RIF1 Is an Essential Chromatin Remodeling Transcription Cofactor in the Maize Anthocyanin Production Pathway

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with Research Distinction in Molecular Genetics in the Undergraduate Colleges of the Ohio State University

by

Kenneth Frame

The Ohio State University

Project Advisor: Dr. Erich Grotewold, Professor, Dept. of Plant Cellular and Molecular Biology
Abstract

My research centers on the characterization of the transcription cofactor RIF1 in Zea mays (corn). The Grotewold lab has already done extensive research into many of the key transcription factors involved in regulating the production of anthocyanins (red pigments created from the flavonoid pathway), in maize. Previous research by our lab has provided strong evidence that the R and C1 transcription factors are absolutely necessary for the production of anthocyanins in maize. RIF1 is a cofactor that has both an Agenet domain and an ENT domain. The presence of an ENT domain suggests that RIF1 may be involved in chromatin remodeling, as this domain is part of a family of motifs that come from a common predecessor that binds methylated DNA substrates. Due to data from protein-protein interaction experiments, it is believed that RIF1 recognizes the bHLH (basic helix-loop-helix) region of R in order to allow chromatin remodeling and subsequent transcription, although the order of events is still not known. This has potential applications in that there is an analogous situation in humans, wherein the protein BRCA2, which is believed to aid against breast cancer, requires the interaction of a cofactor EMSY to work properly and suppress disease. EMSY and RIF1 are similar in that they both contain an ENT domain (EMSY is actually the source of the name of the ENT domain), and they both homodimerize through this domain. The two proteins are also related in that the Agenet domain of RIF1 is similar to the domains found in other EMSY-interacting proteins. I have already been able to suppress anthocyanin formation with RNAi for RIF1 in microparticle bombardment experiments, thus suggesting that RIF1 is essential for anthocyanin formation (Hernandez Et. Al., PNAS 2007), and I am preparing to investigate the phenotype of RIF1 knockouts obtained from the Maize TILLING Project (http://genome.purdue.edu/maizetilling/OtherProjects.htm). Additional experiments will seek to
pinpoint RIF1’s role in chromatin remodeling (if any) by microparticle bombardment experiments with histone deacetylase inhibitors, and also by ChIP experiments using antibodies against RIF1. Yeast two-hybrid experiments will also be carried out to try to identify additional RIF1 partners in maize, perhaps identifying proteins that are recruited to RIF1 in order to carry out chromatin remodeling. The role of other R-interacting factors in maize will also be researched, as well as their corresponding orthologs in Arabidopsis. The resulting thesis will thus present a comprehensive characterization of the maize protein RIF1, and in doing so will provide valuable insight on epigenetics and transcription regulation in general, and on the maize anthocyanin formation pathway in particular, while at the same time shedding light on an analogous protein interaction in humans that could have implications in cancer research.
Acknowledgements

I would like to acknowledge first and foremost the help of my advisor, Dr. Erich Grotewold, who went above the call of simply providing instruction in research techniques and instilled in me the thought process of a scientist. I would also like to thank Antje Feller, without whose knowledge, experience, and endless patience I would have been unable to complete any research at all. Additionally I would like to thank former graduate student Marcela Hernandez and all of the other members of the Grotewold lab for their invaluable assistance.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>1</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>3</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>4</td>
</tr>
<tr>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>11</td>
</tr>
<tr>
<td>Results</td>
<td>14</td>
</tr>
<tr>
<td>Discussion</td>
<td>16</td>
</tr>
<tr>
<td>References</td>
<td>18</td>
</tr>
<tr>
<td>Tables and figures</td>
<td>20</td>
</tr>
</tbody>
</table>
Introduction

Combinatorial control of gene expression refers to the interaction of several transcription factors in varying combinations to elicit a variety of different gene expression responses. Plants, due to their unique needs to respond to their environment, have evolved a highly diverse spectrum of transcription factor groups (5). Indeed, 5.9% of the Arabidopsis genome has been estimated to code for transcriptional regulators, compared to 4.5% of Drosophila, 3.5% of C. elegans, and 3.5% of Saccharomyces cerevisiae (6). In the study of combinatorial gene regulation, the plant flavonoid biosynthesis pathway has proven to be an excellent genetic model. Flavonoids are pigments that, in addition to coloring flowers and plants and thereby attracting pollinators, have functions in defense against pathogens and protection from UV light, and have also been suggested to take part in stress response and signaling. The key enzymes and regulatory genes of the flavonoid pathway are highly conserved across plant species, and it is this conservation, as well as the ease with which pigment accumulation phenotypes are scored for and the general non-lethality of flavonoid biosynthesis mutants, that has made the flavonoid biosynthesis pathway a classic genetic system. (1). Extensive information is therefore available from numerous biochemical and genetic studies of the pathway in several model systems including Zea mays and Arabidopsis thaliana (4).

Of the flavonoid pigments, the anthocyanins are the most common and broadly distributed, and are therefore the best studied and characterized. Anthocyanins are water-soluble, red-to-blue pigments which, in addition to attracting pollinators and protecting against UV light and pathogens, are powerful antioxidants, although the biological importance of their antioxidant properties in plant systems is uncertain as they are located in the vacuole (2,3). Anthocyanins are widely distributed across plant species, and the genes regulating their expression are largely
conserved among plants (7). In Maize, anthocyanins accumulate in the kernel, and in the vegetative and root tissues (1).

The anthocyanins are synthesized from a branch of the flavonoid biosynthetic pathway regulated by the R2R3 myb protein C1 or PL1 in interaction with a bHLH factor of the R/B family (Fig. 1) (8). The designation R2R3 myb designates a protein with a myb-type DNA binding domain consisting of an R2 and an R3 type 51 or 52 amino acid repeat, each of which encodes three α-helices that create a helix-turn-helix structure when bound to DNA (9). C1 and PL1, like some other myb proteins, also encode a C-terminal activation domain (10). C1 and PL1 are similar proteins but are expressed in different parts of the plant, with C1 being expressed in the kernel and in the husk, and PL1 being expressed in the plant body (1). Whereas the synthesis of anthocyanins is regulated by C1/PL1 and R/B, the branch of the flavonoid biosynthesis pathway that creates phlobaphenes requires only the R2R3 myb protein P1 (8). The two regulators may activate the same genes, such as \( A_l \), which encodes dihydroflavonol 4-reductase (DFR); they may also work separately, as with \( Bzl \), which encodes an UDP-glucose flavonoid 3-o-glucosyl transferase (UGFT) and is only activated by C1 in conjunction with R yet not by P (10,11).

The exact mode of action of the C1-R complex is not known; while C1 has been shown to bind DNA and to bind with R via its myb domain, R has also been shown to provide DNA contacts (whether directly or indirectly) and to protect C1 from a transcriptional repressor (12). The fact that R has not yet been shown to directly bind DNA (1), however, and the knowledge that it homodimerizes via a C-terminal ACT-like domain (13), suggest an alternative role for the highly conserved bHLH domain, perhaps in recruiting additional cofactors such as chromatin remodeling proteins for transcriptional activation.
A yeast two-hybrid screen using R as a bait identified the RIF1 (R Interacting Factor 1) protein as a potential binding partner for R (1). RIF1 is a 452 amino acid protein that includes two interesting domains: an Agenet domain, part of the Royal super-family of domains, many members of which have been shown to be involved in chromatin remodeling (19), and an ENT domain, which indicates homology to the human protein EMSY. Interestingly, EMSY has been shown to regulate expression of the human bHLH tumor suppressor protein BRCA2 by recruiting proteins containing Royal super-family domains (14). The presence of an Agenet domain in RIF1, and indeed most other plant Agenet proteins (of which there are 6-9), suggests that, whereas the human EMSY protein may recruit chromatin remodeling proteins to modify transcription, this function may be merged in plants (1).

While yeast two-hybrid experiments with an R construct with a mutation in its bHLH have suggested that R binds RIF1 via its bHLH domain, the region of RIF1 that R binds to has not yet been determined, although yeast two-hybrid experiments have shown that RIF1 can homodimerize through its ENT domain, like EMSY (1,15,16). R and RIF1 have also been shown to localize to the cell nucleus in the same speckled pattern (1). In addition, RIF1 was shown to be recruited to the A1 promoter in the presence of C1 and R but not in their absence (17). Given the evidence for the interaction of R with RIF1, the next question was whether RIF1, like R, is absolutely necessary for anthocyanin accumulation in maize, and what its method of action might be.

As creating knockout mutants in maize is a very difficult and time consuming process (see TILLING below), RNA interference (RNAi) was used to “knock down” RIF1. RNAi is a
technique in which double-stranded RNA (dsRNA) is created within the cell in order to trigger the cell’s own defense machinery, which will then destroy all mRNA homologous to the targeted gene. This resulted in a (presumed) “knock down” of RIF1 expression, which allowed the effect of RIF1 on anthocyanin accumulation to be easily quantified.

• Microparticle Bombardment

As maize and other monocots have proven difficult to transform with Agrobacterium tumefaciens, and as anthocyanin accumulation can not be observed in electroporated protoplasts, the system of transient microparticle bombardment assays in maize callus cells was used. In this system, undifferentiated callus cells in culture are bombarded with gold particles coated with DNA and then assayed for a biological response. This technique had previously been used to determine whether both C1 and R were both absolutely necessary for anthocyanin accumulation. In those experiments, a line of maize callus cells (Black Mexican Sweet, or BMS maize) that is unable to accumulate anthocyanins because it does not produce C1 or R was bombarded with constructs expressing C1 and R under a strong promoter. Whereas transgenic C1+R-expressing maize callus cells accumulate anthocyanins and are therefore red, BMS callus cells can not and are therefore white. This allowed anthocyanin accumulation to be easily assayed for simply by counting the number of red cells in the normally white culture (Fig. 3) and by normalizing the number of red cells to transformation efficiency by cobombarding with the Renilla luciferase reporter. This same system of transient microparticle bombardment assays in maize callus cells was used with the addition of the RIF1 dsRNA as an exogenously introduced construct, by comparing normalized red cell counts in cultures bombarded with R and C1 alone, or with R, C1 and RIF1 dsRNA.

• Luciferase reporter Assay
Cobombarding the maize callus cells with a construct expressing the luciferase reporter under the control of the *A1* promoter (PA1Luc) allowed the question of whether RIF1 was involved in chromatin remodeling to be explored. This was done by comparing the effect of the RIF1 dsRNA on *A1* promoter activation when the promoter DNA was in the chromatin state (as part of the endogenous, nucleosome-bound *A1* gene) and when the promoter DNA was in the non-chromatin state as (as part of the exogenously introduced PA1Luc DNA construct, which was not integrated into the genome or nucleosome-bound).

• Antibodies for ChIP

RIF1 protein has already been expressed and purified and sent for inoculation into rabbits to produce polyclonal antibodies against RIF1, and once these antibodies are received they will be used in ChIP experiments to identify additional target gene promoters that RIF1 may be associated with. In ChIP, or chromatin immunoprecipitation, the various transcription factors and other DNA-binding proteins at work in the cell are first fixed to the DNA they are active on. Next, antibodies against the particular DNA-binding protein of interest (here RIF1) are used to pull down the complex of the protein and the DNA it is bound to. This mixture can then be quantified by binding to a DNA microarray (ChIP-Chip) or by sequencing the attached DNA (ChIP-Seq). Additional target genes for RIF1 are of particular interest because examinations of knockout mutants for a RIF1 ortholog in *Arabidopsis* show a number of developmental defects presumptively unrelated to anthocyanin accumulation; thus RIF1 may be a part of transcriptional complexes regulating other genes, perhaps in unrelated pathways, in maize.

• Affinity Pull-down experiment

In order to find additional binding partners for RIF1, an affinity pull-down experiment will be done. In this experiment, binding partners for a protein are found by expressing a
construct with the DNA coding for the protein along with one or more N- or C- terminal affinity tags in planta and then purifying out the protein complexes via affinity purification and quantifying the bound proteins with mass spectroscopy. RIF1 will be infiltrated into Nicotiana benthamiana leaves in a construct including an N-terminal HIS tag and an IgG binding domain.

•TILLING

As it is problematic to obtain knockout lines in maize except by the arduous process of EMS mutagenesis, RIF1 will also be characterized by the quantification of knockout lines generated by the maize TILLING project (genome.purdue.edu/maizetilling). TILLING (Targeting Induced Local Lesions In Genomes) is a process allowing the identification of EMS-induced mutations in a specific gene region. Mutant seed lines for RIF1 retrieved from the TILLING project will be planted and characterized for phenotypic variation from normal maize plants.
Materials and Methods

•RNAi and Microparticle Bombardment

As no RIF1 knockout line of maize callus cells was available, an RNAi-based approach was used to determine whether RIF1 is absolutely required for anthocyanin formation. A 500 bp fragment of RIF1 (Fig. 2) was cloned in the forward and reverse orientations, separated by the rice waxy-a intron and driven by the powerful 35S CaMV promoter, into the maize binary vector pMCG161 (fig 4.), creating the construct p35SRNAi^{RIF1}. When transcribed in the cell, the mRNA created from this construct binds with itself to form a hairpin of dsRNA, which then induces the cells RNAi response. Black Mexican Sweet (BMS) maize callus cells were then bombarded with either a construct expressing C1 and R under the 35S CaMV promoter and the p35SRNAi^{RIF1} construct, or just the C1+R construct (a blank shot without either C1+R or p35SRNAi^{RIF1} was also performed, as well as a bombardment with C1+R and the empty vector from the p35SRNAi^{RIF1} construct). All of the shots also included a construct expressing the Renilla luciferase reporter under the control of the 35S CaMV promoter for normalizing the efficiency of the bombardment, and a construct expressing the firefly luciferase reporter under the control of the A1 promoter (pA1Luc). 1 µg of each regulator construct and 3 µg of each reporter construct were included used in each shot, and the total amount of DNA was adjusted to 10 µg with p35SBAR for each bombardment in order to equalize the amount of 35S promoter in each bombardment. Red cells were counted 36-48 hours after bombardment and were normalized to Renilla activity, measured with the Promega Dual Luciferase kit. All shots were performed in triplicate, and the entire experiment was repeated twice.

•Antibodies for ChIP
Although ChIP itself is a very challenging experiment, the production of antibodies specific to the protein of interest is very time consuming and typically takes months from start to finish. Antibodies for RIF1 were produced using the same protocol that is also being used to produce antibodies against other plant transcriptional regulators by the Grotewold lab for the Plant Genome Grant. In this protocol, a small region of the target protein is selected that is as unique to the target protein as possible, so as to reduce the possibility of nonspecific antibodies being made. An approximately 100 amino acid region is the smallest region from which polyclonal antibodies can be made using this technique, and the small size of the region helps in reducing homology to other, nontargeted proteins. For RIF1, the region chosen was between amino acids 106 and 241 (Fig. 2), a region stretching from the C-terminus of the Agenet domain to the N-terminus of the ENT domain. By avoiding the highly conserved protein domains, the antibodies will be specific for binding RIF1. The DNA for this region of RIF1 was cloned into the pDest17 IPTG inducible vector with a C-terminal HIS tag and the protein overexpressed in E. coli and the protein purified with a Nickel column. The protein was then run on an SDS-PAGE gel and gel purified, and is now ready to be injected into rabbits for inoculation. When the RIF1 antibodies are ready to be purified from the rabbit sera, a separate construct containing the targeted fragment of RIF1 attached to a GST tag will be expressed in E. coli and the protein bound to a glutathione column. The sera from the rabbits will then be passed through the column, resulting in only antibodies specifically targeting RIF1 (and not the HIS tag, for example) to be recovered. These antibodies will then be used in ChIP assays to identify additional DNA targets for RIF1.

• Affinity Pull-down
The vector used will be the pYL436 pC-TAPa vector (Fig. 6) (18), which includes a HIS tag and an IgG binding domain.

•TILLING

Submitting RIF1 for TILLING first required obtaining a genomic sequence for it and creating an Intron/Exon statement for the gene so that the TILLING project could accurately predict the effect of a mutation in the targeted area. In order to do this, the RIF1 cDNA sequence was blasted against recently sequenced maize BACs, and the highest scoring BAC (AC191252.2, which was mapped to chromosome 2, bin 6) was used to make an intron/exon statement. In addition to the intron/exon statement, primers also needed to be identified that the TILLING project could use to amplify the RIF1 gene out of genomic maize DNA. After these primers were ordered and tested in the lab to make sure that they could reliably amplify the desired region of genomic DNA without significant background amplification, the TILLING order was sent on May 18, 2007.
Results

• Transient Assays

In the transient assays done with microparticle bombardment of maize callus cells (Fig. 7), the p35SRNAi\textsuperscript{RIF1} construct was shown to very significantly ($P<0.01$) reduce the number of Renilla-normalized red cells induced by bombardment of the cells with C1 and R. However, no significant difference in anthocyanin accumulation was observed when cells were bombarded with the empty pMCG161 vector used to create the p35SRNAi\textsuperscript{RIF1} construct and C1+R. This confirms that RIF1 is indeed a vital component of the C1+R transcriptional regulatory complex active at the endogenous flavonoid pathway genes in maize. Whereas the p35SRNAi\textsuperscript{RIF1} construct did reduce the ability of C1 and R to regulate endogenous, chromatin-bound genes required for anthocyanin accumulation, it had no significant effect on the expression of the luciferase reporter gene bombarded into the cell under the control of the $A1$ promoter (PA1Luc). These results conform to a model in which RIF1 has chromatin remodeling functions, as its presence is required to regulate chromatin-bound genes, but not transiently induced, non nucleosome-bound DNA.

• Antibodies for ChIP

The fragment of RIF1 that will be used as an antigen to produce antibodies against was successfully expressed in E. coli, purified with a nickel column that selectively bound only His-tagged proteins such as the RIF1 antigen, and run out on an SDS-gel and gel purified. 600 µg of purified protein has been quantified and sent to Cocalico Biologicals, Inc. for inoculation into rabbits. The protocol used by Cocalico will include an initial injection of 100 µg of the antigen with complete Freund’s adjuvant (day 0), with booster injections of 50 µg of antigen with
incomplete Freund’s adjuvant at days 14, 21, and 49, and test bleeds at days 35 and 56, with the final number of boosts and bleeds depending upon the analysis of the test bleeds.

• TILLING

The first TILLING mutations in RIF1 were identified in January 2008 (Fig. 8). One mutation is in an intron, and therefore not predicted to affect the function of RIF1. The second mutation is a change of a guanine to an adenine at bp 671, resulting in amino acid 224 changing from a glycine to an aspartic acid; however, the SIFT score, which is 0.88 (above 0.05), indicates that the change is not likely to affect RIF1 function. Additional mutations will become available approximately every two months until a mutation with a greater than 95% chance of severely impairing RIF1 function is identified. Seeds from the two available mutations, however, will still be planted and analyzed in order to determine if any phenotypic changes are present.
Discussion

The results of the transient microparticle bombardment assays have shown that RIF1 is indeed a necessary component of the R+C1 transcriptional complex at the flavonoid biosynthetic pathway gene promoters. The fact that RIF1 was shown to be required for the regulation of chromatin-bound genes but not for the regulation of non chromatin-bound DNA also points to a prospective role in chromatin remodeling for RIF1. The identification of a new player in the R+C1 regulatory machinery provides new insight into the classical genetic system of flavonoid pigment production in maize, and is significant for the study of combinatorial transcriptional regulation in general. As mentioned previously, 5.9% of Arabidopsis genes are estimated to play a part in transcriptional regulation (a higher percentage than in Drosophila, C. elegans or yeast), and the study of transcriptional regulation in plants thus may yield important insights about the evolution of gene expression in all eukaryotes (6).

The fact that a knockout mutant of the Arabidopsis RIF1 ortholog ACK1 displays a number of developmental defects seemingly unrelated to the regulation of the flavonoid pathway genes suggests that RIF1 may be recruited to other transcriptional complexes at various other genes, perhaps in order to participate in chromatin remodeling. The creation of a knockout mutant for RIF1 with the maize TILLING project will allow anthocyanin accumulation to be observed in planta, and will also allow the characterization of any additional phenotypes (perhaps like those of the Arabidopsis ACK1 knockout) resulting from the loss of RIF1. In addition, ChIP experiments performed with antibodies currently in production against RIF1 should identify additional DNA targets of RIF1, and affinity pull-down experiments performed in Nicotiana benthamiana will identify binding partners for RIF1. The picture that emerges from the continued research of RIF1 and the maize flavonoid pathway genes, therefore, is one of a
highly interconnected genetic regulatory network that recruits a highly diverse set of cofactors, which work together in various combinations to regulate different transcription activities in a highly plastic manner.
Works Cited


Tables and Figures

Figure 1: The maize flavonoid biosynthetic pathway.

Structural genes are shown with the traditional maize nomenclature and the enzyme name in parenthesis: chalcone synthase (CHS), chalcone isomerase (CHI), dehydroflavonol reductase (DFR), flavonone 3 hydroxylase (F3H), Leucoanthocyanidin dioxygenase (LDOX), UDP-glucose flavonoid 3-o-glucosyl transferase (UGFT), and glutathione S-transferase (GST). The brackets indicate the branches regulated by either P1 or C1 + R (1).

Figure 2: The maize gene RIF1. The blue line represents the fragment used for RNAi (amino acids 55-221), and the red line represents the fragment used for antibody production (amino acids 106-241).
Figure 3: Maize Black Mexican Sweet (BMS) callus cell suspensions

The callus cell cultures shown above, because they are of the Black Mexican Sweet (BMS) line and do not express C1 or R constitutively, do not normally accumulate anthocyanins and are therefore white. When bombarded with C1 alone, no anthocyanins are created because R is absolutely required for anthocyanin accumulation in maize. When C1 and R are bombarded together, however, red cells expressing anthocyanins are visible (1).
Figure 4: The maize binary vector pMCG161.

Binary vectors can be expressed both in maize and in Agrobacterium tumefaciens, a bacteria often used to infiltrate DNA into plant system. The 500 bp fragment of RIF1 was cloned into the Ascl/AvrII site in the forward orientation, and into the Sgfl/Spel site in the reverse orientation, separated by the rice waxy-α intron. RNAi expression is under the control of the strong 35S Cauliflower Mosaic Virus (35S CaMV) promoter. (map at chromdb.org)
Figure 6: The construct that will by used for the affinity pull-down. RIF1 will be inserted into the Gateway cloning site (attR1::Cm\textsuperscript{r}::ccdB:attR2) of the pC-TAP\textsubscript{a} vector (18).

Figure 7: Results from transient microparticle bombardment assays in maize BMS cells, normalized to transformation efficiency with Renilla luciferase (17).
Figure 8: Results from TILLING of RIF1. A sift score of 1.0 means that there is a 0% chance of a deleterious change in protein function, whereas a score of 0% represents a 100% chance of such a change.