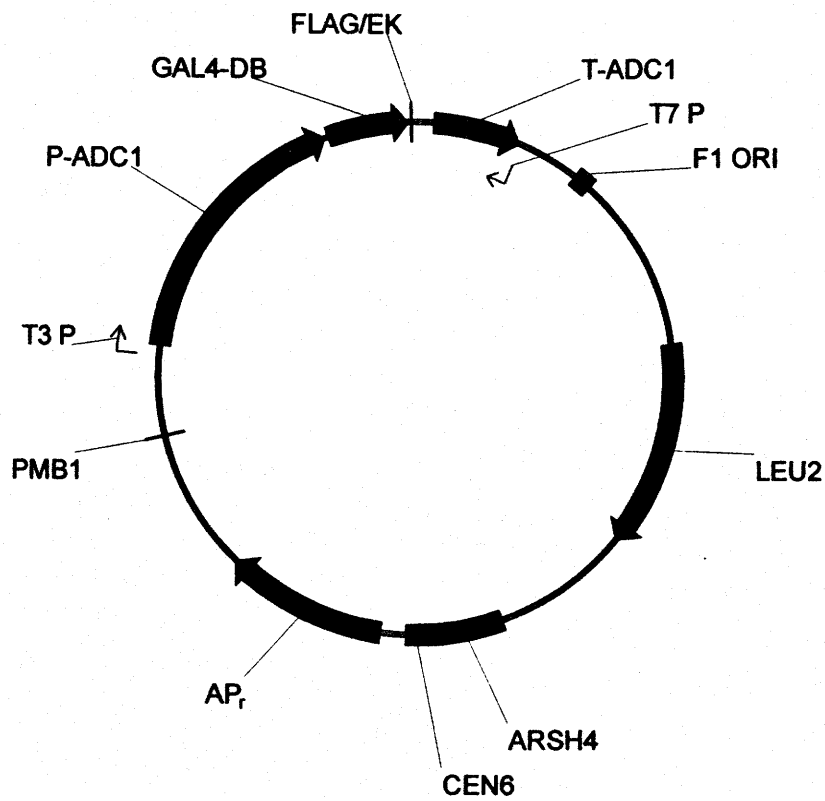
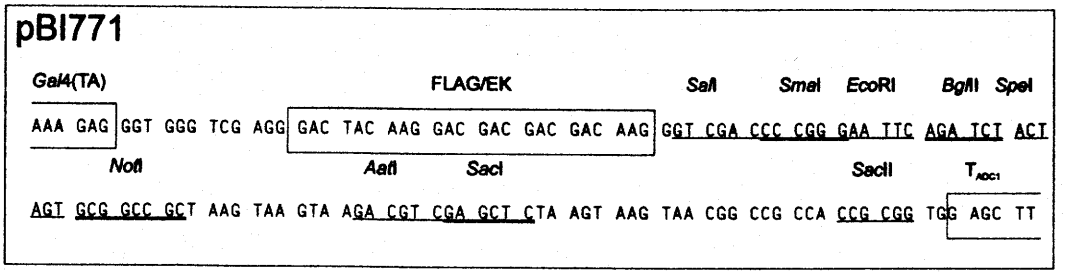
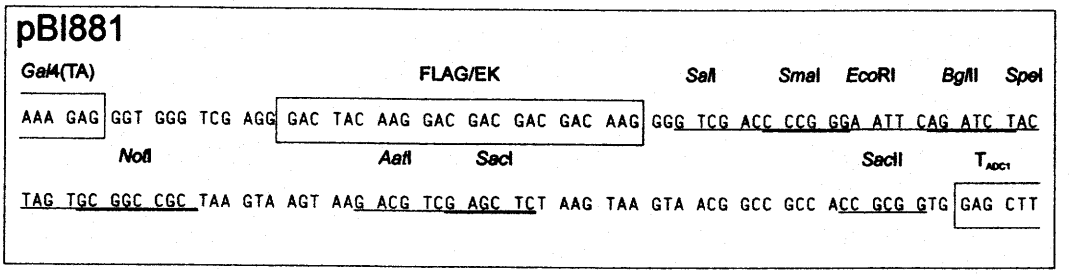


Figure 2.7a. Yeast Two-Hybrid System Vectors pBI770 and pBI880. pBI880 contains an extra G following the FLAG/EK site to generate a +1 reading frame. All inserts were clones as *SalI-NotI* fragments.





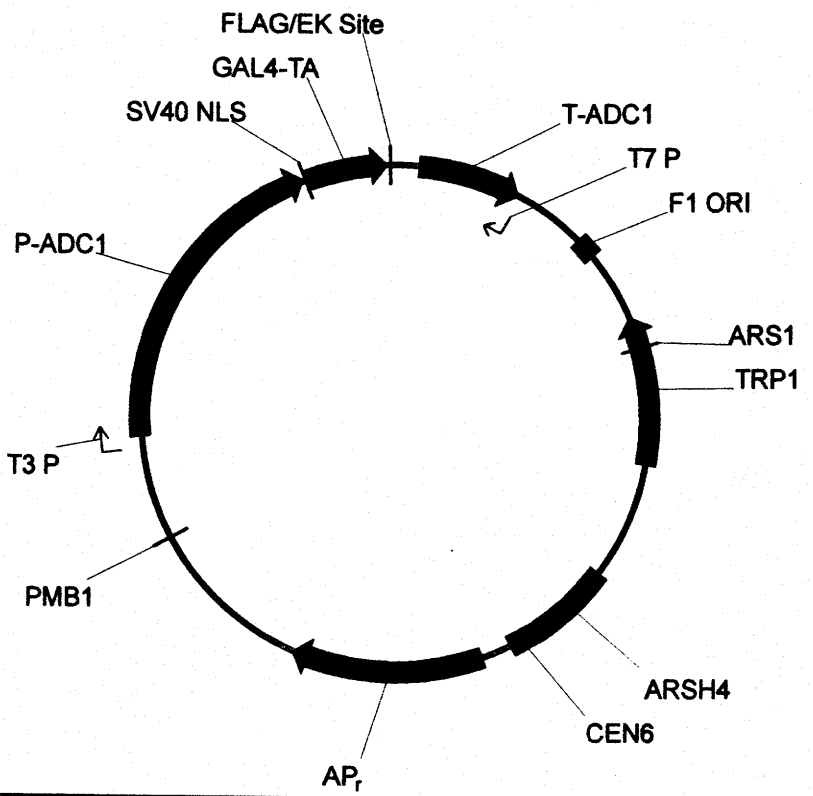


Figure 2.7b. Yeast Two-Hybrid System Vectors pBI771 and pBI881. pBI881 contains an extra G following the FLAG/EK site to generate a +1 reading frame. All inserts were clones as *SalI-NotI* fragments.

2.6.2 Advantages and disadvantages of the yeast two-hybrid system

The yeast two-hybrid system is a powerful tool to identify novel protein-protein interactions through library screens, critical aa through point mutation studies, and minimal domains required for interaction through deletion analysis. The system can detect interactions of nuclear, cytoplasmic, mitochondrial, membrane associated and extracellular proteins (43). An advantage of the system is its ability to screen a large number of colonies and select, through genetic analysis, only a relatively small number for further analysis. The manipulations of this system are all performed at the DNA level forestalling the problems associated with protein purifications. Novel protein-protein interactions can be confirmed using complementary approaches such as affinity chromatography or co-immunoprecipitation (43).

The two-hybrid system is not without its disadvantages. For a protein-protein interaction to be detected by the yeast two-hybrid, the proteins must be localized to the nucleus. In addition, proteins that are post-translationally modified in their natural hosts may not form the proper conformation in the yeast cell. The two-hybrid system makes use of protein fusions that may exclude certain domains from being presented for interaction and some proteins, like AGL2, activate transcription of marker genes when presented as 'bait' fusions and must be modified prior to any assays for

protein-protein interaction. In light of the advantages, and despite the disadvantages presented here, the yeast two-hybrid system is a powerful approach to identify and study protein interaction.

Under sub-optimal conditions the yeast two-hybrid system can deliver an unacceptably high frequency of false positives. These false positives generally appear as colonies that grow on selection plates (i.e. SD-leu-trp-his+3-AT) but fail to show co-activation of the *lac*Z marker gene. Others have shown that careful optimization of 3-AT concentration and plating density greatly reduced the number of false positives encountered (26). The system is hindered by the fact that some 'bait' fusion proteins activate both marker genes in the absence of an interacting partner in a TA vector. The reverse situation has also been encountered where a TA fusion construct activated transcription of the marker genes in the absence of an interacting DB plasmid (75), binding directly or indirectly to the UAS sequences present on the yeast chromosome. To minimize artifacts, each 'bait' plasmid, prior to screening, must be assayed for its ability to independently activate marker genes. Similarly each putative interacting TA plasmid should be tested in a series of controls including the absence of a DB construct, DB vector with no inserted gene, and a DB with a known non-interacting gene insert. In these latter constructs, the seed storage protein, *CRUCIFERIN* was used as a biological control, since it has no known role in

flower development. Together, these controls eliminate the majority of cDNAs that activate marker genes independent of a *bona fide* protein-protein interaction.

It is important to note that the two-hybrid system is designed to detect binary protein-protein interactions. Thus systems, such as the three-hybrid system(74, 58), that detect protein interactions requiring higher quaternary complexes can be used to characterize multi-component protein complexes.

2.7 Two-hybrid Variations

2.7.1 'Reverse' two-hybrid system

The yeast 'reverse' two-hybrid system was designed to select for *loss* of interaction between two proteins (70, 71, 32). The system uses the same basic plasmid set as used in this study for conventional two-hybrid screening but involves the use of specialized yeast strains, such as MaV95 (Table 2.1), that contains a *Spo13* promoter-*URA3* marker construct under the control of *GAL4* UAS (*SPAL5::URA*; 66, 67) The *Spo* promoter is a sporulation specific promoter that is completely inactive during cell growth and division of the haploid, with the *URA3* gene only active in response to *GAL4* activation. In those cells where the protein interaction persists, *URA3* is activated and the host cell are killed in the presence of 5-fluoroorotic acid (5-FOA). Therefore only those colonies in which there is no protein -protein interaction are recovered (Table 2.3).These colonies can be further tested for 3-AT and *lacZ* phenotypes. In this system, the gene for one member of an interacting protein pair is mutagenized, using either non-

directed or directed approaches, and used to screen for loss-of-interaction mutants with the wild type interacting partner. cDNA's resulting from the loss of interaction can then be sequenced to determine the mutation that resulted in the loss in interaction capability.

An alternative to the 'reverse' two-hybrid system used here has been developed that uses the cycloheximide resistant yeast strain CL9 (a derivative of JC981 - α *gal4 gal80 his3 trp1-901 ura3-52 URA3::GAl1→lacZ leu2-3, 112 cyh2*) where the *CHY2* construct was placed under the control of *GAL4* UAS (32) in place of *URA3* (31). In these host cells, the interaction of the DB-TA fusion protein complex activates transcription of *CYH2*, resulting in sensitivity to cycloheximide thus allowing for the selection of loss of interaction in the presence of cycloheximide.

2.7.2 One-hybrid systems

The one-hybrid system was first developed to isolate and characterize *ORC6*, a component of the *Saccharomyces cerevisiae* origin of replication complex (33). This system is similar to the conventional two-hybrid system in that cDNA library clones are expressed as fusions to the TA domain of GAL4. Using standard transformation protocols, the library can be introduced into a host strain carrying a reporter gene in *cis* with nucleotide binding sites for the protein of interest, in this case the **A**utonomous **R**eplicating **S**equence (ARS)**C**onsensus Sequence (ARC) from yeast. In these experiments reporter

Table 2.3. Conventional versus 'Reverse' Two-Hybrid Selection

Protein Interaction		No Protein Interaction	
Conventional System	'Reverse' System	Conventional System	'Reverse' System
3-AT resistant	3-AT resistant	3-AT sensitive	3-AT sensitive
	5-FOA sensitive		5-FOA resistant

genes are activated when one of the protein fusions from the TA library is bound to the ARC. This mimics the DB domain of the two-hybrid system resulting in transcriptional activation by the fused TA domain. Others have developed a similar system using the *HIS3* selectable marker (73,19) . In this system a dual selection scheme was utilized in their search for a **M**etal **R**esponse **E**lement (MRE) binding protein where a second reporter construct was incorporated carrying the *HIS3* reporter gene. These constructs contained either *TRP1* or *URA3* markers for selection in yeast, 4 MRE sequences, a promoter, and a selectable marker. Interaction of the TA fusion library with the target DNA domain activated marker gene expression, resulting in 3-AT resistance (in the case of *HIS3* activation) as well as *lacZ* expression. This system was used to show that the BZLF1 protein of Epstein-Barr virus was capable of binding the igH intronic enhancer (13) and to characterize the *GATA-1* DNA binding domain through mutational analysis (73).

3. METHODS AND MATERIALS

3.1 Generation of Point Mutation Libraries

3.1.1 Heat shock transformation of mutator strain KD1122pCJ

(Modified from Sambrook *et. al.*;(53)

When using strain KD1122pCJ, mutator (2YT media) and non-mutator (M9 media) growth conditions need to be carefully determined. The strain should only be grown under mutator conditions when it carries the construct to be mutated or when testing for induction of the mutator effect. Non-mutant growth conditions are used when preparing competent cells and during transformation to minimize the accumulation of point mutations in the *E. coli* genome. To do this a 3 ml 24 h pre-culture of *E. coli* strain KD1122pCJ was grown in M9 media (non-mutator conditions) and 100μL was used to inoculate 100 ml of M9 media and grown to an OD_{600}=0.5-0.8 at 37°C. The cells were centrifuged at 5000 rpm in a Saroval GSA rotor at 4°C for 15 min, the supernatant discarded, and the cells resuspended in 10 ml of ice cold 50 mM $CaCl_2$ in an Oakridge tube. The cells were pelleted and resuspended in 2 ml of ice cold $CaCl_2$ while the supernatant was discarded. The centrifugation was repeated and cells resuspended in 0.5 ml of 50 mM $CaCl_2$. Aliquots of cells (50 μl) were combined with 0.5 μg of plasmid DNA (pBI881-*AGL2*, pBI881-*AGL4*, or pBI880-*AG*). The mixtures were incubated on ice for 30 min,

heat shocked at 42°C for 45 s, and returned to ice for 2 min. 2YT (280 μl) was added and cells were transferred to a 15 ml culture tube where they were incubated with vigorous shaking at 37°C for 30 min before 0.2 ml was plated onto M9+amp$_{100}$ media.

3.1.2 Introduction and assay of point mutations

A 3 ml pre-culture of M9+amp$_{100}$ media was inoculated with a single colony of KD1122pCJ carrying the plasmid construct to be mutated. 2YT+amp$_{100}$ (mutator conditions; 100 ml) and 3 ml of M9+amp$_{100}$ (non-mutator conditions) were inoculated with 100 μl of the pre-culture and incubated overnight at 37°C with shaking. In order to measure the mutator effect, cultures in 2YT or M9 media were assessed for the reversion frequency of a chromosomal histidine auxotrophic marker point mutation. From each culture, 3 ml of cells were centrifuged in a desktop centrifuge and resuspended in 3 ml of M9-his+amp$_{100}$ media. Serial dilutions in M9-his+amp$_{100}$ media were prepared, plated onto M9+amp$_{100}$ and M9-his+amp$_{100}$ media and incubated at 37°C for 2-3 days. Viable cell titres as CFU were calculated for each of the four plating conditions. Reversion frequencies for the *his*$^-$ allele, determined under mutator and non-mutator conditions were defined as the CFU on M9-his+amp$_{100}$ divided by the CFU on M9+amp$_{100}$. The reversion frequency was routinely 10-100 fold higher under mutator culture conditions. Mutagenized plasmid DNA was isolated from the remainder of the mutated 100 ml 2YT+amp$_{100}$ culture using a Maxi Prep system (Qiagen).

3.1.3 Mutated insert isolation and ligation into the non-mutant vector

To ensure changes in interaction capabilities in the two-hybrid system were due to the mutations in the recombinant donor cDNA and not elsewhere in the construct, insert sequences were subcloned into the non-mutant vector prior to screening. Mutant coding region constructs were digested with *Sal*I and *Not*I in a reaction that contained 10 μg DNA, 4 μl of Gibco BRL buffer 3, 10 U of *Sal*I, 10 U of *Not*I, and made up to 40 μl with sterile distilled water. The restriction digests were incubated at 37°C for 2 hours prior to electrophoresis through 0.8% preparative agarose gels. DNA bands representing the excised insert were isolated from the agarose using a commercial kit (GeneClean® II, Bio 101 Inc. Bio/Can Scientific). Aliquots of the insert (20 μl) were ligated with 20 μl of wild type vector prepared in an identical manner with 40 μl of 5X ligation buffer (Gibco BRL), 120 Weiss Units of T4 ligase (Gibco BRL) made up to 200 μl with sterile distilled water and incubated overnight at 16°C. The ligation reactions were precipitated with the addition of 100 μl of 7.5 M ammonium acetate plus 600 μl of 95% ethanol at -85°C for 20 minutes. The tubes were centrifuged in a microcentrifuge for 20 min at 4°C and the pellet washed, dried, and resuspended in 20 μl of sterile distilled water.

3.1.4 Electroporation into DH10B

Ten 2 μl aliquots of the ligation mixtures were electroporated into 25 μl aliquots of 'Electromax®' DH10B Competent Cells (Gibco BRL) using a BioRad Electroporator (1500 V, 50 W, 50 mA, 250 Ω , 40 μF to max 2500 DC) and cells

were allowed to recover for 60 min in 9.5 ml 2YT + 1.0% Glucose. Cells were centrifuged for 10 min and resuspended in 1 ml of 2YT while the supernatant was discarded. Approximately 20 μl of the cells were reserved for titering and the remaining 980 μl were plated on a 20 cm square petri dish containing 250 ml of 2YT+amp$_{100}$ solid media and incubated at 37°C overnight.

3.1.5 Calculations of library size

The 20 μl of cells that were reserved in the previous step were used to make serial dilutions prior to plating onto 2YT+amp$_{100}$ media. Plates were incubated at 37°C overnight and the viable cell count of these titre plates was calculated to determine the number of independent clones in the library.

3.1.6 Storage of library and calculation of aliquot titre

The library colonies were scraped from the plate using a glass rod and resuspended in 10 ml of 2YT +15% glycerol. Aliquots (0.5 ml) were quick-frozen and stored at -85°C. A 5 μl aliquot was used to make serial dilutions, plated onto 2YT+amp$_{100}$ media, and incubated overnight at 37°C. The titre was used to calculate the ampicillin resistant viable cell counts of the frozen stocks.

3.1.7 Calculation of percent recombinants in the library

Twenty colonies from the titre plates were individually inoculated into 3 ml 2YT+amp$_{100}$ media and grown overnight at 37°C with shaking . Plasmid DNA was isolated, restricted with *Sal*I and *Not*I, and electrophoresed through a 0.8 % agarose gel to determine the number of colonies containing plasmids with DNA inserts.

3.1.8 Preparation of library DNA

1 L of 2YT+amp$_{100}$ media was inoculated with 1 ml of frozen *E. coli* cells containing the library DNA and incubated at 37°C with shaking overnight. Maxi Preps™ (Qiagen) were used to isolate high purity DNA which was subsequently precipitated with 2 volumes of ethanol and 0.1 volume of 3 M sodium acetate. Samples were placed at -20°C overnight and centrifuged for 20 min at 4°C. The supernatant was discarded, and the DNA stored in a dried state until resuspended to 1 μg/μl with sterile distilled water.

3.2 High Efficiency Lithium Acetate Transformation of Yeast

3.2.1 Preparation of carrier DNA

(Modified from Sambrook *et. al.*; 53).

100 mg of salmon-sperm DNA was dissolved in 10 ml of TE buffer (pH 8.0) and incubated overnight at 4°C with stirring. The DNA was sonicated twice for 30 s with a large probe at ¾ power (OMNI International GLH). A sample of the DNA was electrophoresed through an agarose gel to determine that the average size of the fragments was 7 Kb and that they ranged in size from 2-15 Kb. The solution was aliquoted out in 0.5 ml portions and stored at -20°C.

3.2.2 Preparation of competent yeast cells

(Modified from Schiestl and Gietz; 55)

A 3 ml pre-culture of SD media (SD-leu for 'bait' vectors and SD-trp for 'prey' vectors) was inoculated with 10 μl of stock culture, previously frozen at

-80°C, and incubated at 30°C with shaking overnight. Cell number in the resulting preculture was determined with a hemacytometer and used to calculate the inoculum of a larger culture used for the preparation of competent yeast cells. A desired culture cell density of 7×10^6 cells/ml was selected to ensure that cells would be in log phase upon harvest. In Equation 3.1, the calculation is based on: a preculture cell density of 6.1×10^7 cells/ml, a generation time of 1.8 h, a final culture volume of 500 ml, and a 13.5 h growth period.

$$\text{Inoculation Volume} = \frac{\text{Current Cell Density}}{\text{Desired Cell Density}} \times \frac{\text{Culture Volume}}{2^{\text{No. Generations}}} \qquad (3.1)$$

$$= \frac{7 \times 10^6 \text{ cells / ml}}{6.1 \times 10^7 \text{ cells / ml}} \times \frac{500 \text{ ml}}{2^{7.5}}$$

$$= 317\ \mu\text{L}$$

Log phase yeast cells were harvested by centrifugation for 5 min at 5000 rpm in a GSA rotor at room temperature. The supernatant was discarded and the cells resuspended in 1/30 their original volume in sterile distilled water, transferred to 40 ml Oakridge tubes, and centrifuged for 5 min at 7000 rpm in an SS34 rotor. Again the supernatant was discarded and the cells were resuspended in 1/200 of their original volume in lithium acetate solution (0.1 M lithium acetate in 1xTE buffer, pH 7.5) followed by incubation for 1 hr at 30°C with shaking before use in transformation experiments, or frozen in 15% glycerol at -85°C.

3.2.3 Transformation of yeast

(Modified from Schiestl and Gietz; 55)

Competent cells (200 μl) were combined with 20 μl of carrier DNA and up to 5 μg of DNA was added to a microfuge tube followed by an incubation for 30 min with shaking at 30°C. PEG solution (1.2 ml of 40% PEG 4000, 0.1 M lithium acetate, 0.01 M TE buffer, pH 7.5) was added, the tubes mixed by inversion and again incubated for 30 min at 30°C. Cells were placed in a 42°C water bath for 15 min and cooled on ice for approximately 3 min. Tubes were centrifuged for 5 s, the supernatant removed by aspiration and cells resuspended in 0.5 ml of TE buffer (pH 8.0) using the wide end of a flat toothpick. Cells were washed twice in 0.5 ml of TE buffer (pH 8.0) and resuspended to a final volume of 1.0 ml TE buffer (pH 8.0). Aliquots of the resuspended cells (0.2 ml) were plated onto appropriately supplemented SD plates (SD-leu-trp-his+5 mM 3-AT for forward selection, or SD-leu-trp+0.2%(w/v)5-FOA for reverse selection). Plates were incubated at 30°C until colonies appeared (three to seven days). For library screens, all volumes were doubled and the transformations were done in a set of five 2.0 ml microcentrifuge tubes using 10 μg of library DNA per tube. Since the tubes were not large enough to add a doubled volume of PEG solution, they were filled to the top (approximately 1.6 ml) with no deleterious effect on transformation efficiencies. For the large tubes, transformations were resuspended to a final volume of 1.0 ml TE, allowing for the application of four 0.25 ml aliquots to individual plates.

3.2.4 Calculation of transformation efficiency and number of mutants screened

5 μl of the transformation mix was serially diluted and plated on appropriately supplemented SD media at 30°C to determine the titer of viable and transformed cells. Colony counts from SD-leu-trp media were used to calculate the number of transformants that were screened, while SD-leu or SD-trp plate counts were used to calculate the number of viable CFUs plated in the experiment. The ratio of transformants to viable cells was used to calculate the transformation frequency.

3.3 Colony Characterization

3.3.1 Initial characterizations

Colonies were picked using the wide end of a flat toothpick and streaked on two SD-leu-trp plates and incubated at 30°C. Cells scraped from one plate were used to prepare 1 ml 50% glycerol stocks, while the second plate was used for X-gal filter assays.

3.3.2 X-gal filter assay

(Modified from Chevray and Nathans; 7)

In the lid of a petri dish, 1.8 ml of Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$), 5 μl of 2-mercaptoethanol, and 20 μl of X-gal (100 mg/ml in dimethylformamide) were mixed and used to soak a trimmed Whatman No. 1 filter paper. A nylon-backed nitrocellulose filter (MSI NitroPlus) was placed over the colonies and allowed to wet completely. The filter

was removed, placed cell side down in liquid nitrogen for approximately 10 s., and then placed cell side up on the reagent - soaked filter paper. The plates were sealed with Parafilm® and placed in a 30°C incubator. Results were checked at hourly intervals, and again after approximately 16-24 hours.

3.3.3 Secondary characterization

Frozen aliquots (5 μl) of β-galactosidase expressing colonies were spotted to SD-leu-trp and SD-leu-trp-his+3-AT (SD-leu-trp+5-FOA media in the case of 'reverse' two-hybrid experiments) to recheck *lacZ* activation and growth patterns (3-AT^R for the conventional two-hybrid system, 3-AT^S and 5-FOA^R for the 'reverse' two-hybrid system). For those colonies showing appropriate marker gene activation, DNA was isolated using a protocol modified from Robin Wright (Dept. Zoology, University of Washington, Seattle) as outlined in section 3.4.

3.4 Isolation of Yeast Plasmid DNA

A 5 ml yeast pre-culture was grown to stationary phase in the appropriate media. Cells were centrifuged at 3500 rpm for 10 min in a bench top centrifuge, the supernatant was discarded, cells were resuspended in 1 ml of Zymolyase solution (2 mg/ml Zymolyase 20T (ICN Biomedicals Inc.) or Yeast Lytic Enzyme (ICN Biomedicals Inc.), 100 mM KPO_4, pH 7.5, 1.2 M Sorbitol) and incubated at 37°C with aeration for 2 hours. The resulting spheroplasts were centrifuged for 5 s in a microcentrifuge, the supernatant was discarded and the pellet resuspended in 200 μl of cell lysis solution (0.2 M NaOH, 1% (w/v) SDS). The solution was incubated at 65°C for 5 min and then 200 μl of neutralization

solution (2.55 M potassium acetate, pH 4.8) was added. Samples were mixed by gentle inversion, placed on ice for 15 min., and then centrifuged in a microcentrifuge for 20 min. The supernatant was transferred to a new microcentrifuge tube and the DNA was precipitated by the addition of 1 ml of 95% ethanol. Samples were centrifuged in a microcentrifuge for 2 min, the supernatant aspirated off, and the pellet resuspended in 25 μl of sterile distilled water. *E. coli* strain DH10B was transformed with 2 μl of yeast DNA by electroporation.

3.5 Electroporation of Yeast DNA Samples

3.5.1 *E. coli* competent cells

(Modified from Sambrook *et. al.;* 53).

A 3 ml pre-culture of 2YT media was inoculated with 10 μl of DH10B stock and grown overnight at 37°C. An aliquot (500 μl) of the pre-culture was added to 500 ml of 2YT and cultured at 37°C until the cell density reached OD_{600}=0.5-0.7. Cells were placed on ice for 30 min, centrifuged at 2400 rpm for 10 minutes in a GSA rotor at 4°C, and the supernatant was discarded. Cells were resuspended in 500 ml of ice cold sterile water, centrifuged at 2800 rpm for 10 min in a GSA rotor at 4°C, and the supernatant discarded. Cells were resuspended in 250 ml of ice cold sterile water, centrifuged at 3000 rpm for 10 min in a GSA rotor at 4°C, and the supernatant was discarded. This final resuspension, centrifugation and decantation was repeated and the cells were

resuspended in a final volume of 2.5 ml in water. Cells were left on ice for 1 to 3 hours to improve electroporation efficiencies. Any cells not required immediately were frozen at -85°C in 15% glycerol and then thawed on ice when required.

3.5.2 Electroporation

Aliquots (2 μl) of yeast DNA were added to 0.2 cm gap electroporation cuvette (BioRad) and placed on ice. 'Homemade competent cells' (100 μl, see section 3.2.2) or 25 μl of commercially prepared DH10B cells (Gibco BRL 'Electromax®') were added, and the samples submitted to a 1500 V pulse using a BioRad Electroporator (1500 V, 50 W, 50 mA, 250 Ω , 40 μF to max 2500 DC). Cells were allowed to recover in 1 ml of 2YT for 30 min before plating on 2YT+amp$_{100}$ plates and incubating at 37°C.

3.6 PCR from *E. coli* colonies

In order to ensure that the desired 'bait' or 'prey' vector was present in selected *E. coli* colonies a **P**olymerase **C**hain **R**eaction (PCR) was done using vector-specific primers. The narrow end of a flat toothpick was touched to the *E. coli* colony and placed in a microcentrifuge tube containing 50μl of sterile distilled water and 5 μl of this was used directly as a template in a PCR reaction. As soon as the 5μl had been removed, the remaining cells were recovered by adding 955 μl of 2YT+amp$_{100}$ and further cultured at 37°C for several hours. Recovered cultures were used to inoculate 3 ml 2YT+amp$_{100}$ cultures for DNA isolation or to make frozen cell stocks. A typical PCR reaction consisted of 5 μl of dilute cells as template, 2.5 μl of 10X Taq Buffer (Gibco BRL), 0.5 μl of

10 mM dNTP mix (Pharmacia), 0.25 μl of 100 pmol/μl ADH1-terminator primer JN069, and 0.25 μl of 100 pmol/μl of either 'bait' specific primer BC293 or 'prey' specific primer BC304 (See Table 3.1), 5 U Taq Polymerase (Gibco BRL), and made up to 25 μl with sterile distilled water. The PCR reactions were covered by 50 μl of mineral oil and then incubated in a thermocycler (MJC) running the following program: 96°C for 2 min, 32 cycles of 96°C for 45 s, 60°C for 60 s, and 75°C for 2 min, followed by 10 min at 75°C. Samples were electrophoresed through 0.8% (w/v) agarose gels to visualize the amplified DNA products. 2YT+amp$_{100}$ cultures (3 ml) were inoculated with 10 μl of rescued sample for those that corresponded to lanes with amplified inserts and grown at 37°C.

3.7 Isolation of plasmid DNA from *E. coli*

E. coli DNA was isolated by using Wizard DNA Preps (Promega) or by using a method modified from Ish-Horowicz and Burke (20). From each 3 ml overnight culture, 1.5 ml was transferred to a microcentrifuge tube for preparation of plasmid DNA, and 0.5 ml was used to make 1 ml 50% glycerol stocks. Gene inserts, from purified plasmid, were sequenced in an in-house sequencing facility to determine the position of mutations, or introduced into yeast cells to reconfirm the loss- or gain-of-interaction phenotype.

3.8 Primers

All primers used in this project were synthesized by the DNA Technology Unit at PBI-NRC using a Beckman Oligo 1000M and standard phosphoramidite chemistry. Oligos were purified by HPLC (Waters) on a Delta Pak 15m C18 300A column (Waters).

3.9 Sequencing of DNA samples

All sequencing of DNA samples was done by the DNA Technologies Unit at the NRC-PBI using a Perkin Elmer Applied Biosystems 373 DNA Sequencer STRETCH™ in conjunction with the ABI PRISM™ Dye Terminator Cycle Sequencing Kits in combination with primers listed in Table 3.1. Print outs and computer files of the sequence were analyzed using DNASTAR™ software to align the sequenced fragments to their wild type counterparts. In order for a sequence to be scored as mutated, the mutation must be clearly apparent in two separate strands producing overlapping sequence alignments.

3.10 Site-directed mutagenesis

The protocol was a modification of the QuickChange™ Site-Directed Mutagenesis Kit manufactured by Stratagene (Catalog # 200518).

Table 3.1. Primers Used For PCR and Sequencing

Primer Target	Primer Name	Primer Sequence (5'-3')
ADH Terminator (in both 'prey' and 'bait' vectors)	JN069	TTGATTGGAGACTTGACC
GAL4 DB	BC293	GAATAAGTGCGACATCATC
GAL4 TA	BC304	CTATTCGATGATGAAGATACC
AG	LR001	TGCACAGTATTATCAAC
AG	LR002	CTTGTTGATAATACTGTGC
AG, AGL2	SK017	CAAGACACAGTATATGC
AGL2, AGL4	LV001	CAACATGCTCAAGACAC
AGL2	SK022	ATGACTCTCAGAGCATC
AGL2	SK027	TATTGGTTTCAAGCAAC
AGL4	LV002	CTTCCAGCTTCATTGAC
AGL4	LV003	TGTAGCTGTTCTCAAGC
AGL4	LV004	TCAGCATCGCAGAGAAC
AGL4	LV005	TGAAGCAAGTTCGCTGC

3.10.1 Primer design

For each point mutation identified, primer pairs of about 35 bases in length, at least 40 % GC content, and ending in one or more Gs or Cs were designed (Table 3.2.). The point mutation was incorporated as close to the middle of the primer as possible (base 16-18).

3.10.2 PCR

Each nucleotide substitution identified in *AGL2* that resulted in loss of interaction with AG was recreated by site-directed mutagenesis to ensure that the identified mutation was required and sufficient for the observed loss of interaction. Wild type *AGL2* DNA in a pBI881 vector backbone was amplified using pairs of site-directed mutagenesis primers to recreate each nucleotide substitution identified using the 'reverse' two-hybrid screen. The reaction consisted of 5 µl of 10X PCR buffer (Gibco BRL), 10 µl of wild type *AGL2* DNA (0.01 µg/µl), 1 µl of 1 µg/µl of each primer, 1 µl of 10 mM dNTP (Pharmacia), 2.5 U of Native *Pfu* DNA polymerase (Stratagene #600135), and 31 µl of sterile distilled water. The samples were subjected to the following PCR cycle: 95°C for 30 s, 12 cycles of 95°C for 30 s, 55°C for 60 s, and 68°C for 90 s and then cooled to 4°C. Following the PCR reaction, 10 U of the restriction enzyme *Dpn*I (New England Biolabs) was added and each tube was incubated at 37°C for 1 h to digest the parental non-mutated, methylated DNA. An aliquot (1 μl) of the digest was mixed with 40 μl of Epicurian Coli XL2-Blue Ultra Competent Cells® (Stratagene) in a pre-chilled microcentrifuge tube and incubated on ice for 30 min. The samples were heat shocked for 30 s at 42°C

Table 3.2. Primers Used For Site-Directed Mutagenesis

Mutant Number	Primer Name	Location of Mutation in AGL2 cDNA Sequence [nt]	Sequence [5'-3']
1	LR009	380	AGGAGTTAGAGCAGC**c**TGAGCGTCAACTGGACGGC
	LR010**		GCCGTCCAGTTGACGCTCA**g**GCTGCTCTAACTCCT
2	LR011	418	CAAGCAAGTTCGG**c**CCATCAAGACACAGTACATGC
	LR012		GCATGTACTGTGTCTTGATGG**g**CCGAACTTGCTTG
3	LR013	335	CGTCAACAGAGAAATCTTC**c**TGGGGAGGATTTAGG
	LR014		CCTAAATCCTCCCCA**g**GAAGATTTCTCTGTTGACG
4	LR015	392	GCAGCTTGAGCGTCAAC**c**GGACGGCTCTCTCAAGC
	LR016		GCTTGAGAGAGCCGTCC**g**GTTGACGCTCAAGCTGC

* Nucleotides in bold and lowercase type denote the position of the point mutation

** Even numbered primers are the compliments of the preceding odd number primer. e. g. LR010 is the perfect complement of LR009

and then cooled on ice for 2 min. 2YT (0.5 ml) preheated to 42°C was added to rescue the transformants, followed by incubation at 37°C for one hour with shaking. Aliquots (250 μl) were plated onto 2YT+amp$_{100}$ plates and incubated overnight at 37°C. The following day, five - 3 ml 2YT+amp$_{100}$ cultures were inoculated from each separate mutagenesis experiment and DNA was isolated as outlined in Section 3.7.

3.10.3 Confirmation of mutation by restriction analysis

The DNA sequence of *AGL2* was analyzed to determine if the introduction of the individual point mutations would alter the restriction pattern of the sequence, thus allowing a facile initial screen for introduction of the desired point mutation. For all mutants, 2 µl of DNA was restricted with the addition of 1 U appropriate restriction enzyme, 1 µl 10X buffer (supplied with enzyme), and 6 µl sterile distilled water and incubated at 37°C for one hour. Agarose gel electrophoresis was used to assess the introduction of mutations as indicated by alterations to the restriction pattern.

3.10.4 Confirmation of interaction capabilities

Constructs that carried directed *AGL2* mutations were transformed into competent MaV95 cells containing an *AG* gene 'bait' construct as outlined in 3.2.3. Transformants were subjected to marker assays (Section 3.3) in order to determine if the directed mutations in the *AGL2* gene resulted in the predicted loss of interaction with AG as indicated by reduced marker gene expression.

4. RESULTS

Structure-function analysis of the AG-AGL2 and AG-AGL4 interactions involved the production and analysis of non-directed mutant libraries. Three types of mutations were pursued using the 'forward' and 'reverse' two-hybrid systems: *agl4* mutants that had gained the ability to interact with AG, *agl2* mutants that had lost the ability to interact with AG, and *ag* mutants that had lost the ability to interact with AGL2. The position of the mutations were determined through sequence analysis and recreated using site-directed mutagenesis to confirm the altered interaction properties of the gene products.

4.1 Generation of Non-Directed Mutant Libraries

Three independent point mutation libraries were constructed for use in conventional screens (*agl4* mutant library), or 'reverse' two-hybrid systems (*agl2* and *ag* mutant libraries).

4.1.1 Reversion frequency calculations and the induction of point mutations

Reversion frequencies are shown in Table 4.1. An increase of approximately 10 times over non-mutant conditions was considered to be an adequate induction of point mutations throughout the construct. The *ag* library had the highest induction of mutation rate of about 200 times while that of *agl4* was the lowest at only 20 times.

Table 4.1. Point Mutation Library Reversion Frequencies

Point Mutation Library	Reversion Frequency in M9 Media	Reversion Frequency in 2YT Media
ag	3.3×10^{-5}	6.7×10^{-3}
agl2	5.0×10^{-6}	3.8×10^{-4}
agl4	6.7×10^{-6}	1.3×10^{-4}

4.1.2 Library size and recombination frequencies

Library sizes and recombination frequencies are listed in Table 4.2. The independent library size varied by two orders of magnitude with *agl4* being the largest at 1.9×10^6 independent clones and *agl2* the smallest (2.1×10^4 clones). The recombinant frequency of the *ag* mutant library was the lowest, where 90% of clones contained mutant *ag* coding regions. 100% of the *agl2* library vectors contained a recombinant insert while 95% of the *agl4* constructs were recombinant. To prepare mutagenized plasmid DNA for use in subsequent genetic screens DNA was prepared from media innoculated by approximately 1×10^{10} CFU.

4.2 *Agl4* Mutants

Two separate conventional two-hybrid screens were completed using the *agl4* library to screen for mutants that had gained the ability to interact with AG. In the first screen, 1.3×10^5 transformed cells were screened for protein-protein interaction (Table 4.3) of which 215 putative gain-of-interaction mutants were recovered and stored for further analysis. Of these colonies, 34% were positive for *lacZ* activation and 87% were phenotypically 3-AT resistant when grown on SD-leu-trp-his+3-AT media. 34 putative interactors (16%) were again positive for interaction when re-introduced into YPB2 carrying the *AG* 'bait' vector. Upon sequence analysis however, no mutations could be found within the *AGL4* coding domain of these putative mutants.

Table 4.2. Point Mutation Library Sizes and Recombination Frequencies

Point Mutation Library	Library Size (Independent Clones)	Recombinant Frequency (%)
ag	7.0×10^5	90
agl2	2.4×10^4	100
agl4	1.9×10^6	95

Table 4.3. Viable Cells and Number of Transformed Cells per Library Screen

Library Screen	Total Viable Cells (CFU)	Transformed Cells (CFU)
62*AG*-86*agl4*	9.3×10^6	1.3×10^5
62*AG*-86*agl2*	3.5×10^8	3.2×10^4
62*AGL2Δ2*-86*ag*	1.2×10^8	1.2×10^4
62*ag*-86*AGL2*	4.8×10^8	8.2×10^5
62*ag*-86*AGL2* (mutant backbone)	4.7×10^8	3.4×10^5

Transformed Cells is a measure of the total number of clones that were screened for gain- or loss-of-interaction.

4.3 *Agl2* Mutants

From the 3.2 x 10^4 transformants screened from the *agl2* mutant library, 107 putative AGL2-AG loss-of-interaction mutants were identified (Table 4.3). Of the original 107 putative mutants, 89% were negative for *lacZ* expression and displayed only weak 3-AT resistance. When re-introduced into MaV95 containing wild type *AG* in a 'bait' vector, 11% were *lacZ* negative, showed 3-AT sensitivity and were significantly 5-FOA resistant.

4.4 Sequence Analysis

After sequencing, five single-nucleotide C-T transitions that altered either leucine or serine codons to proline were identified, as summarized in Figure 4.1. Note that all of these mutations were localized to the K-box domain, for example agl2-Pro_1 and agl2-Pro_2 represent a change at aa 4 of the K-box in the AGL2 coding region, agl2-Pro_3 is at aa 19, AGL2-Pro_4 is at aa 23, and finally, agl2-Pro_5 is at aa 32. AGL2-Pro_1 and AGL2-Pro_2 represent the recovery of the same point mutation and may be indicative of 'sibs' within the library.

4.4.1 Site-directed mutagenesis

Each of the point mutations identified was predicted to alter the restriction pattern for selected enzymes in the coding region. Therefore, a confirmatory restriction analysis was carried out to pre-screen prospective mutant clones. agl2-Pro_1 and agl2-Pro_2 introduced an additional *Dde*I site into *AGL2*, agl2-Pro_3 gained a *Hae*III site, agl2-Pro_4 gained an *EcoR*II site, and agl2-Pro_5 gained

AGL2	RNLLGEDLGPLNSKELEQLERQLDGSLKQVRSIKTQYMLDQLSDLQNKEQYML
agl2-Pro$_1$	RNL**p**GEDLGPLNSKELEQLERQLDGSLKQVRSIKTQYMLDQLSDLQNKEQYML
agl2-Pro$_2$	RNL**p**GEDLGPLNSKELEQLERQLDGSLKQVRSIKTQYMLDQLSDLQNKEQYML
agl2-Pro$_3$	RNLLGEDLGPLNSKELEQ**p**ERQLDGSLKQVRSIKTQYMLDQLSDLQNKEQYML
agl2-Pro$_4$	RNLLGEDLGPLNSKELEQLERQ**p**DGSLKQVRSIKTQYMLDQLSDLQNKEQYML
agl2-Pro$_5$	RNLLGEDLGPLNSKELEQLERQLDGSLKQVR**p**IKTQYMLDQLSDLQNKEQYML

Figure 4.1. Alignment of K-Boxes of Wild Type AGL2 and agl2-Pro Mutants.
Lowercase bold letters denote positions of amino acid substitutions.

an *Msp*I site. In each case, 2 of the 5 putative site directed mutagenesis colonies analyzed were *bona fide* mutants as indicated by restriction analysis. The identity and loss of interaction properties of these mutants were subsequently confirmed by sequence analysis, transformation and marker gene assays in the two-hybrid system.

4.4.2 *Agl2* mutants and the self-activation of markers

Each *agl2* mutant was subcloned to a 'bait' vector construct and confirmed by sequence analysis prior to transformation into MaV95 or YPB2 competent cells, with or without an AG fusion protein. Further assays were performed to determine if the mutant agl2 fusion was capable of activating marker genes. In the presence or absence of an AG 'prey', all 5 agl2 fusions were unable to activate marker genes.

4.5 *Ag* Mutants

4.5.1 *Ag* mutant library in the 'prey' vector

The initial *ag* mutant library was prepared in the 'prey' vector and screened against the non-self activating version of *AGL2* (*AGL2Δ2*) in the 'bait' vector. *Agl2Δ2* is an N terminus deletion of amino acids 198 to 248 therefore not affecting the MADS-box or K-box domains. When preparing this construct, two erroneous amino acids were added to the N terminus after the deletion was made. When this was discovered, the protein-protein interaction profile of the

construct was analyzed in the two-hybrid system and found to be identical to the wild type AGL2 (27). 1.2 x 10^4 transformants (Table 4.3) were screened for loss of interaction using the mutant *ag* library and 137 putative loss-of-interaction *ag* mutants were identified. 22% were found to lack activation of the *lacZ* marker and 94% displayed sensitivity to 3-AT. The 22% of mutants negative for *lacZ* activation were re-assayed for all demonstrated loss of interaction. However, subsequent sequencing failed to identify mutations in the *AG* coding region among these putative mutants. After examining control transformants (i.e. non mutated forms of Agl2Δ2 in 'bait' and *AG* in 'prey') an inherent irreproducibility in the interaction was observed, since levels of marker activation varied significantly among individual yeast colonies.

4.5.2 *Ag* mutant library in the 'bait' vector

The reciprocal screen was undertaken where a wild type *AGL2* 'prey' construct was screened against a mutant *ag* library in the 'bait' vector. Of the 8.2 x 10^5 colonies screened (Table 4.3), 70 putative mutants were selected of which only 9 showed the appropriate phenotype when re-transformed and assayed for marker activation. 52 demonstrated strong 3-AT resistance and 8 showed strong *lacZ* activation. Many of the colonies that were picked were from the edge of the petri plates and this is an area in which plating density can be higher than normal allowing the growth of colonies on media that would not normally support their growth. After sequence analysis, no mutations within the

AG coding region were identified. Some of the putative loss-of-interaction mutants were the result of vectors with no insert, since these are strongly selected for in a 'reverse' two-hybrid screen and after electroporation of the 70 putative mutant DNAs into *E. coli*, only 9 of these were found to have an insert. Following sequence analysis none of these 9 were found to contain mutations within the *AG* coding region.

4.5.3 *Ag* mutant library in mutant 'bait' vector

In the third attempt to identify *ag* mutants, the library of *ag* mutations in the 'bait' vector was screened for interaction before being subcloned into a non-mutant vector background. From the 3.4×10^5 transformants screened, 100 putative loss-of-interaction mutants were initially identified on SD-leu-trp+0.2% 5-FOA plates, of which 28 were negative for *lacZ* expression and sensitive to 3-AT (Table 4.3). By PCR analysis, 21 clones were found to contain an intact *AG* coding region and sequence analysis confirmed that 20 of these were *AG*. Among these were ten wild type *AG* inserts and the remaining 10 included frame shifts, silent mutations, and nonsense mutations as summarized in Table 4.4.

Table 4.4. *Ag* Point Mutations Resulting in Loss of Interaction with AGL2

Mutant	Type of Mutation	Number of Independent Isolates	Codon Alterations
$ag_{STOP199}$	nonsense	2	CAA → u*AA
$ag_{SHIFT55}$	frameshift	4	insertion of an a
$ag_{STOP154;\ SILENT118}$	nonsense; silent	2	AGA → **u**GA ; CAA → **u**AA
$ag_{SHIFT\ 84}$	frameshift	2	deletion of a c

*Nucleotides in bold and lowercase denote the positions of point mutations

5. DISCUSSION

5.1 The MADS-Box Gene Family

The first MADS-box genes characterized in plants were *DEF* (61) and *GLOBOSA (GLO;* 64) from *Antirrhinum majus*. In *def* and *glo* mutant plants, homeotic transformations occur where petals were converted to sepals, and stamens to carpels (63, 64). Only after the genes were cloned and sequenced were the important structural features of the genes revealed. Since then, numerous plant MADS-box genes have been cloned and sequenced including the large *AGL* gene family of *Arabidopsis*. Most members of this family have been defined and characterized on the basis of their sequence similarity to the *AG* gene sequence. Since in many cases no mutants have been identified, it is not known if all genes in the family affect floral organogenesis. The expression patterns of AGL transcripts are not restricted to the flower, rather some are specific to roots, or leaves (37) and many show extensive expression overlap (31). The question therefore arises whether the *AGL* family is an example of familial genetic redundancy, or whether it represents a group of genes that act in concert, each with a specific role in the regulation of plant development.

The term MADS-box was coined by Schwarz-Sommer in 1990 (57) to describe a family of related sequences in *A. thaliana*, *Antirrhinum majus*, *H. sapiens*, and frog with homology to the *SRF* and *MCM1* transcription factors,

suggesting that these genes played a role in cell differentiation. These and related transcription factors, are generally believed to have two basic functions; the ability to bind to DNA at specific sites and to subsequently recruit RNA polymerase to transcribe nearby genes. In the case of MADS-box proteins, the CarG box has been identified as a consensus DNA binding site (9) but has not been implicated in transcription initiation.

5.2 Protein-Protein Interaction and DNA Binding Activities Among MADS-Box Proteins

Others have studied the ability of several MADS-box proteins, including AP1, AP3, PI, and AG, to bind CarG box sequences (49). These studies used *in vitro* binding and gel retardation assay approaches to show that AP1:AP1, AG:AG, or AP3:PI proteins could bind to the consensus site. Through deletion studies, the K-box or C terminus of these proteins was found not to be required to maintain DNA binding activity *in vitro* (49; 48). These authors also studied the binary protein-protein interaction properties between AG, AP3, PI, and AP1 using immunoprecipitation and electrophoretic mobility shift assays. The results suggested that these four proteins all interacted with each other, providing the K-box or C terminus of the protein was present. These studies concluded that the MADS-box and intervening region functioned as a DNA binding domain, while the K-box mediated non-specific protein-protein interactions and was involved in transcription initiation.

Through AGL2 deletion studies, the K-box domain was found to be required for interaction with AG in the two-hybrid system (31). This supports the results of previous studies in which the intervening region and more C terminal sequences, including the K-box, were concluded to be important for protein interaction (48). Further support for this conclusion arose upon the finding that all AGL cDNAs recovered from two-hybrid screens using AG as a 'bait' construct minimally retained a complete K-box domain (31). However, studies of the AG-AGL24 protein interaction have shown that the MADS-box domain of AG is required for interaction with AGL24 *in vitro* (56). Taken together, these results suggest that different domains may mediate specific MADS-box protein-protein interactions, including the possibility for simultaneous protein-protein interactions with multiple partners.

Previous reports of homo- and hetero-dimer formation between MADS-box proteins have in part been conducted by observing complexes that bind to DNA. These studies therefore examined protein-protein interactions indirectly, since protein-protein interactions were defined in the presence of DNA - a potential third binding component in the reaction (48;49). Thus the 'non-specificity' demonstrated by these studies could be a result of DNA-mediated interaction.

A recent publication has reported interaction between AG and AGL4 in a two-hybrid system utilizing multi-copy 2 μm origin-of-replication vectors (36). This result is at direct odds with our findings, and may be due to multiple copies of both AG and AGL4 expression constructs within the cell. This may result in a

relatively high steady-state concentration of interacting protein partners, thus amplifying or saturating marker gene activation. Under our conditions, β-gal marker gene activation by the AG-AGL2 interaction was 44.2 Miller Units while that of the AG-AGL4 interaction was only 0.2 Miller Units (72). This recent published study did not present quantitative data for marker gene activation, thus a quantitative comparison of the results cannot be made.

5.3 The Two-hybrid System and Point Mutation Analysis

A structure - function analysis of two specific AGL protein-protein interactions, AG-AGL2 and AG-AGL4, was undertaken in an attempt to localize specific amino acid domains important in the interactions.

5.3.1 *Agl4* mutants

Despite having the largest point mutation library size (Table 4.2), no *agl4* mutants that had gained the ability to interact with AG were identified. Several factors can be put forward to explain this. First, only approximately one tenth of the mutant library diversity was screened (1.3×10^5 transformants from a library of 1.9×10^6; Tables 4.2 and 4.3). The mutation rate was good, with a chromosomal his^+ reversion frequency of 1.3×10^{-4} under mutator conditions (Table 4.1). Since the AGL4 cDNA is approximately 1250 bp, only 1 in 10 would carry a point mutation (1.3×10^4 / 1250), and given that 1.3×10^5 transformants were screened (Table 4.3), this is the equivalent to approximately 1.3×10^4 point mutations screened ($1.3. \times 10^5$ / 10). Few clones in the library would be predicted to have more than a single point mutation within the *AGL4* coding

region and it is unlikely that any single mutation event would result in a gain of interaction mutant with AG. Therefore a picture emerges in which a low number of mutations were introduced into the gene and the library was partially screened. Further studies could be done in which a larger portion of the library was screened in an attempt to identify interacting clones. It may also be useful to remake the library using alternate methods to introduce a higher frequency of point mutations that would result in multiple point mutations in each *agl4* coding region domain.

5.3.2 *Agl2* mutants

Although, the *agl2* point mutation library was almost 100-fold smaller than the *agl4* library (Table 4.2), the screen was successful in the identification of 5 *AGL2* missense mutants that exhibited a reduced ability to interact with AG. The reversion frequency of 3.8×10^{-4} was very similar to the *agl4* mutant library (Table 4.1) however the fraction of the library diversity that was screened in each case varied considerably. The *agl4* library screen surveyed approximately one tenth of the library diversity, while in the *agl2* screen, a more complete survey of the mutant library was achieved (3.2×10^4 transformants from a library of 2.4×10^4; Table 4.3), however only 1.0×10^3 point mutations were screened. (See 5.3.1 for calculation). Despite only screening one-tenth as many point mutations, the screen for *agl2* mutants was successful. From this it appears that is easier to matutinally interfere with an interaction that to create one.

5.3.3 Agl2 mutants, the K-box, and loss of transcription activation

In this study, a non-directed mutagenesis approach was used in order to avoid assumptions about the structure-function relationships mediating protein interaction and transcriptional activation among the proteins being studied. The recovery of partial *agl2* cDNAs encoding proteins capable of interacting with AG have previously suggested the importance of the K-box domain for protein-protein interactions to occur. In addition, earlier C-terminal deletion studies localized the transcription activation capabilities of AGL2 to this region (27). Using a non-directed mutagenesis approach, our results support this conclusion by finding that all five loss-of-interaction missense mutations were located in the K-box domain (Figure 4.1). These mutants, all of which replace leucine and serine residues with proline, are predicted to disrupt the putative coiled-coil structure of the K-box by inserting a β-turn in the protein and thereby abolishing the capability for protein - protein interactions. The suggested disruption affected not only the ability to interact with AG, but the agl2 mutant proteins were no longer able to activate transcription in the two-hybrid system when expressed from bait vectors. The interaction of the mutant *agl2* protein expressed as a 'bait' fusion with wild type AG as a 'prey' fusion also failed to show activation of the marker genes. I suggest that the disruption of K-box tertiary structure in these mutants may likewise disrupt the protein's capacity for

interaction with AG. It would seem logical that the region C-terminal to the K-box would have a role in the different activities of AGL2 and AGL4 in the two-hybrid system in light of their reduced sequence similarity C-terminal to the K-box domain (Figure 2.5).

5.3.4 *Ag* mutants

5.3.4.1 Aspects of the agl2Δ2 - AG interaction

The agl2Δ2 derivative was constructed in an attempt to identify an agl2 deletion construct which lacked the ability to activate transcription in the two-hybrid system. The agl2Δ2 deletion lacked the C terminal 52 amino acids of the wild type protein, leaving the MADS-box and K-box regions intact. Importantly, the deletion protein exhibited the same protein-protein interaction profile vis-a-vis AGL8, AGL15, and AGL22 as did the wild type AGL2 protein in the two-hybrid system (25). However, upon further analysis it was determined that the interaction between agl2Δ2 and AG was variable in the two-hybrid system (28). Putative ag mutants identified in a reverse two-hybrid screen of 1.2×10^4 clones, or 1.8×10^3 mutations (see 5.3.1 for calculation), with *agl2Δ2* as the bait vector (Table 4.3) yielded results that initially were scored positive for interaction but later appeared to be the result of unstable marker gene activation. It should be noted that the library diversity of 7.0×10^5 independent clones was acceptable (Table 4.2), however, the frequency of null (non-recombinant) constructs was high, with 10% of clones lacking inserts (Table 4.2).

5.3.4.2 Screening for *ag* mutants

A second approach to screen for ag loss-of-interaction mutants employed a wild type *AGL2* gene in the prey vector, screened against an *ag* point mutation library in the bait vector. Again, several putative loss of interaction mutants were identified in a screen of 8.2×10^5 independent clones, corresponding to approximately 1.2×10^5 mutations (Table 4.3; see 5.3.1 for calculation). However, in all cases, the putative 'mutant' vector was found to be either non-recombinant or its sequence was indistinguishable from wild type AG.

The problem of the reverse two-hybrid system selecting for null vectors is inherent to the system. Although null vectors number less than 10% of the library, they do not interact with the bait protein and are therefore selected as phenotypic (FOA^R) loss of interaction mutants. The empty vectors presumably arose via self-ligation of the non-mutant input vector. To circumvent these problems, the library was screened prior to the subcloning stage in an attempt to eliminate the selection of empty vectors as putative mutant constructs.

5.3.4.3 Screening the AG mutant library in a mutated vector backbone

Nine mutant *ag* clones were identified from this third and final screen of $3.4.x\ 10^5$ clones (Table 4.3), or $5.1. \times 10^4$ mutations (see 5.3.1 for calculation), but they provided little useful information. The mutants were characterized by frameshift mutations including the insertion of an 'A' at nucleotide 55 of the cDNA in four clones and the deletion of a 'C' at nucleotide 84 of the cDNA in two clones. In addition two nonsense mutations caused by point mutations at

nucleotides 154 and 199 (Table 4.4) were predicted to prematurely terminate the protein prior to the K-Box domain and consequently abolish any capacity for protein interaction. If missense mutations had been identified, or if the introduction of stop codons had occurred later in the cDNA (perhaps in the K-box or in the intervening region between the MADS-box and K-box) a clearer picture of the important domains may have been elucidated.

5.3.5 False-positives

Throughout all of these screens a number of false-positives were initially picked as loss- or gain-of interaction mutants. Several reasons for this have been offered in the preceding text, including the selection for null vectors and the irreproducibility of the agl2Δ2-AG interaction. Many of the problems with the selection of false positives arose from a lack of experience with the two-hybrid system and the inherent qualitative colony selection criteria. When plating transformed cells onto agar plates pooling can occur on the plate, resulting in a higher localized plating density. The control of plating density is known to be an important factor that must be carefully controlled to avoid two-hybrid artifacts (26). Under sub-optimal plating conditions colonies that grow on selective plates may not display the same characteristics when subsequently re-checked for marker gene activation. It can also be difficult to determine what constitutes 'adequate' marker activation. The delineation between growth and absence of

growth on selective media and colour development in an X-gal assay are both qualitative and time-dependent. The selection of false-positives can be reduced by experience with the two-hybrid system, including careful plating techniques and standardized marker gene assays.

5.4 Specificity of Function in the AGL Gene Family

Previous work by others has implied that protein interactions between MADS-box proteins are general and non-specific (54, 49). However, the two-hybrid analyses presented here contradicts this interpretation in several respects. Screens of an *A. thaliana* two-hybrid cDNA expression library using MADS-box 'bait' proteins such as AG revealed a specific matrix of interacting MADS-box proteins, all of which display overlapping *in situ* expression patterns (31). However, the converse was not true; not all proteins with overlapping expression patterns were found to interact. In addition, when directly testing for binary interaction between MADS-box proteins in the two-hybrid system, no interaction was detected between canonical members of the family including AGL2-AGL4, AGL4-AP1, or AG-AP1, to name a few (72). These findings point to a specificity of interaction between MADS-box proteins as defined by the yeast two-hybrid system.

Our findings show that AGL2 is a strong transcriptional activator in the yeast two-hybrid system while AGL4 is not. To our knowledge, PI and AP1 are the only other MADS-box proteins that exhibit this property out of a growing

number of proteins tested including AGL9, AGL15, AGL18, AGL20, AGL22, AGL24, CAL and AP3 (29).These findings point to a specificity of transactivation function of MADS-box proteins as defined by the yeast two-hybrid system.

Studies of the AG-AGL24 and AG-AGL2 protein interactions have shown that the AG MADS-box domain is required for interaction with AGL24 (56). These results suggest that multiple domains can be involved in mediating MADS-box protein-protein interactions, and may allow for simultaneous protein-protein interactions with multiple partners via independent domains.

It is tempting to draw certain parallels between MADS-box proteins, as putative transcription factors, and the functional properties of TAFs as members of the transcription factor TFIID complex in eukaryotes (46). These proteins play specific roles in transcription initiation largely mediated by protein-protein interactions (76). TAFs may exhibit specificity of function including the ability to bind DNA or to interact with portions of the RNA polymerase II holoenzyme complex. A similar diversity of function with respect to protein-protein interactions has been demonstrated for the members of the AGL gene family.

As implied by the 'protein recruitment' model of TAF function, these properties of protein-protein interactions, including specificity and affinity, could alter the transcription specificity exhibited by the protein complex. For example, the relatively strong AG-AGL2 interaction (as defined by β-gal assays) would form relatively more stable complexes early in floral morphogenesis than the weaker AG-AGL4 interaction, with possible downstream consequences for selective gene transcription required for establishing organ identity. One can

also envisage direct competition for binding and differential transcription activation properties based simply on the binding interaction affinities among these putative transcription factors.

The apparent redundancy within the *AGL* gene family could be viewed in a different perspective. The specific properties of auto-activation, protein-protein interactions, and the domains which mediate them together argue that MADS-box proteins may exert different functions in the regulation of transcription. Furthermore, varied properties of these protein-protein interactions could influence the specificity for selective gene activation by these putative transcription factors. As with TAFs, the *AGL* family has demonstrated diversity in function as assessed by the two-hybrid system.

The protein complexes defined by this gene family would be able to perform a variety of functions ultimately leading to the initiation of specific OIGs in different developmental compartments. Due to the overlapping expression patterns of the AGLs and the capability of protein interaction with multiple partners, the combinatorial diversity of complexes capable of being formed is large. This specialization of function exhibited by members of the AGL gene family, together with the large number of protein complexes capable of being formed in and between specific developmental compartments within the flower, could lead to precise control of diverse gene expression

5.5 Further Studies

Studies described here were entirely conducted using the yeast two-hybrid system. It is therefore important that the functional properties of these proteins be reconciled *in planta*. Do the identified interactions, and properties of transactivation persist in the plant? Do the loss-of-interaction mutants, defined in yeast, have any consequences for floral development? In order for these findings and implications to be validated, further studies must be done.

5.5.1 Confirmation of two-hybrid data *in planta*

As well as their protein-protein interaction properties, the transcription activation properties of proteins like AGL2 need to be verified *in planta.* Experiments using GAL4 mediated marker gene cassettes in which a DB-*AGL2* and *GAL4-UAS -GUS* or *-Green Flourescent Protein* cassette could be introduced into plant tissue by particle bombardment and used to assess activation function *in planta*. Transcriptional activation by AGL2 could be approached by expression of a GAL4 DB:AG fusion, that would be predicted to co-activate a GAL4-regulated reporter gene construct when co-expressed with AGL2. Further work on assessing the transcriptional activation properties of AGL2 *in planta* is currently underway.

To verify protein-protein interactions, direct biochemical approaches could be taken where protein complexes would be isolated from transgenic plants expressing epitope tagged fusions and the composition of these complexes determined. Difficulties could arise with genes expressed early in

flower development where limited amounts of sample would be available. Determining the precise location of where the complexes are formed would also be complicated by the asynchronous nature of flower development where not all floral primordia are in the same stage of development.

5.5.2 MADS-box genes as transcription factors

By virtue of their general structure, MADS-box genes fall into the category of putative transcription factors, some of which are known to play a role in floral development. In this regard, *ag* missense mutants could be assayed in complementation tests using mutants altered in their interaction capabilities vis-a-vis known partners such as AGL2, AGL24, AGL9, or AGL15 and their subsequent effects upon floral development determined. This loss of interaction may result in a change of gene regulation and give rise to phenotypic changes in the plant. A PCR-based differential display strategy could be used to identify target genes altered in transcript abundance (69). Mutants with a decreased affinity for protein interaction (e.g. agI2 mutants with a decreased affinity for interaction with ag) could be studied in a similar manner for their affects on morphogenesis. In this case, phenotypic changes would be expected to be less severe than if the interaction was completely absent. Studies such as these may shed further light on the specific functions of selected members of the AGL gene family in *A thaliana*.

5.6 Conclusion

This study has revealed specific properties of functional specialization among the MADS-box family of proteins in *A. thaliana*. Rather than genetic redundancy, this family of structurally similar genes might be more profitably viewed by examining the protein-protein interaction and transcriptional activation properties determined by the subtle differences among its members. The 7% difference in primary sequence identity between AGL2 and AGL4 may be more significant than initially supposed especially when, as seen here, a single amino acid substitution can dramatically affect the protein-protein interaction and transcription activation properties of representative proteins such as AGL2. Notions of "redundancy' in the AGL family may well yield to a closer examination of the subtle functional *differences* between these gene products. By furthering the research beyond the findings and interpretations presented here, much could be learned about the mechanistic role of the products of the *AGL* gene family in *Arabidopsis* development.

6. REFERENCES

1. **Ainsworth, C., S. Crossley, V. Buchanan-Wollaston, M. Thangavelu, and J. Parker**. 1995. Male and female flowers of the dioecious plant Sorrel show different patterns of MADS-Box gene expression. Plant Cell **7**:1583-1598.

2. **Aspenstrom, P. and M.F. Olson**. 1995. Yeast two-hybrid system to detect protein-protein interactions with Rho GTPases. Small GTPases and their Regulators **256**:228-241.

3. **Bartel, P.L., C.-T. Chien, R. Sternglanz, and S. Fields**. 1993. Using the two-hybrid system to detect protein-protein interactions, p. 153-179. In D.A. Hartley (ed.), Cellular Interactions in Development: A Practical Approach. Oxford, Oxford University Press.

4. **Bowman, J.L., H. Sakai, T. Jack, D. Weigel, U.X. Mayer, and E.M. Meyerowitz**. 1992. SUPERMAN, a regulator of floral homeotic genes in *Arabidopsis*. Develop. **114**:599-615.

5. **Brondyk, W.H. and I.G. Macara**. 1995. Use of two-hybrid system to identify Rab binding proteins. Small GTPases and their Regulators **257**:200-208.

6. **Castello, A., V. Brocheriou, P. Chafey, A. Kahn, and H. Gilgenkrantz**. 1996. Characterization of the dystrophin-syntrophin interaction using the two-hybrid system in yeast. FEBS Lett. **383**:124-128.

7. **Chevray, P.M. and D. Nathans**. 1992. Protein interaction cloning in yeast: Identification of mammalian proteins that react with the leucine zipper of Jun. Proc. Natl. Acad. Sci. USA **89**:5789-5793.

8. **Clemens, K.E., R. Brent, J. Gyuris, and K.E. Münger**. 1995. Dimerization of the human papillomavirus E7 oncoprotein *in vivo*. Virol. **214**:289-293.

9. **Davies, B. and Z. Schwarz-Sommer**. 1994. Control of floral organ identity by homeotic MADS-box transcription factors, p. 235-258. In Results and Problems in Cell Differentiation. Berlin, Heidelberg, Springer Verlag.

10. **Dhordain, P., O. Albagli, S. Ansieau, M.H.M. Koken, C. Deweindt, S. Quief, D. Lantoine, A. Leutz, J.P. Kerckaert, and D. Leprince**. 1995. The BTB/POZ domain targets the LAZ3/BCL6 oncoprotein to nuclear dots and mediates homomerisation in vivo. Oncogene **11**:2689-2697.

11. **Fields, S. and O.K. Song**. 1989. A novel genetic system to detect protein-protein interactions. Nature **340**:245-246.

12. **Fischer, A., N. Baum, H. Saedler, and G. Theiβen**. 1995. Chromosomal mapping of the MADS-box multigene family of *Zea mays* reveals dispersed distribution of allelic genes as well as transposed copies. Nucl. Acids Res. **23**:1901-1911.

13. **Gstaiger, M., C. Hovens, O. Georgiev, L. Knoepfel, and W. Schaffner**. 1996. BZLF1 (ZEBRA, Zta) protein of epstein-barr virus selected in a yeast one-hybrid system by binding to a consensus site in the IgH intronic enhancer: a role in immunoglobulin expression? Biol. Chem. **377**:669-673.

14. **Hardenack, S., D. Ye, H. Saedler, and S. Grant**. 1994. Comparison of MADS box gene expression in developing male and female flowers of the dioecious plant white campion. Plant Cell **6**:1775-1787.

15. **Heard, J. and K. Dunn**. 1995. Symbiotic induction of a MADS-box gene during development of alfalfa root nodules. Proc. Natl. Acad. Sci. USA **92**:5273-5277.

16. **Heck, G.R., S.E. Perry, K.W. Nichols, and D.E. Fernandez**. 1995. AGL15, a MADS domain protein expressed in developing embryos. Plant Cell **7**:1271-1282.

17. **Hollinger, P.** 1993. Yeast Strain KD1122pCJ. Personal communication

18. **Hong, Y., K. Levay, J.F. Murphy, P.G. Klein, J.G. Shaw, and A.G. Hunt**. 1995. A potyvirus polymerase interacts with the viral coat protein and VPg in yeast cells. Virol. **214**:159-166.

19. **Inouye, C., P. Remondelli, M. Karin, and S. Elledge**. 1994. Isolation of a cDNA encoding a metal response element binding protein using a novel expression cloning procedure: the one hybrid system. DNA and Cell Biology **13**:731-742.

20. **Ish-Horowicz, D. and J.F. Burke**. 1981. Rapid and efficient cosmid cloning. Nucl. Acids Res. **9**:2989-2998.

21. **Jack, T., L.L. Brockmann, and E.M. Meyerowitz**. 1992. The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. Cell **68**:683-697.

22. **Kang, H.G., Y.S. Noh, Y.Y. Chung, M.A. Costa, K.S. An, and G.H. An**. 1995. Phenotypic alterations of petal and sepal by ectopic expression of a rice MADS box gene in tobacco. Plant Mol. Biol. **29**:1-10.

23. **Keegan, K. and J.A. Cooper**. 1996. Use of the two hybrid system to detect the association of the protein-tyrosine-phosphatase, SHPTP2, with another SH2- containing protein, Grb7. Oncogene **12**:1537-1544.

24. **Kempin, S.A., M.A. Mandel, and M.F. Yanofsky**. 1993. Conversion of perianth into reproductive organs by ectopic expression of the tobacco floral homeotic gene *NAG1*. Plant Physiol. **103**:1041-1046.

25. **Kohalmi, S. E.** 1995. *Agl2Δ2* has normal interaction capabilities but does not activate transcription. Personal communication.

26. **Kohalmi, S. E.** 1996. Plating density and its effect of false positives recovered in the two-hybrid system. Personal communication.

27. **Kohalmi, S. E.** 1997a. *Agl2Δ2* sequence and interaction capabilities. Personal communication

28. **Kohalmi, S. E.** 1997. The Agl2Δ2 - AG interaction is less stable than the AGL2-AG interaction. Personal communication

29. **Kohalmi, S. E.** 1997c. Two-hybrid 'Bait' fusion proteins capable of transcriptional activation. Personal communication

30. **Kohalmi, S. E., J. Nowak, and W.L. Crosby**. 1997. The Yeast Two-Hybrid System, p. 63-82. In E. Hansen and G. Harper (ed.), Differentially Expressed Genes In Plants: A Bench Manual. Taylor & Francis Ltd., London.

31. **Kohalmi, S. E., P. Schorr, J. Nowak, L.J.V. Reader, and W.L. Crosby**. 1997. The product of *AGAMOUS* interacts with select members of the MADS-box family of proteins in *Arabidopsis thaliana*. Plant Journal (Submitted).

32. **Leanna, C.A. and M. Hannink**. 1996. The reverse two-hybrid system: A genetic scheme for selection against specific protein/protein interactions. Nucl. Acids Res. **24**:3341-3347.

33. **Li, J.J. and I. Herskowitz**. 1993. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science **262**:1870-1874.

34. **Liu, Z.C. and E.M. Meyerowitz**. 1995. LEUNIG regulates AGAMOUS expression in *Arabidopsis* flowers. Develop. **121**:975-991.

35. **Lu, Z.-X., M. Wu, C.-S. Loh, C.-Y. Yeong, and C.-J. Goh**. 1993. Nucleotide sequence of a flower-specific MADS box cDNA clone from orchid. Plant Mol. Biol. **23**:901-904.

36. **Ma, H.** 1997. Plant J. (In Press)

37. **Ma, H., M.F. Yanofsky, and E.M. Meyerowitz**. 1991. *AGL1-AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. Genes Develop. **5**:484-495.

38. **Marykwas, D.L., S.A. Schmidt, and H.C. Berg**. 1996. Interacting components of the flagellar motor of Escherichia coli revealed by the two-hybrid system in yeast. J. Mol. Biol. **256**:564-576.

39. **Miller, H.G., T.D. Kocher, and B. Loy**. 1995. New MADS box domains in *Asparagus officinalis* L. Sexual Plant Reproduction **8**:318-320.

40. **Mizukami, Y., H. Huang, M. Tudor, Y. Hu, and H. Ma**. 1996. Functional domains of the floral regulator AGAMOUS: Characterization of the DNA binding domain and analysis of dominant negative mutations. Plant Cell **8**:831-845.

41. **Mosteller, R.D., W. Park, and D. Broek**. 1995. Analysis of interaction between Ras and CDC25 guanine nucleotide exchange factor using yeast GAL4 two-hybrid system. Small GTPases and their regulators **255**:135-148.

42. **O'Reilly, E.K., N.J. Tang, P. Ahlquist, and C.C. Kao**. 1995. Biochemical and genetic analyses of the interaction between the helicase-like and polymerase-like proteins of the brome mosaic virus. Virol. **214**:59-71.

43. **Phizicky, E.M. and S. Fields**. 1995. Protein-protein interactions: methods for detection and analysis. Microbiol. Rev. **59**:94-123.

44. **Pnueli, L., A.D. Hareven, S.D. Rounsley, M.F. Yanofsky, and E. Lifschitz**. 1994. Isolation of the tomato AGAMOUS gene TAG1 and analysis of its homeotic role in transgenic plants. The Plant Cell **6**:163-173.

45. **Poullet, P. and F. Tamanoi**. 1995. Use of yeast two-hybrid system to evaluate Ras interactions with neurofibromin-GTPase-activating protein. Small GTPases and Their Regulators **255**:488-497.

46. **Ptashne, M. and A. Gann**. 1997. Transcriptional activation by recruitment. Nature **386**:569-577.

47. **Purugganan, M.D., S.D. Rounsley, R.J. Schmidt, and M.F. Yanofsky**. 1995. Molecular evolution of flower development: Diversification of the plant MADS-box regulatory gene family. Genetics **140**:345-356.

48. **Riechmann, J.L., B.A. Krizek, and E.M. Meyerowitz**. 1996. Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. Proc. Natl. Acad. Sci. USA **93**:4793-4798.

49. **Riechmann, J.L., M.Q. Wang, and E.M. Meyerowitz**. 1996. DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucl. Acids. Res. **24**:3134-3141.

50. **Rocchi, S., S. Tartare-Deckert, D. Sawka-Verhelle, A. Gamha, and E. Van Obberghen**. 1996. Interaction of SH2-containing protein tyrosine phosphatase 2 with the insulin receptor and the insulin-like growth factor-I receptor: Studies of the domains involved using the yeast two-hybrid system. Endocrinology **137**:4944-4952.

51. **Rossi, F., A. Gallina, and G. Milanesi**. 1996. NEF-CD4 physical interaction sensed with the yeast two- hybrid system. Virol. **217**:397-403.

52. **Rounsley, S.D., G.S. Ditta, and M.F. Yanofsky**. 1995. Diverse roles for MADS box genes in arabidopsis development. The Plant Cell **7**:1259-1269.

53. **Sambrook, J., Fritsch, E. F., and Maniatis, T.** 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

54. **Sardet, C., M. Vidal, D. Cobrink, Y. Geng, C. Onufryk, A. Chen, and R.A. Weinberg**. 1995. E2F-4 and E2F-5, two members of the E2F family are expressed in the early phases of the cell cycle. Proc. Natl. Acad. Sci. USA **92**:2403-2407.

55. **Schiestl, R.H. and R.D. Gietz**. 1989. High efficiency transformation of intact yeast cells using single stranded nucelic acids as a carrier. Curr. Genet. **16**:339-346.

56. **Schorr, P. and Crosby, W. L.** 1997. Verification of two-hybrid interactions. Personal communication.

57. **Schwarz-Sommer, Z., P. Huijser, W. Nacken, H. Saedler, and H. Sommer**. 1990. Genetic control of flower development by homeotic genes in *Antirrhinum majus*. Science **250**:936

58. **Sengupta, D.J., B.L. Zhang, B. Kraemer, P. Pochart, S. Fields, and M. Wickens**. 1996. A three-hybrid system to detect RNA-protein interactions *in vivo*. Proc. Natl. Acad. Sci. USA **93**:8496-8501.

59. **Shore, P. and A.D. Sharrocks**. 1995. The MADS-box family of transcription factors. Eur. J. Biochem. **229**:1-13.

60. **Sieburth, L., M.P. Running, and E. Meyerowitz**. 1997. Genetic separation of third and fourth whorl functions of *AGAMOUS*. The Plant Cell **7**:1249-1258.

61. **Sommer, H., J.-P. Beltran, P. Huijser, H. Pape, W.-E. Lönnig, H. Saedler, and Z. Schwarz-Sommer**. 1990. *Deficiens*: a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. EMBO J. **9**:605-613.

62. **Steinert, P.M. and D.R. Roop**. 1988. Molecular and cellular biology of intermediate filaments. Annu. Rev. Biochem. **57**:593-625.

63. **Treisman, R. and G. Ammerer**. 1992. The SRF and MCM1 transcription factors. Curr. Opin. Genet. Develop. **2**:221-226.

64. **Tröbner, W., L. Ramirez, P. Motte, I. Hue, P. Huijser, W.-E. Lönnig, H. Saedler, H. Sommer, and Z. Schwarz-Sommer**. 1992. *Globosa*: a homeotic gene which interacts with *Deficiens* in the control of *Antirrhinum* floral organogenesis. EMBO J. **11** :4693-4704.

65. **Tsuchimoto, S., A. van der Krol, and N.-H. Chua**. 1993. Ectopic Expression of pMADS3 in Transgenic Petunia Phenocopies the Petunia blind Mutant. The Plant Cell **5**:843-853.

66. **van Tunen, A.J.** 1996. Molecular control of floral organogenesis in higher plants: The ABCD model. Neth. J. Zool. **46**:91-96.

67. **Vidal, M., P. Braun, E. Chen, J.D. Boeke, and E. Harlow**. 1996. Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. Proc. Natl. Acad. Sci. USA **93**:10321-10326.

68. **Vojtek, A.B. and S.M. Hollenberg**. 1995. Ras-Raf interaction: Two-hybrid analysis. Small GTPases and Their Regulators **255**:331-342.

69. **Walker, S.S., W. Shen, J.C. Reese, L.M. Apone, and M.R. Green**. 1997. Yeast $TAF_{II}145$ required for transcription of G1/S genes and regulated by the cellular growth state. Cell **90**:607-614.

70. **Wang, H., G.A. Peters, X. Zeng, M.L. Tang, I. Wallace, and S.A. Khan**. 1995. Yeast two-hybrid system demonstrates that estrogen receptor dimerization is ligand-dependent *in vivo*. J. Biol. Chem. **270**:23322-23329.

71. **Weigel, D. and E.M. Meyerowitz** . 1994. The ABCs of floral homeotic genes. Cell **78**:203-209.

72. **Wigness, M.** 1997. β-gal assays. Personal communication

73. **Yang, H. and T. Evans**. 1995. Homotypic interations of chicken GATA-1 can mediate transcriptional activation. Mol. Cell. Biol. **15**:1353-1363.

74. **Zhang, J. and S. Lautar**. 1996. A yeast three-hybrid method to clone ternary protein complex components. Anal. Biochem. **242**:68-72.

75. **Zhao, R.** 1997. TA insert activates marker genes in the two-hybrid system in the absence of a DB vector. Personal communication

7. APPENDICES

7.1 Media

7.1.1 2YT media

For one liter of medium, 16.0 g of tryptone, 10.0 g yeast extract, 5.0 g sodium chloride, and 0.5 ml of 5N sodium hydroxide were mixed and made up to a final volume of 1 L before autoclaving. For solid medium, 15 g of agar was added prior to autoclaving. Plates containing ampicillin could also be made by adding 10 ml of a 10 mg/ml sterile stock solution of ampicillin for a final concentration of 100μg amp/ml 2YT.

7.1.2 M9 media

To make 1 L of medium, 750 ml of sterile water, 20 ml of 50 mM $CaCl_2$, 10 ml of 10% glucose, 200 ml of M9 salts, 2 ml of 1 M $MgSO_4$, 2 ml each of 10 mg/ml aa stocks of arginine, threonine, thiamine, leucine, and histidine, 4 ml of 5 mg/ml stock solution of tryptophan, 880 μl 34 mg/ml stock solution of chloramphenicol and 5.0 ml of 10 mg/ml solution of ampicillin were mixed and made up to 1 L. For solid medium, 15 g of agar was added before autoclaving and the solution was cooled to 50°C before the addition of the remaining ingredients. This medium can also be made without histidine for selection purposes and/or ampicillin (100 μg amp/ml media) to accommodate for the presence or absence of various plasmids.

7.1.3 SD media

For 1 L of medium, mix 20.0 g of glucose, 6.7 g of yeast nitrogen base without amino acids, 1.5 g of the appropriate aa dropout powder, and one drop of 5 N NaOH before autoclaving to adjust pH to 5.5±0.2 (25°C). After autoclaving and cooling to 50°C, 2.5 ml of a 2 M stock solution of 3-AT can be added for a final concentration of 5 mM or 2 g of 5-FOA powder could be added to the cooled media for a final concentration of 0.2% 5-FOA (w/v) for selection purposes.

7.1.3.1 Amino acid dropout powder

To make dropout powder the components in Table 7.1 are combined and stored at room temperature. When making selective media (e.g. SD-leu-trp) specific aa are left out of the mixture - in this case L-leucine and L-tryptophan.

7.1.3.2 Yeast nitrogen base without amino acids

Yeast nitrogen base without amino acids can be bought commercially but it is much more economical to mix the ingredients in the lab. For 1 kg of yeast nitrogen base without amino acids the items listed in Table 7.2 were placed in a mill (U.S. Stoneware) for 24 h at 40% power. At this time, the yeast nitrogen base was checked to determine if a homogenous mixture had been formed. If clumps of the material were present on the side or lid of the mill, it was scraped off and milled for a further 24 h. When the milling process was complete, the base was transferred to a 1L plastic bottle and stored at room temperature.

Table 7.1. Amino Acid Drop Out Powder for SD Media

Component	Mass (g)
Adenine sulphate	4.0
L-arginine HCN	2.0
L-aspartic acid	10.0
L-glutamic acid	10.0
L-histidine HCl	2.0
L-isoleucine	3.0
L-leucine	6.0
L-lysine HCl	3.0
L-methionine	2.0
L-phenylalanine	5.0
L-serine	37.4
L-threonine	20.0
L-tryptophan	4.0
L-lysine	3.0
Uracil	2.0
Valine	15.0

This recipe makes 128.4 g which is enough for 85.6 L of SD media (1.5 g/L). Drop out powders lacking certain aa (e.g. leucine or tryptophan) are prepared to select for 'bait' or 'prey' vectors in the yeast two-hybrid system

Table 7.2. Yeast Nitrogen Base Without Amino Acids (1kg)

Component	Mass (g)
$(NH_4)_2SO_4$	746
KH_2PO_4	150
$MgSO_4 \cdot 7H_2O$	75
NaCl	15
$CaCl_2 \cdot 2H_2O$	15
Inositol	0.3
H_3BO_3	0.075
$ZnSO_4 \cdot 7H_2O$	0.06
$MnSO4 \cdot H_2O$	0.06
Thiamine·HCl	0.06
Pyridoxine·HCl	0.06
Niacin	0.06
Calcium Pantothenate	0.06
p-aminobenzoic acid	0.03
Riboflavin	0.03
$FeCl_3$	0.03
$Na_2MoO_4 \cdot 2H_2O$	0.03
KI	0.015
$CuSO_4 \cdot 5H2O$	0.006
Folic Acid	0.0003
Biotin	0.0003