The Rho Family GTPase:

Determining GEF specificity through recombinant expression

Farrah Steinke

4/26/2010

GTPases are signaling proteins activated through the binding of guanine nucleotide-exchange factors, or GEFs, which exchange protein bound GDP for GTP. Expression of two novel GEFs in bacterial vectors and purification of the GEF proteins allows for the testing for Rho family GTPase activation by the GEFs. Early experimental obstacles impeded the successful cloning of the GEF genes into the plasmid; this prevented the expression of the GEF proteins. Without the protein product, the assay for GTPase identification was incomplete and the results inconclusive.

Table of Contents

Figures

1. Introduction

2. Materials and Methods

1. Introduction

1.1 Rho-family proteins and function

Rho-family proteins include GTPases, guanine nucleotide exchange factors (GEFs), and GTPase activating proteins (GAPs) all of which regulate the formation and rearrangement of the actin cytoskeleton. They are activated by various cell-surface receptors including tyrosine kinase receptors, cell adhesion receptors and G-protein-coupled receptors (GPCRs) [6]. Regulation of the actin cytoskeletal proteins implicates Rho-family proteins in cellular processes such as cell adhesion, phagocytosis, cytokinesis, and cell migration. [3].

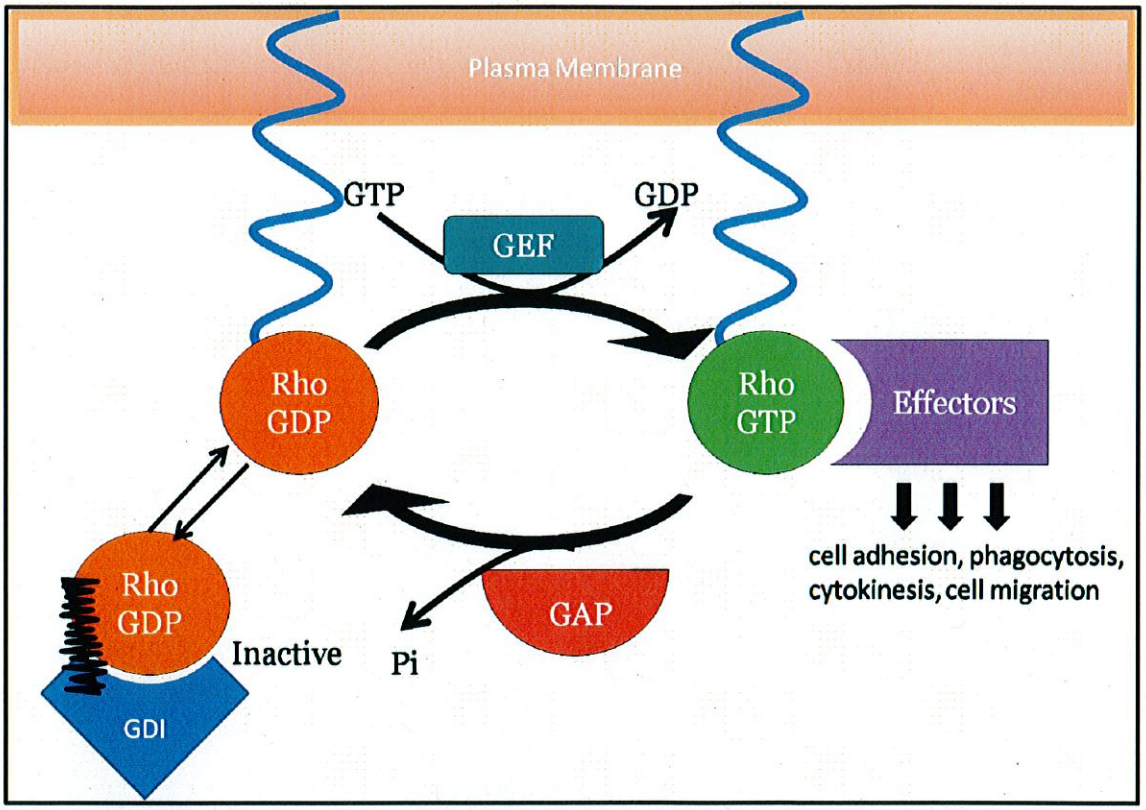
The Rho-family of GTPases is comprised of 22 small GTPase (~21 kDa) proteins [4]. Generally, GTPases are proteins that function as switches and respond to the binding of the nucleotides GDP or GTP. Conformational changes occur in the protein when GTP is exchanged for the bound GDP within the nucleotide binding pocket and activates the GTPase's signaling activity, producing downstream effects within the cell. The intrinsic function of the GTPase, hydrolysis of GTP to GDP, returns the GTPase to the inactive conformation and stops the signaling pathway.

1.2 GTPase regulation

As shown in figure 1, exchange of GTP for GDP is catalyzed by a Rho family GEF binding to the GTPase, increasing Rho-GTPase related activities such as those described above. Alternatively, the activity of Rho GTPases is suppressed by GAPs, which enhance the GTPase's hydrolysis of GTP to GDP, returning the GTPase to an inactive state [2]. A third type of GTPase regulatory protein is the guanine dissociation inhibitor (GDI). In figure 1 GDIs are shown as inhibitors of the dissociation of the GDP, preventing the association of the GTP and keeping the GTPase in the off conformation. The GDI also

prevents association of the GTPase with the plasma membrane through its lipid tail (shown as blue, curvy line in figure 1). Alternatively, the GDI can also prevent the intrinsic function of the GTPase to hydrolyze GTP bound nucleotide, maintaining the active conformation. [1].

Figure 1: Interaction of GEF, GAP, and GDI with GTPases

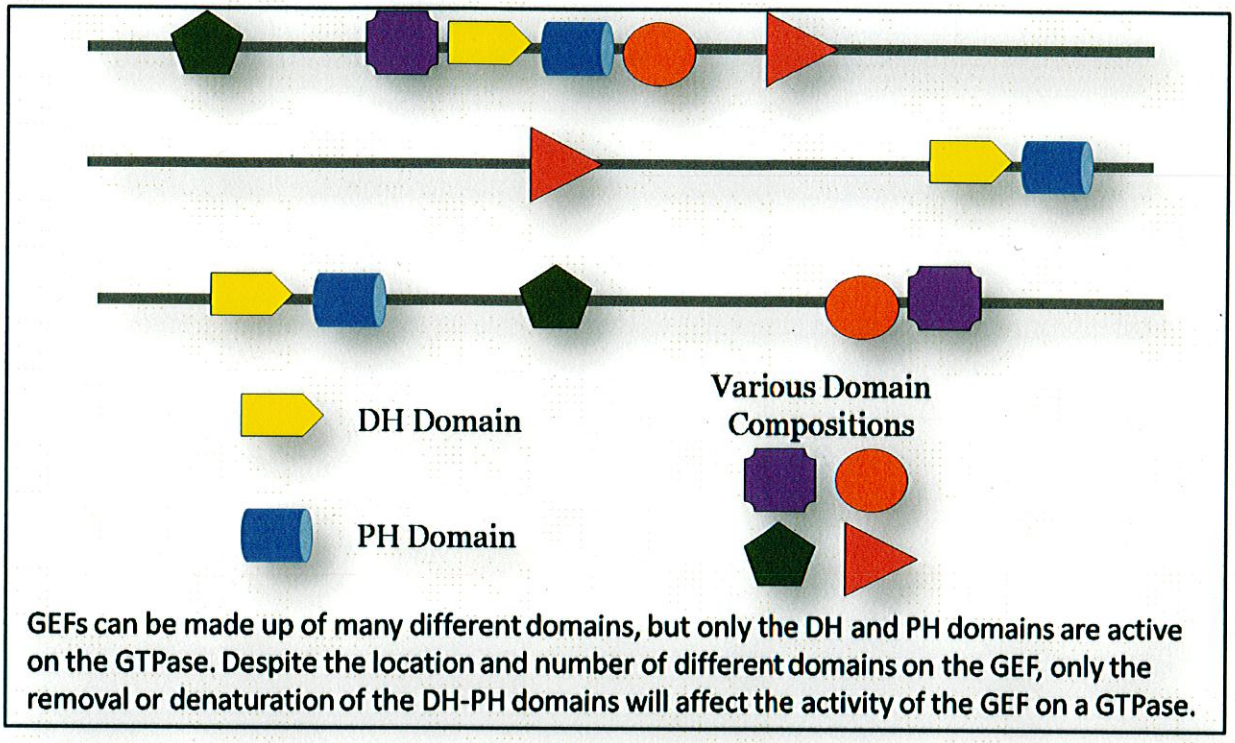


1.3 Rho-family GEF structure and function

Rho-family GEFs, the class of proteins that catalyzes the exchange of GTP for GDP and activates GTPase activity, contain two domains: the Dbl homology (DH) domain (about 200 amino acids) and the pleckstrin homology (PH) domain (about 100 amino acids) [4]. These two domains are typically only a small part of the entire GEF protein but are the portion responsible for the activation of GTPases. Regions outside

the tandem DH-PH domains do not pertain directly to GTPase activation, but to other cellular functions unique to the various Rho-family GEFs (see Figure 2).

Figure 2: GEF Protein Domain Schematic



The DH domain of the Rho-family GEF is comprised mostly of three conserved regions that interact directly with structural features on the GTPase called switch regions. Switch regions have different conformations dependent on whether guanosine *di*phosphate or guanosine *tri*phosphate is bound to the GTPase. At the interface between the GEF and the GTPase, significant physical changes occur to both the switch region and a Mg^{2+} binding pocket found on the GTPase; the conformational change of the GTPase removes the GDP and Mg^{2+} as well as exposes the nucleotide-binding site to the cytosol [4]. Because GTP is actually found in cells at higher concentrations than GDP

it is preferentially loaded into the GTPase but only when the nucleotide-binding pocket is empty.

The other critical domain for GEF-GTPase interaction is the PH domain, which may also physically participate in GTPase-binding with the GEF, but unlike the DH domain it has no catalytic activity towards a GTPase by itself. For some GEFs, the PH domain acts as a membrane anchor and also as a docking site for proteins that facilitate amplification of the GTPase's signal near the plasma membrane. Initially, the function of PH domains were thought to be limited exclusively to the binding of lipids, but studies have shown that PH domains act as protein docking sites which, when studied further, may provide answers to questions about specific signaling pathways and sites of GTPase activity [4].

Interestingly, removal or denaturation of the PH domain from a Rho GEF results in a significant decrease in GTPase activation; however, the catalytic activity of some GEFs is restored if the PH domain is re-established or replaced with another sequence that directs localization of the GEF to plasma membranes [4]. Questions arise if the localization of Dbl-family GEFs to the plasma membrane is due only to PH domain and phospholipid interaction alone; because of the PH domain's, apparent, but low binding affinity for phospholipids. The PH domain has been verified as a key part of the GEF structure and as a membrane anchor, but phospholipids alone are not what seem to significantly regulate GEF-GTPase interactions at the plasma membrane.

Phosphoinositides found in plasma membranes play a role in the orientation of both the PH domain as well as the GTPase. Conformations between the two DH and PH domains can be altered by the binding differences between amino acid residues on both the PH domain binding to the GTPase, and the position of a phosphate found on a

membrane-bound phosphoinositol. Phosphoinositides can expose a Rho-GTPase's binding surface, favoring interactions with the GEF and stabilizing the GTPase intermediate that forms when GDP and Mg^{2+} are removed [4].

1.4 *Identifying the GTPase specificity of Gef 5 and Ect 2*

So far, of the mammalian Rho family proteins, 22 GTPases and 69 GEFs have been identified, and many Rho GEF/GTPase pairs have been established. In many cases, one GEF can activate more than one GTPase, and alternatively, one GTPase can be activated by multiple GEFs. The GTPase specificity for each Rho GEF has not been entirely determined; and most of the analyses have been focused on the GTPases RhoA, Rac1 and Cdc42. Of the 69 GEFs in the Rho family, 48 have been tested for some sort of GTPase specificity, and most have only been shown to activate RhoA, Rac1, or Cdc42 or a combination of these three [4]. In most studies, these identifications have been made with GEFs that have been produced in bacteria through recombinant expression, purified, and placed in solution with individual recombinant GTPases, which can then be assayed for nucleotide exchange activity.

Two novel Rho family GEFs, herein called Gef 5 and Ect 2 (GenBank accession numbers AAH26778 and BAA91741) have been identified as DH-PH containing sequences, but lack any known GTPase specificity [4]. Genetic sequencing of the Gef 5 and Ect 2 proteins verifies the presence of the DH-PH domains grouping them in the Dbl family. Purified protein products of Gef 5 and Ect 2 are hypothesized to activate one or more of these three thoroughly studied GTPases, leading to the formation of different types of cytoskeletal structures signaled by the activation of these small G proteins.

2. Materials and Methods

2.1 *Identification of target domains of Gef 5 and Ect 2*

As mentioned in section 1.4, Gef 5 and Ect 2 have assigned accession numbers of AAH26778 and BAA91741, respectively. The coding sequences for these proteins have been fully sequenced, and DH-PH domains have been located within these sequences. Access to the genetic information of Gef 5 and Ect 2 was through GenBank. The entire amino acid sequence was provided and the residues that are included in the DH-PH domains are identified by numbers designated to each residue.

The coding, nucleotide sequence for the DH-PH domains were used to design oligonucleotide primers to PCR amplify the DH-PH domains of Gef 5 and Ect 2. Restriction enzyme cut sites were added to the primer sequences to facilitate cloning of the coding sequences into an appropriate expression plasmid (see below). The final sequences for the primers for Gef 5 and Ect 2 were:

Gef 5 DH-PH

Not1 (3') 5' – CGA CCG TGC AGC GGC CGC TTA CTG TTG CAG CTG GTT CTG GGC - 3'

HindIII (5') 5' –CGA CCG TGC AAA GCT TGA TGG GCA TGA GAA GCT GAC- 3'

Ect 2 DH-PH

Not1 (3') 5' –CGA CCG TGC AGC GGC CGC TCA CTT CTC CAT ACT GCT TTT G- 3'

HindIII (5') 5' –CGA CCG TGC AAA GCT TAA GGA CTC AGC AGA AAA GCG- 3'

The primers were used in the amplification of the target DNA sequence, which is outlined in more detail in the next section. The primers were designed to bind to the 5' end of the gene sequence for the DH domain and also bind to the 3' end of the PH domain to encapsulate and synthesize the target DNA. Additionally, the primers provide sites where restriction enzymes can cut the sequence encoding only these domains from a larger fragment of DNA in order to isolate the target DNA with significant purity and in high quantities.

2.2 *PCR*

The primers designed above were synthesized by IDT (Integrated DNA Technologies) and were used for amplification of the DNA sequences that code for Gef 5 and Ect 2 through the technique of polymerase chain reaction (PCR). For the PCR reaction primers at 30 pmol/μL, cDNA isolated from HeLa human cells, high fidelity DNA polymerase, and a PCR buffer that controls pH as well as contains the dNTPs, were added to thermalcycler tubes and placed in the thermalcycler. In order to amplify product DNA using PCR, the thermalcycler temperatures were initially set at 95° C for 30 sec, 55° C for 30 sec and 72° C for 2 min, and then varied from these temperatures to try and improve the yield of DNA product (which is discussed in section 4).

Verification of a successful DNA amplification was conducted through the use of gel electrophoresis. A 1% gel made from agarose and TBE (tris borate EDTA) was used to separate DNA fragments based on molecular weight. Larger fragments move more slowly through the gel than smaller fragments, and were compared to a molecular weight ladder. Syber safe dye was used for viewing bands of DNA in the gel under UV light. Only a small fraction of the PCR reaction from a given experiment was applied to agarose gels; the remaining sample was used for the next section.

2.3 *Purification of target DNA*

Following PCR amplification of the DH-PH target sequence, the target DNA was purified from all unused ingredients in the PCR reaction. The DNA was bound to a silica based resin and washed several times with an ethanol based wash solution. The resin was pelleted from the wash solution and the DNA was eluted from the resin with sterile TE.

The restriction enzymes NotI and HindIII were used to digest the target DNA at the restriction enzyme sites using Promega buffer E, which promotes optimal activity of both of the restriction enzymes, and stabilizes the purified PCR product. The digest was allowed to incubate at 37° C for 2 hours.

The purified and digested target DNA was isolated from the remaining DNA fragments and the restriction enzymes through agarose gel electrophoresis. TAE (tris acetate EDTA) was used as the running buffer and ethidium bromide as the DNA dye. Bands containing the target DNA were visualized under UV light and cut from the gel. The agarose gel slice was dissolved using 6M sodium iodide (NaI). Ground glass was used to bind the DNA as the NaI is washed away with an ethanol based wash solution. Sterile TE was added to keep the target DNA in solution for storage.

2.4 *Ligation and cloning*

For expressing the DH-PH regions of the Gef 5 and Ect 2 as proteins, the target DNA was inserted into the pMAL plasmid adjacent to a maltose binding protein (MBP) sequence to create a recombinant plasmid. The pMAL plasmid was cut with HindIII and NotI to create 'sticky ends' that complementarily base pair with the cut sites on the DH-PH target sequence. DNA eluted from the glass milk was added to ligation buffer, the cut pMAL plasmid, and T4 DNA ligase. After mixing thoroughly, the reaction was incubated at room temperature for 45 minutes before continuing to the bacterial transformation.

2.5 *Bacterial Transformation*

Bacterial transformation entails inducing a bacterial host, in this case E. coli (strain DH5 α), to take up the recombinant pMAL plasmid DNA. A small amount of the ligation reaction was added to the competent bacteria and incubated on ice for 15 minutes. The tube was then immediately placed in a 42° C water bath for 50 sec, then

immediately transferred back to the ice for an additional 2 minutes. After the final ice incubation, 0.5 mL of sterile LB media was added to the tube and incubated for 1 hr at 37° C. The bacteria were centrifuged to the bottom of the tube and 400 µL of the supernatant is removed. After re-suspending the bacteria in the remaining media it was spread on an agar plate containing ampicillin.

After incubating the agar plates for 24 hrs in a 37° C environment, individual colonies were selected, placed in glass tubes containing sterile LB media, and grown in a 37° C incubator, with shaking, over night. In order to detect for the presence of the recombinant plasmid, one mL of the bacteria containing media was placed in a microcentrifuge tube and spun down, and plasmid DNA was isolated using standard miniprep procedures. The isolated plasmid DNA was digested with NotI and HindIII restriction enzymes, and purified with the same procedure used after PCR amplification (section 2.2). The DNA was run through a 1% agarose gel containing TBE buffer, and syber safe dye; bands corresponding to the size of the target DNA and the larger plasmid fragment indicate successful transformation of the recombinant vector into the bacterial host.

2.6 *Protein purification*

The bacterial host does not produce the target fusion protein along with the various other proteins until a repressor protein located on the pMAL plasmid is removed. The repressor protein is bound to the operator and prevents the translation of the fusion protein, yet it can be removed with a chemical called IPTG. Isolation of the synthesized target protein is made possible by the fused MBP. MBP is easily synthesized by the vector, soluble, prevents aggregation of GEF's DH-PH domains, and is isolated using a semi-solid resin in a one step isolation.

Bacteria verified to contain the recombinant vector were treated with 100 μM IPTG and grown over night; 50 μL of the bacterial culture were placed in centrifuge tube and spun down. Supernatant liquid was removed; the bacteria pellet was re-suspended with MBP lysis buffer (20 mM Tris and 50 mM NaCl) and allowed to sit at room temperature for 30 minutes. The physical lysis of the bacterial cells is accomplished with the use of sound waves. Three, 1 minute cycles of sonication bursts open the cells, which were then spun in an ultracentrifuge for 15 minutes to pellet the bacterial residues, leaving the soluble, MBP fusion proteins in the supernatant liquid.

Amylose resin was placed in a separate tube and washed once with de-ionized water. The supernatant from the tube containing the sonicated bacteria was added to the amylose resin and incubated on ice for 30 min; this incubation ensures the binding of the MBP to the amylose. Centrifugation and removal of supernatant liquid follows, along with the addition of an ethanol based wash solution. After washing, the resin was spun down again and as much of the supernatant discarded as possible.

The MBP fusion protein, currently bound to the amylose resin, was eluted with a 500 mM solution of maltose and allowed to set for 1 min. Maltose competitively binds the site that the amylose is bound to on the MBP and releases the target protein from the semi-solid and back into solution. The liquid accumulates below the resin and was removed with a pipette and placed in a new microcentrifuge tube.

The purity of the target protein in the solution was determined with the use of SDS-PAGE. The target protein product along with a molecular weight marker and a solution containing un-fused MBP, to serve as a control, were loaded into the wells of the gel. Viewing of the bands formed on the gel was possible by dying the gel with coomassie blue.

2.7 Mant-GTP exchange assay

Mant-GTP (2'-O-(N-methylanthraniloyl) guanosine 5'-triphosphate) is a fluorescently tagged form of GTP. Fluorescence of the Mant-GTP occurs when it is loaded into a GTPase due to interaction of the fluorescent label with the tyrosine side chains that are oriented near the tag during binding. Addition of a GEF specific to the GTPase will enhance the fluorescence of the solution, which can be measured by a fluorescence detector.

Rho A, Rac1 and Cdc42 were the three GTPases of interest to be tested for activation by Gef 5 and Ect 2. A Mant-GTP exchange assay of these GTPases with the GEFs would determine the GTPase specificity of the GEFs. Three exchange reactions were prepped for each GEF containing: exchange reaction buffer (20 mM Tris, 50 mM NaCl and 5 mM $MgCl_2$), and equal parts of the Mant-GTP with one of the test GTPases (RhoA, Rac1, and Cdc42).

A reading was taken of each of the tubes without the addition of the purified GEF proteins and recorded. Immediately after adding each of the GEF proteins, a series of reading were taken and recorded. Minute changes in the level of fluorescence is expected without any GTPase specificity due to the high concentration of Mant-GTP, but this does not indicate GTPase activation by GEFs; whereas a dramatic increase in the level of fluorescence does indicate GTPase activation by the GEF proteins and identifies a specificity.

3. Results

Struggles with PCR amplification of the target DNA, the very first step, led to inconclusive results, and the experiment did not make it past section 2.4. After the

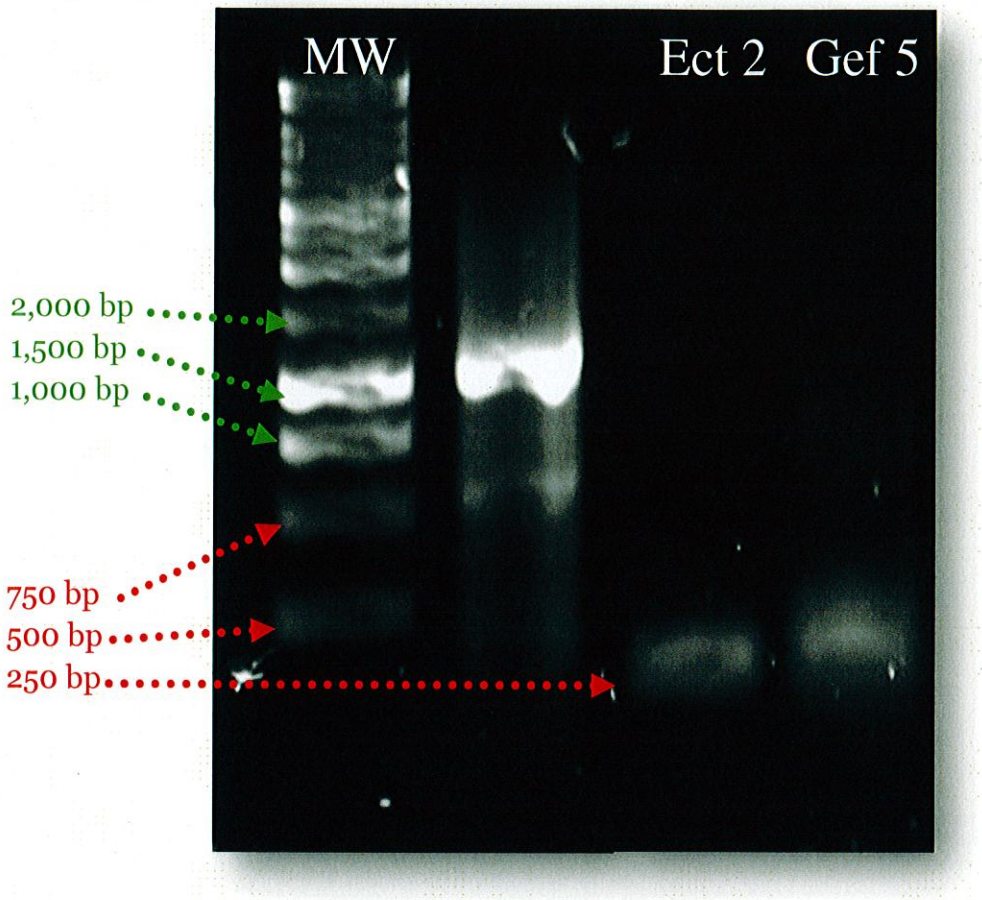
bacterial transformation, electrophoresis of DNA isolated from the bacteria showed negative results for the presence of any recombinant plasmid within the bacteria.

Figure 3 is an image of a gel electrophoresis run after one of the many initial attempts for PCR amplification. In the first lane, labeled MW, the molecular weight ladder is shown; several bands within ladder are labeled with nucleotide base pairs (bps) for size comparisons. The second lane contains successfully amplified DNA from a control DH-PH sequence, which produced bright, heavy bands between the 1,000 and 2,000 bp bands of the ladder. Lanes 3 and 4 contain the PCR products of Ect 2 and Gef 5; with this gel there is clearly a lack of any bands of the appropriate size, indicating no product was amplified. Ect 2 and Gef 5 both produced bands that are located very near the bottom of the gel signifying that the DNA sequence that was amplified is far too short. As mentioned in section 1.3, the DH-PH domains are made up of about 300 amino acids total; this equivocates to at least 900 nucleotide base pairs. The dim, wide bands found between 500 and 250 bp indicate that various, undersized sequences were produced in small quantities, and the target DNA was not successfully amplified.

The PCR procedure was repeated several times and thermalcycler temperatures were altered to improve the primer's ability to properly adhere to the target DNA and facilitate amplification. After using fairly extreme temperatures for the PCR, faint bands were detected on the gels within the molecular weight range correlating to the target sequences. Although the bands indicated that only a small amount of target DNA was present. Regardless of the relatively small amount of PCR product, this DNA was purified and ligated into the pMAL plasmid, and subsequently used for bacterial transformation. After growing several bacterial colonies, minipreping, and analyzing the

plasmid DNA by gel electrophoresis, it was determined that the target DNA did not ligate into pMAL plasmid in these instances

Figure 3: Gel electrophoresis analysis



4. Discussion

The false impression of successful amplification of target DNA could have been an effect of the extreme temperatures used. With a low enough annealing temperature, the primers may have adhered to the DNA in locations that did not exactly match the base pairing. Forced synthesis of a DNA sequence that was not the target sequence, but similar in size may have occurred. As was mentioned in section 3, after attempting PCR with various temperatures, some faint bands on a gel showed up in the correct size region, but the amount of DNA required for the success during the ligation procedure

was never synthesized. Analysis of the DNA found in the transformed bacteria with gel electrophoresis demonstrated that the target sequence was not ligated into the pMAL vector.

There are many possibilities as to why the target sequence did not amplify, and one is that the target sequence was not present in the cDNA pool. It is possible that the HeLa cells from which the cDNA was extracted did not code for either Gef 5 or Ect 2. The two GEFs are very similar, genetically, and if the sequence for one wasn't present, then it is likely that the sequence for the other would not be present either.

Annealing temperatures of the primers to the target sequence can be optimized based on the A-T and G-C content. Potentially, the proper annealing temperature was never employed. Without adhesion of the primers to the appropriate sequence, the DNA polymerase cannot synthesize new strands of DNA. Also, excessive G-C, content could make the synthesis of DNA more difficult for the polymerase, and complementary base pairing of the target sequence within itself, could both have posed problems. The DNA polymerase can easily be dislodged from its track when encountering one of these hairpins, halting the amplification process. Further analysis of the sequence would determine if these problems were possibilities.

The GTPase specificity of Gef 5 and Ect 2 remains to be determined. If solutions to the problems with DNA amplification were to be found, the rest of the experiment would probably have been successful. The bacterial vector used had successfully taken up other recombinant plasmids and produced the target MBP-GEF fusion protein. In addition, the proteolytic removal of the MBP from the GEF after the purification step may be a good ancillary step which would ensure that the large MBP does not interfere with the binding of the GTPase with the GEF during a Mant-GTP assay.

5. References

[1] DerMardirossian, Celine, and Gary Bokoch. "GDIs: Central Regulatory Molecules in Rho GTPase Activation." *TRENDS in Cell Biology* 15.7 (2005): 356-63. *Science Direct*. Web. Mar. 2010.

[2] Gamblin, Steven J., and Stephen J. Smerdon. "GTPase-activating Proteins and Their Complexes." *Current Opinion in Structural Biology* (1998): 195-201. *Science Direct*. Web. Mar. 2010.

[3] Moon, Sun Young, and Yi Zheng. "Rho GTPase-activating Proteins in Cell Regualtion." *TRENDS in Cell Biology* 13.1 (2003): 13-22. *Science Direct*. Web. Mar. 2010.

[4] Rossman, Kent L., Channing J. Der, and John Sondek. "GEF Means Go: Turning on Rho GTPases with Guanine Nucleotide-Exchange Factors." *Nature Reviews* 6 (2005): 167-80. Print.

[5] Shan, Zhihong, Thomas Haaf, and Nicholas C. Popescu. "Identificaton and Characterization of a Gene Encoding a Putative Mouse Rho GTPase Activating Proein Gene 8 Arhgap8." *Gene* 303 (2003): 55-61. *Science Direct*. Web. Mar. 2010.

[6] Sinha, Soniya, and Wannian Yang. "Cellular Signaling for Activation of Rho GTPase Cdc42." *Cellular Signaling* 20 (2008): 1927-934. *Science Direct*. Web. Mar. 2010.

[7] Tapon, Nicolas, and Alan Hall. "Rho, Rac, and Cdc42 GTPases Regulate the Organization of the Actin Cytoskeleton." *Current Opinion in Structural Biology* 9 (1997): 86-92. *Science Direct*. Web. Mar. 2010.