Project Title: Development of Improved Chemicals and Plastics from Oilseeds

Award Number: DE-FC36-01ID14213

Recipient: The Dow Chemical Company

Project Location: Midland, MI and San Diego, CA

Reporting Period: Final Technical Report

Date of Report: July 31, 2006

Written by: Patricia Nugent

Program manager: Patricia Nugent

Principle Investigators: Zenon Lysenko, Paul Roessler

Subcontractors: USDA-ARS Western Regional Research Center (WRRC); Castor Oil, Inc. (COI)

Cost-Sharing Partners: None

Project Contact(s): Patricia Nugent (PD), 989-636-5181, panugent@dow.com

Robert Falardeau (BO), 989-636-9047, erfalardeau@dow.com

DOE Project Team:

DOE-HQ contact, Valerie Sarisky-Reed

DOE Field Project Officer, Fred Gerdeman

DOE Contract Specialist, Margo Gorin

DOE Project Monitor, Bryna Berendzen

DOE Analysis:

DISCLAIMER

Development of Improved Chemicals and Plastics from Oilseeds

Executive Summary

The overall objective of this program was to develop technology that can be applied to the production of various chemicals and plastics from seed oils. This research and development program included activities in all four key barrier areas identified in the US DOE Technology Roadmap for Plant/Crop-Based Renewable Resources, namely Plant Science, Production, Processing, and Utilization. Participants in the project included The Dow Chemical Company, Castor Oil, Inc., and the USDA Western Regional Research Center (WRRC).

- Plant Science research involved the isolation and characterization of genes and enzymes involved in seed oil metabolism in the industrial oilseed crop *Ricinus communis* (castor), along with the development of tools that will facilitate the production of improved castor varieties through recombinant DNA technologies.
- Production research included efforts directed at improving the agronomics and agricultural practices associated with castor, including analysis and breeding of cultivars, harvesting, and oil processing.
- Processing research entailed the use and development of catalysts and processes useful for converting various seed oils to monomers that can be utilized in several different applications. The main chemical conversion process studied, metathesis, provides two co-products: terminally functionalized fatty acids and linear alpha olefins. These materials are useful for formulating a variety of plastics, foams, coatings, etc.
- Utilization research was conducted to determine market opportunities for the various seed oil derivatives envisioned using the conversion processes developed in this project. In addition, a Life Cycle Analysis (LCA) was conducted to assess the sustainability of several representative processes and products derived from seed oils.

Additional details for each of these research areas are provided below.

Plant Science: In the area of Plant Science, Tasks 1.1 and 1.2 were successfully completed and all milestones associated with these tasks were achieved. In the genomics area, approximately 15,000 independent castor seed cDNA clones were processed for high throughput sequencing, which led to the determination of over 3300 unique (non-overlapping) sequences. The predicted gene products (*i.e.*, proteins) were annotated based on similarity to protein sequences in the public databases, leading to the identification of numerous genes involved in oil metabolism and general carbon metabolism; both of these areas of metabolism are extremely relevant to the genetic engineering of castor for the production of industrially-important oils.

In addition to the castor gene identification and annotation work, we isolated numerous gene promoters from castor that can potentially be used for the genetic engineering of castor and other crop plants. Several of these promoters were tested, some of which showed highest activity in seeds and others which exhibited constitutive activity or higher activity in vegetative (leaf) tissue.

Tasks 1.4 to 1.7 were the tasks WRRC worked on in the study. For the castor transformation task, the primary achievement of the work was generation of treatments that allow multiple shoot regeneration from explants derived from different parts of the plant,

including shoots and mature embryos. WRRC solved both halves needed to achieve the goal of castor transformation, but were not able to make both halves work together.

The WRRC group cloned genes for a two important types of enzymes involved in lipid biosynthesis, a diacylglycerol acyltrasferase (Type 1 DGAT) and several acyl-CoA synthetase (ACS) isoforms. DGAT catalyze the final acylation of diacylglycerol (DG) to form triacylglycerol (TG), the main component of seed oil. ACS enzymes produce essential metabolites (acyl-CoAs) involved in acyl transfer and beta-oxidation of fatty acids. Two of the cloned genes expressed enzymatically active proteins in a yeast expression system. Together, the ACS and the Type 1 DGAT displayed a six-fold preference for utilizing ricinoleate versus oleate for TG synthesis. This was similar to the preference observed for castor microsomes incorporating ricinoleate into TG versus oleate.

Production: Various castor lines were tested in order to identify lines exhibiting favorable traits.   The identified desirable traits included drought tolerance, high oil content, low ricin content, uniform seed size, pest resistance, disease resistance, non-dehiscence, uniform crop morphology, and higher yields. Several hundred lines of castor were tested by planting isolated plots of each line, identifying the presence or absence of an identified trait, harvesting the seed, and extracting the oil.

The crosses developed during the project are still under evaluation to determine if there is a yield advantage, as well as to identify other desirable agronomic attributes. Some of the lines exhibited a 15% yield advantage over open-pollinated varieties in initial trials, but more trials are needed to confirm the yield data since other factors can affect the yield data, including the presence of weeds that compete for water. Furthermore, it was shown that the pH of the soil needed be between 6.0 and 7.5 to ensure good seed yields. An additional factor that was shown to play a role in determining yields was maintenance of the proper level of usable nutrients so that the plants did not remain in a vegetative state.

Although castor breeding programs cannot be completed in four years (the duration of the grant period), progress was made in this project. Most of the work that has been carried out in the past with castor germplasm improvement programs has focused on collecting wild castor seed from different parts of the world as opposed to true breeding operations.   We were fortunate enough to recognize and know other true castor breeders, so we have been able to collaborate with them. These breeders (Dr. D. L. Van Horn, Dr. Raymond Brigham, and Dr. Harold Muller) have been invaluable.

Castor has great potential to be a very profitable crop for farmers in certain regions of the country, once some of the few remaining problems are eliminated. Not all of these problems can be solved in four years, but because of this opportunity we have been able to gain a better understanding of the obstacles that must be overcome, and have made significant progress that will facilitate the reintroduction of castor into the United States.

Processing: The objective of this task was to evaluate and develop metathesis catalyst technology as a means of utilizing seed oils as feedstocks for the chemical industry. Specifically, ethenolysis of fatty acid methyl esters, FAME's, leads to α,ω-functionalized derivatives. These serve as valuable starting points for materials which cascade into a variety of applications, many of which Dow has a market presence in. The relatively recent discovery and commercial availability of a family of metathesis catalysts by Grubbs *et al.* which are

tolerant of polar functional groups and the acquisition and implementation of high throughput synthesis and screening infrastructure at Dow led to a prime opportunity to investigate this project area.

The course of this project at Dow had four major thrust areas:

(1)   Evaluation of the Grubbs family of catalysts for carrying out the FAME conversion.

(2)   Development of the Grubbs catalyst for optimal performance.

(3)   Process developments for optimal catalyst and economic conversion.

(4)   New catalyst discovery for accomplishing the conversion.

Key discoveries and findings include:

(1)   Determination of kinetic models for direct ethenolysis using the two commercial Grubbs catalysts. Kinetic modeling results include:

   a. Kinetic parameters and rate laws as a function of key process variables.

   b. Catalyst lifetime determination.

   c. The magnitude and importance of product inhibition.

(2)   Determination of the relative impact of various feed impurities on catalyst performance.

(3)   Discovery of feed composition requirements and a means of achieving this quality from a crude, industrially relevant FAME stream.

(4)   Determination of the structure activity relationships for the Grubbs family of catalysts as they relate to direct ethenolysis catalyst performance using high-throughput methodology.

(5)   Demonstration of catalysts recovery/recycle using selective membrane technology.

(6)   Discovery of indirect ethenolysis as an improved conversion process for methyl oleate ethenolysis.

.

Utilization: Dow compared the sustainability of flexible foam polyols made through conventional petrochemical routes to a product with the same performance and attributes but made largely from renewable materials (castor oil or soy oil) by doing a partial LCA that follows closely the ISO standards. Three basic routes were compared. The calculated life cycle inventories (energy, mass & water) and impacts (greenhouse gas and acid gas emissions) for the three compared routes to make a functionally identical polymer solution for flexible foam polyurethanes were determined. The three routes included: a petrochemical route (APME), soy oil route, and a castor oil route, based on irrigated farming in Texas and production of useful fertilizer and oil from castor seeds (no low value or useless co-product).

NBID Associates was contracted to explore the commercial viability of a new oleochemical-based monomer, methyl decenoate, and the co-product decene. In fields

experiencing strong discontinuities, they looked for new business opportunities in terms of market openings and needs, what was required for competitive advantage and a business model for capturing value.

Comparison of Accomplishments with Goals of Project

Plant Science: This task identified seven different subtasks for deliverables in the project. Of the seven, **Tasks 1.1** and **1.2** were successfully completed and all milestones associated with these tasks were achieved. In the genomics area, approximately 15,000 independent castor seed cDNA clones were processed for high throughput sequencing, which led to the determination of over 3300 unique (non-overlapping) sequences. For **Task 1.3**, we identified a limited set of genes (~200) expressed in castor seed that are known to be involved in intermediary metabolism and in the metabolism of lipids and storage proteins. Genes involved in the following categories of lipid metabolism were cataloged: fatty acid biosynthesis, membrane lipid biosynthesis (both in plastids and endoplasmic reticulum), storage oil biosynthesis, wax and cutin biosynthesis, lipid and fatty acid degradation, and lipid signaling. In addition, representative genes from the following categories were identified: amino acid biosynthesis and degradation, carbohydrate biosynthesis and degradation, nucleoside biosynthesis and degradation, vitamin/cofactor biosynthesis and degradation, and secondary metabolism. The design of gene-specific oligonucleotides for the microarray had been initiated as part of this task, but a reduction in funding for the project resulted in halting additional activities for this task.

Progress was made on **Task 1.4** castor transformation and progress on **Task 1.5** was minor due to lack of progress on castor transformation. The preference of a Type 1 DGAT for DG containing hydroxy fatty acids (FAs) is the subject of a patent application in Task 1.6 and the work on **Task 1.6** was shifted to identification of biochemical roadblocks to facilitate incorporation of target fatty acids into triglycerides at relevant levels. **Task 1.7** required the castor transformation system development to be competed.

Production: **Task 2.1** was completed with over 200 lines of castor tested. **Task 2.2** was completed and four or five good high yielding hybrids were identified. For **Task 2.3** on improved harvest technology, COI built a harvester that would remove only the capsule and hull from the plant and not take the whole plant through the combine. In **Task 2.4,** improvements were made in oil extraction and progress was made on the identification of new uses for the meal.

Processing: Task 3.1 identified a series of potential impurities in the methyl oleate feed. The impurities were examined for their impact on catalyst performance and several methods were developed to address removing the impurities. **Task 3.2** began work on improving the Grubbs catalyst. During the course of the project intellectual property issues arose due to the partnership agreement formed between Materia and Cargill to develop Grubbs catalyst for the conversion of seed oils to chemical feedstocks. At this point, no more Materia catalyst was commercially available to Dow and a research license was not granted. This situation, coupled with our inability to improve Grubbs catalytic performance for the direct ethenolysis of methyl

oleate, caused a shift of our research direction away from these systems toward the discovery of non-Ru based olefin metathesis catalysts. The current state of the new catalyst effort, however, is that no catalyst system has been identified which can compete with the Grubbs catalyst for the ethenolysis, direct or indirect, of methyl oleate. **Task 3.3** was completed successfully on 2 separate occasions and **Task 3.4** was evaluated as detailed in the task and it was determined that heterogeneous catalyst development would be discontinued. **Task 3.5** and **Task 3.6** included economic evaluation and reaction engineering. The completion of the 2 tasks determined that the catalyst turnover for the 1st reactor of this new process was 70,000 on a consistent basis. This approaches the targeted catalyst performance we originally established for commercial viability.

Utilization: Task 4.1 and 4.2 were completed with a combined modeling and life cycle analysis comparing polyol production from 3 different feedstocks. **Task 4.3** was completed with an external market study and initial validation with customers was completed.

Task 1 Plant Science: Castor Genomics, Biochemistry, and Metabolic Engineering

The castor plant has the potential to be an excellent host plant for the production of a variety of industrial oils, including hydroxy fatty acids and epoxy fatty acids, *via* the use of recombinant DNA technologies. There are two primary reasons that castor would likely be able to fulfill this role: 1) the lipid metabolizing biochemical machinery from castor is uniquely adapted for producing and utilizing oxygen-functionalized fatty acids, and 2) castor is grown exclusively as an industrial crop, and therefore it is extremely unlikely that castor oil and food oil would become co-mingled.

In order to produce different fatty acids in castor, several technological advancements must be realized. In particular, the amount of genetic information about this crop needs to be greatly expanded and a robust and efficient genetic transformation system must be developed. The overarching objective of Task 1 was to expand the technologies available for the genetic manipulation of castor plants.

DOW

1.1 - Castor seed EST sequencing and bioinformatics

Of particular importance in expanding the genetic toolbox available to castor oilseed researchers is the availability of DNA sequences of castor genes that relate to carbon metabolism and seed oil production. It is also important to understand the regulation of these genes and the relative abundance of particular gene transcripts. These goals were met by the successful completion of a large-scale castor EST sequencing project. (Note: randomly chosen cDNA sequences are commonly referred to as "ESTs," or "Expressed Sequence Tags.")

Total RNA isolated from castor endosperm tissue harvested at various stages of development was provided by the McKeon lab at the USDA ARS Western Regional Research Center. The castor variety used, Accession PI 215769, was originally obtained from the USDA ARS Southern Regional Plant Introduction Station, Griffin, GA. A sample of the RNA was provided to Invitrogen Corp. for mRNA isolation and cDNA library construction. The library was produced in the plasmid vector pCMV-SPORT6.1, using *E. coli* DH10B-TONA host cells, such that each recombinant *E. coli* cell contained an independent cDNA clone.

After comparing several potential vendors through the use of trial sequencing runs to assess sequence quality and cost attributes, we decided to utilize the Michigan State University Genomics Technology Support Facility (GTSF) (http://genomics.msu.edu/index.html) for high throughput sequencing. A preliminary sequencing project involving ~500 independent clones was carried out in order to identify the most highly expressed genes; clones corresponding to these abundantly expressed genes could then be bypassed when selecting clones to sequence, improving the total genome coverage. The preliminary sequencing data indicated that approximately 40-50% of the clones in the library coded for various seed storage proteins. In order to identify and avoid these abundant clones during the full-scale sequencing effort, individual colonies from the library were robotically transferred from agar plates to 96-well (deep-well) primary growth plates containing LB medium. After overnight growth, a small aliquot from each well was spotted onto nylon membranes in ordered arrays ("macroarrays"). Next, labeled nucleic acid hybridization probes corresponding to the 10 most abundant cDNA sequences were produced via the polymerase chain reaction (PCR) and used to probe the macroarray membranes. An automated optical detection system was then used to relay the

coordinates of non-hybridizing clones to a computer-controlled liquid sampling system that removed aliquots from the appropriate wells of the primary deep-well plates and inoculated new cultures in secondary deep-well plates for clone propagation. Next, plasmids were isolated from each secondary clonal culture in order to provide sequencing templates. Sequencing reactions were performed using "Big-Dye Terminator" cycle sequencing technology (Applied Biosystems, Inc.). High throughput sequencing was then performed on ABI 3700 and ABI 3730 automated capillary sequencer units (Applied Biosystems, Inc.). A total of approximately 15,000 clones were processed in this manner.

The initial analysis of raw data generated *via* automated sequencing involved three primary steps: 1) nucleotide base calling, 2) assembly of the EST sequences into overlapping contiguous stretches of DNA ("contigs"), and 3) draft annotation of the contigs and non-overlapping "singletons" by comparison with known gene and protein sequences in the public genetic databases (*e.g.*, GenBank and UniProt). Base calling was done by computer analysis of the sequencer chromatogram output files using the Phred program (Ewing and Green 1998) and assembly of contigs was accomplished by using the Phrap program (http://www.phrap.org/phredphrapconsed.html#block_phrap). The entire set of contigs and singletons was then converted into a format that could be analyzed by the "GCG Wisconsin Package" set of programs included in SeqLab (Accelrys, Inc.). The final dataset included 1465 contigs and 1909 singletons, representing a total of 7248 independent clones.

The EST sequences generated in this project were annotated via homology comparisons with genes and proteins in the public databases. Specifically, the BLASTX program (Gish and States 1993) was used to compare the amino acid sequences deduced through translation of each EST sequence in all six possible reading frames with the UniProt protein database, release 1.8 (Wu *et al.* 2006) and also against the predicted protein sequences encoded by the entire genome of the model plant *Arabidopsis thaliana*, which were downloaded from the TAIR ("The Arabidopsis Information Resource") website (http://www.arabidopsis.org/home.html). The results of these analyses were imported into a Microsoft Access database, which enabled searching for specific text strings in the "hit" description fields (*e.g.*, "lipid," "acyl," "storage," etc.). The database also contained the actual DNA sequences of both contigs and singletons, a listing of the various independent clones that constitute each contig, and the the BLASTX-generated pairwise comparison output for each unique sequence in the database.

As a result of this EST sequencing project, a large number of genes that encode proteins involved in various metabolic pathways in developing castor seeds were identified. We were particularly interested in enzymes involved in lipid metabolism, and Table 1-1 includes a list of genes associated with various aspects of plant lipid metabolism for which homologs were present in the castor EST dataset.   A number of the contigs produced through this EST project represented full-length clones.

Table 1-1. Genes involved in plant lipid metabolism with significant homology to castor EST sequences.

Acetyl-CoA carboxylase, a-carboxyltransferase	Methylcrotonyl-CoA Carboxylase, biotinylated subunit
Acetyl-CoA carboxylase, b-carboxyltransferase	Mitochondrial Acyl Carrier Protein
Acetyl-CoA carboxylase, biotin carboxyl carrier protein	Mitochondrial Dihydrolipoamide Dehydrogenase,
Acetyl-CoA Synthetase	Mitochondrial Enoyl-CoA Hydratase
Acyl-ACP Thioesterase FatA	Mitochondrial Glycerol-3-Phosphate Dehydrogenase
Acyl-ACP Thioesterase FatB	Mitochondrial Ketoacyl-ACP Reductase
Acyl-CoA : Diacylglycerol Acyltransferase	Mitochondrial Ketoacyl-ACP Synthase
Acyl-CoA Binding Protein	Mitochondrial Lipoate Synthase
Acyl-CoA Binding Protein	Mitochondrial Lipoyltransferase
Acyl-CoA Binding Protein	Monoacylglycerol Lipase
Acyltransferase	Monogalactosyldiacylglycerol Desaturase (palmitate-specific, FAD5)
Acyl-CoA Oxidase	NAD+ Oxidoreductase (involved in FA alpha-oxidation)
alpha-Ketoacid Decarboxylase E1 beta subunit	Non specific Phospholipase C
ATP Citrate Lyase A subunit	Oil-body Oleosin
ATP Citrate Lyase B subunit	Oleate Hydroxylase
CER1 Protein involved in wax synthesis	Patatin-like Acyl-Hydrolase
Ceramide Sphingobase delta-8 Desaturase	Peroxisomal Long-Chain Acyl-CoA Synthetase
CTP : Ethanolaminephosphate Cytidylyltransferase	Phosphatidylcholine : Diacylglycerol Acyltransferase
Cytosolic Homomeric Acetyl-CoA Carboxylase	Phosphatidylinositol Phosphate Kinase type I/II B
Cytosolic Phospholipase A2	Phosphatidylinositol Phosphate Kinase type III
DAD1-like Acylhydrolase	Phosphatidylinositol-4-Kinase gamma
Diacylglycerol Kinase	Phosphoethanolamine N-Methyltransferase
Dihydrolipoamide Transacylase	Phospholipase A2-activating Protein (activity not documented in plants)
Epoxide Hydrolase	Phospholipase D beta
ER 2-Lysophosphatidate Acyltransferase (LPAAT)	Phospholipid : Acyl acceptor Acyltransferase
ER CDP-Diacylglycerol Synthetase	Plastidial Acyl Carrier Protein
ER Oleate Desaturase	Plastidial Dihydrolipoamide Dehydrogenase, pyruvate DH complex
Fatty Acid Alcohol Oxidase	Plastidial Dihydroxyacetone-Phosphate Reductase
Fatty Acid Amide Hydrolase	Plastidial Enoyl-ACP Reductase
Fatty Acid Omega-Hydroxylase	Plastidial Ketoacyl-ACP Reductase
Isovaleryl-CoA Dehydrogenase	Plastidial Linoleate Desaturase (FAD7/FAD8)
Jasmonic Acid Carboxyl Methyltransferase	Plastidial Lipoxygenase
Ketoacyl-ACP Synthase I	Plastidial Lipoyltransferase
Ketoacyl-ACP Synthase II	Plastidial Long-Chain Acyl-CoA Synthetase
Ketoacyl-ACP Synthase III	Plastidial Oleate Desaturase (FAD6)
Ketoacyl-CoA thiolase	Plastidial Pyruvate Dehydrogenase E1alpha subunit
Lipid Acylhydrolase-like	Plastidial Pyruvate Dehydrogenase E1beta subunit
Lipid Transfer Protein type 1	Pollen-surface Oleosin
Lipid Transfer Protein type 2	PTEN-like Phosphoinositide 3-Phosphatase
Lipid Transfer Protein type 3	Putative Transcription Factor CER2 involved in wax biosynthesis
Lipid Transfer Protein type 5	Serine Palmitoyltransferase (LCB2)
Lipid Transfer Protein type 6	Sphingobase 1-Phosphate Lyase
Lipid Transfer Protein type 7	Stearoyl-ACP Desaturase
Long-Chain Acyl-CoA Synthetase	Sulfolipid Synthase
Lysophospholipase	Triacylglycerol Lipase
Malonyl-CoA : ACP Malonyltransferase	

1.2 – Full-length gene and promoter cloning

The availability of several gene promoters that function in the host plant of interest is a critical need for the recombinant modification of crops. Because we are interested in producing unique varieties of castor through genetic engineering, a major goal of Task 1 was the isolation of native gene promoters from this crop.

For oilseed genetic engineering, it is desirable to have a suite of promoters that drive both constitutive expression (*i.e.*, expression in all tissues at all times) and seed-specific expression (*i.e.*, expression primarily in seed tissues). Constitutive expression is useful when it is desired to introduce a trait that is manifest throughout the plant (*e.g.*, herbicide tolerance or

insect resistance) and is useful to drive the expression of the selectable marker used to identify transformed plants. Seed-specific promoters are often used to modify the properties of seed oils by driving the expression of enzymes involved in lipid synthesis, especially in cases where production of those enzymes in the vegetative portions of the plant would be deleterious.

Promoter Isolation:

The analysis of the castor EST sequences from Task 1.1, especially those obtained prior to the hybridization-based normalization process, was useful to identify highly expressed genes; such genes are typically expected to have strong promoters. This analysis also pointed to certain genes that would be expected to be expressed either constitutively or in a seed-specific manner, based on the expected role of the encoded protein in the plant, which in many cases has been documented in the scientific literature. We chose five genes that were expected to be expressed constitutively and five that were expected to be expressed primarily in seeds. The genes chosen are provided in Table 1-2.
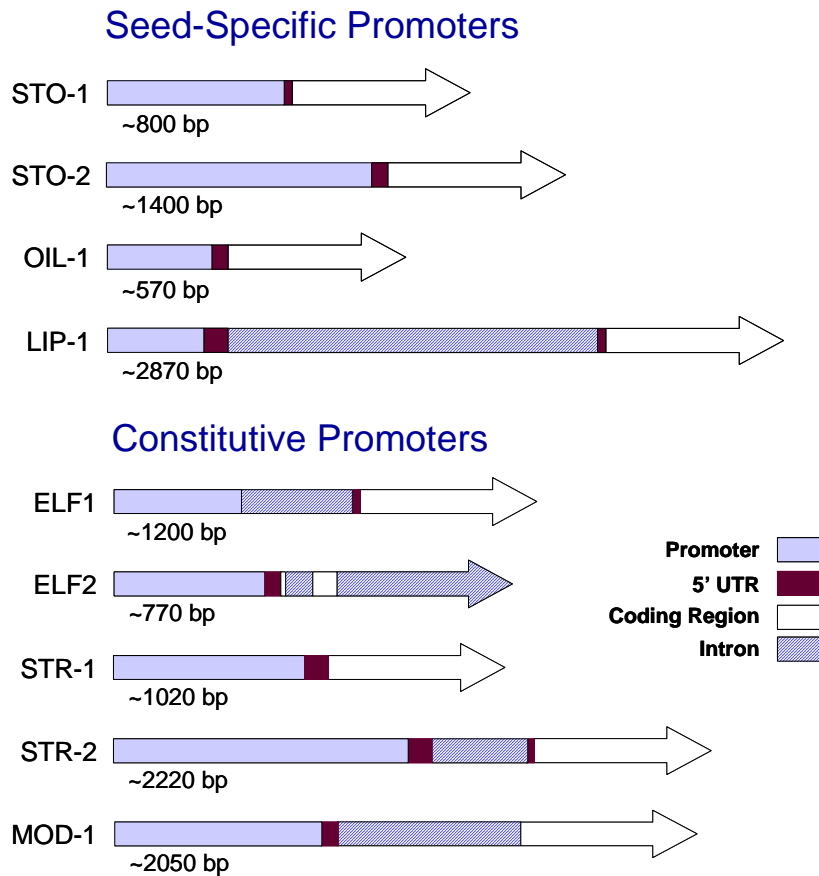
Table 1-2. Genes chosen for elucidation of promoter sequences.

Gene Designation	Expected Function of Gene Product	Expected Promoter Type
STO-1	Seed storage protein	Seed-specific
STO-2	Seed storage protein	Seed-specific
STO-3	Seed storage protein	Seed-specific
LIP-1	Lipid metabolism	Seed-specific
OIL-1	Lipid metabolism	Seed-specific
ELF-1	Protein synthesis	Constitutive
ELF-2	Protein synthesis	Constitutive
STR-1	Structural protein	Constitutive
STR-2	Structural protein	Constitutive
MOD-1	Protein modification	Constitutive

The promoters for these selected genes were isolated via PCR-based "genome walking," using a GenomeWalker™ kit (Clontech Laboratories, Inc.) according to the manufacturer's instructions. In this technique, short oligonucleotide adapter molecules with a known sequence were first ligated onto the ends of castor genomic DNA that had been digested with several different blunt-cutting restriction enzymes. A primary PCR reaction was carried out using a forward PCR primer that annealed to the adapter molecule and a reverse PCR primer that annealed specifically to the DNA within the coding region of each selected gene. A second round of PCR using nested primers, also based on the known adapter and gene-specific sequences, was then carried out. Using this strategy, DNA sequences upstream from the selected coding sequences were isolated and inserted into the cloning vector pCR2.1 (Invitrogen Corp.) and sequenced. The promoter sequences were confirmed by direct PCR amplification from genomic DNA.

Comparison of the genomic sequences obtained in this process with the assembled EST (cDNA) sequences revealed the presence of introns in a few of the cloned fragments. In some cases, the intron was found in the 5' untranslated region (5'UTR) of the respective gene and in other cases the intron was located in the coding region slightly downstream from the start codon. It has been established that introns in these locations can play a role in regulating gene expression (Rose 2004; Chung *et al.* 2006). A schematic drawing indicating the genomic DNA cloned, including coding regions and regions known to comprise at least a portion of the 5'UTR based on analysis of the corresponding cDNA sequences, is provided in Figure 1-1. (Note that the results involving the STO-3 gene were quite complex, in that four different versions of the DNA sequence upstream of the coding sequence were identified; further analysis of this promoter family was thus not conducted and this information has been omitted from Figure 1-1.)

Figure 1-1. Promoter regions isolated from various castor genes. The lengths of the isolated fragments upstream of the coding region are provided.

Seed-Specific Promoters

STO-1 — ~800 bp

STO-2 — ~1400 bp

OIL-1 — ~570 bp

LIP-1 — ~2870 bp

Constitutive Promoters

ELF1 — ~1200 bp

ELF2 — ~770 bp

STR-1 — ~1020 bp

STR-2 — ~2220 bp

MOD-1 — ~2050 bp

Promoter

5' UTR

Coding Region

Intron

Promoter Testing:

Transient expression: Four of the promoter regions isolated during this project were selected for functional testing. Initial studies utilized the firefly luciferase gene as a reporter

gene; the gene product (luciferase enzyme) emits light when active, and therefore the amount of luciferase produced as a result of ligation to the castor promoter regions can in theory be quantified by the use of a luminometer. Initial transient expression studies attempted to detect luciferase activity in tobacco cells (*Nicotiana tabacum* NT1) into which plasmid DNA containing the promoter-luciferase expression cassette was introduced *via* electroporation. The luciferase gene-containing vector pGL3 (Promega Corp.) was used as the parent plasmid in these studies, and the ELF-2, STR-1, STO-2, and OIL-1 promoter regions were ligated directly upstream of the luciferase gene. Experiments were conducted using constructs containing either the SV40 terminator present in the parent pGL3 vector or an introduced Cauliflower Mosaic Virus 35S (CaMV35S) terminator region. Unfortunately, the proportion of cells that take up DNA in such experiments is typically very low, and thus the luciferase activity, measured with a SpectraMax Gemini XS microtiter plate reader using Promega's Bright-Glo Luciferase Assay System, was too low to detect. We were successful, however, in utilizing luciferase expression as a measure of promoter activity in stably transformed *Arabidopsis thaliana* plants (see below).

Unlike the case with the luciferase experiments, transient expression assays in tobacco suspension cells were successfully performed using reporter genes that encode fluorescent proteins. The genes for Cop-Green Fluorescent Protein (CopGFP) from the copepod *Pontellina plumata* and Phi-Yellow Fluorescent Protein (PhiYFP) from the jellyfish *Phialidium sp.* were obtained from Evrogen (Moscow, Russia). The CopGFP and PhiYFP genes were amplified from the source plasmids *via* PCR such that appropriate restriction enzymes were introduced at each end of the coding region, allowing the replacement of the luciferase gene in the pGL3-derived vectors containing the ELF-2, STR-1, STO-2, and OIL-1 promoters and CaMV35S terminator with the CopGFP or PhiYFP genes. The various constructs were introduced into tobacco NT1 cells *via* electroporation and then examined with fluorescence microscopy to identify cells expressing the fluorescent proteins. For these analyses, we used a Nikon Eclipse E800 microscope with a VFM Epi-Fluorescence Attachment containing Filter Block B-2E/C (excitation filter = 465-495 nm, dichromatic mirror cut-off = 505 nm, barrier filter = 515-555 nm). Although this is a non-quantitative assay, the ability of the ELF-2, STR-1, STO-2, and OIL-1 promoters to drive the expression of the reporter genes was clearly shown. Representative photomicrographs of cells expressing the fluorescent proteins are shown in Figures 1-2 to 1-4.

Stable Expression in Plants: Since the transient expression assays in tobacco cells indicated that the four castor promoters tested were functional, the next step was to test the promoters in stably-transformed whole plants, which enables an assessment of the temporal and spatial expression patterns of genes placed under the control of the promoters. Plant binary expression vectors were constructed by inserting the various 'castor promoter-CopGFP-CaMV35S terminator' expression cassettes cleaved out from the transient expression vectors described above into a basic binary vector backbone. Similar constructs were made using expression cassettes containing the OIL-1 and STR-1 promoters coupled to the luciferase gene. The binary vectors contained a phosphinothricin acetyltransferase (PAT) selectable marker gene to enable growth of transformed plants in the presence of glufosinate, a spectinomycin resistance marker gene to enable selection of transformed bacteria containing the plasmid, and T-DNA border sequences that allow *Agrobacterium*-mediated recombination of the bordered DNA (including the CopGFP expression cassette) into the plant's chromosomes.

Testing was performed in *Arabidopsis thaliana*, a small plant in the mustard family that has been very heavily utilized in the plant research community as a model system because of its well established genetic system, fast growth cycle, and ease of genetic transformation. Transgenic *Arabidopsis* plants were produced by the *Agrobacterium*-based floral dip method

13

(Clough and Bent 1998) In order to determine which portions of the plant the various promoters were active in, certain tissues were harvested from first generation (T1) transformants, including leaves, developing siliques (seed pods), and mature seeds. After grinding in specific lysis buffers, these tissues were tested for CopGFP and luciferase activity using a microtitler plate reader (SpectraMax Gemini XS) in fluorescence and luminescence modes, respectively.

In order to determine whether the promoters continued to function in subsequent generations, the seeds from the T1 plants were planted to generate second generation (T2) plants, from which leaves, developing siliques, and mature seeds were also taken and analyzed.   The results of the studies utilizing CopGFP as the reporter gene in conjunction with the STR-1, OIL-1, and STO-2 promoters in T2 plants are presented in Figures 1-5 to 1-7. Testing of the ELF-2 promoter was not carried out with T2 plants because the T1 results indicated that the activity of this promoter was quite low (*i.e.,* less than 0.04 RFU/mg protein for all tissues). For the STR-1 promoter, expression was substantially higher in leaf tissue relative to siliques and seeds.   For both the STO-2 and OIL-1 promoters, there was clearly much more activity in seeds than in vegetative leaf material. The developing siliques contain both vegetative tissue and developing seeds, and activities of the STO-2 and OIL-1 promoters in these tissues were typically in between those of the leaf tissue and mature seeds. Experiments were also carried out in which the OIL-1 and STR-1 promoters were operably linked to the luciferase gene; the outcome of these studies was similar to the results obtained with CopGFP as the reporter gene (data not shown).

1.3 - Gene expression studies

The goal of Task 1.3 was to conduct gene expression studies with castor developing seeds to evaluate which genes contribute to various phenotypes exhibited by different castor varieties *(e.g.*, oil content, ricinoleic acid content, etc.). The plan was to use small (~200 gene) oligonucleotide-based microarrays containing regions of selected genes involved in oil and intermediary metabolism to evaluate the expression of these genes in several different varieties. To initiate this task, we collected developing seeds and leaf samples from six different castor varieties growing in the plant nursery of our collaborator Castor Oil, Inc. in Plainview, Texas. The samples were immediately frozen in liquid N_2, and then shipped on dry ice back to the San Diego labs.

We also initiated a bioinformatics analysis of our castor gene dataset in order to design oligonucleotide-based microarrays to investigate expression of genes that may be important for industrial oil accumulation in castor seed. To this end, we identified a limited set of genes (~200) expressed in castor seed that are known to be involved in intermediary metabolism and in the metabolism of lipids and storage proteins. Genes involved in the following categories of lipid metabolism were cataloged: fatty acid biosynthesis, membrane lipid biosynthesis (both in plastids and endoplasmic reticulum), storage oil biosynthesis, wax and cutin biosynthesis, lipid and fatty acid degradation, and lipid signaling. In addition, representative genes from the following categories were identified: amino acid biosynthesis and degradation, carbohydrate biosynthesis and degradation, nucleoside biosynthesis and degradation, vitamin/cofactor biosynthesis and degradation, and secondary metabolism.

The design of gene-specific oligonucleotides for the microarray had been initiated as part of this task, but a reduction in funding for the project resulted in halting additional activities for this task

Figure 1-2. Photomicrographs of tobacco NT1 cells transiently expressing CopGFP under the control of the OIL-1 promoter. A. Brightfield microscopy. B. Fluorescent microscopy.
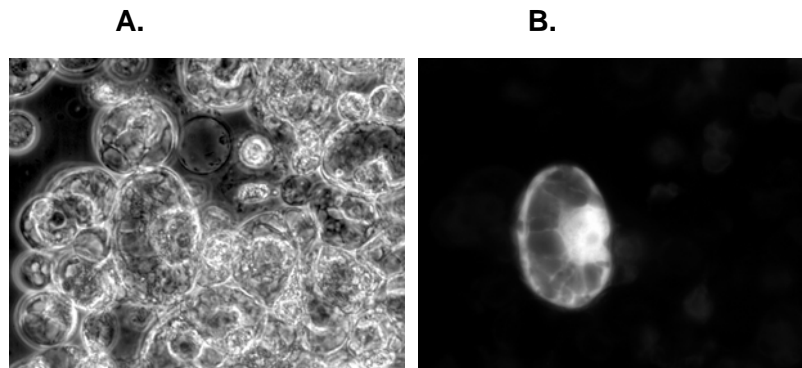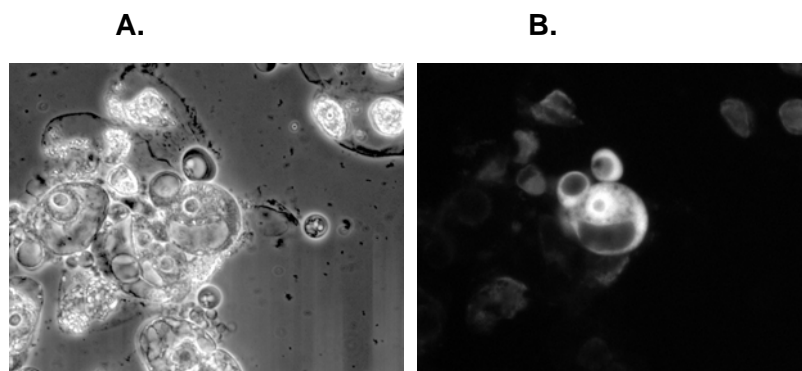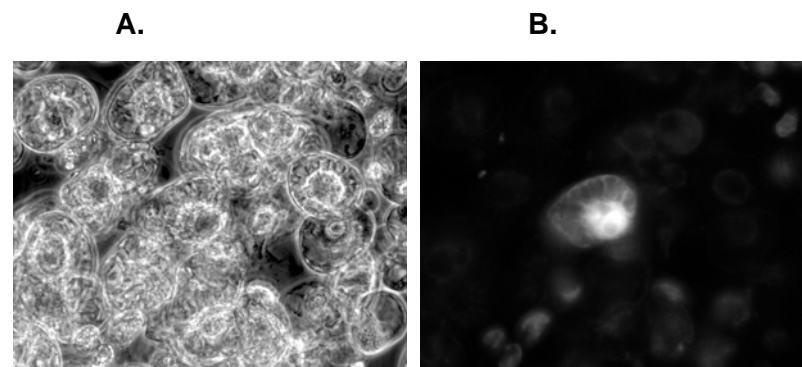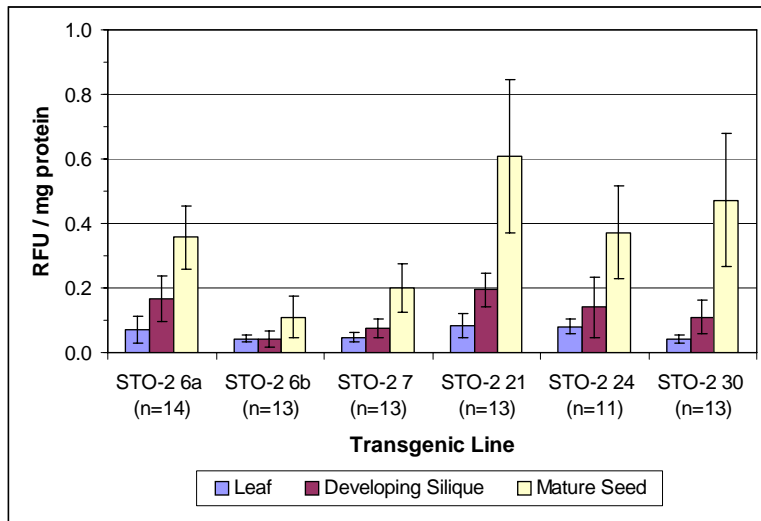
A.   **B.**

Figure 1-3. Photomicrographs of tobacco NT1 cells transiently expressing CopGFP under the control of the STR-1 promoter. A. Brightfield microscopy. B. Fluorescent microscopy.

A.   **B.**

Figure 1-4. Photomicrographs of tobacco NT1 cells transiently expressing PhiYFP under the control of the STO-2 promoter. A. Brightfield microscopy. B. Fluorescent microscopy.

A.   **B.**

Figure 1-5. Expression of CopGFP coupled to the castor STO-2 promoter in various tissues of T2 *A. thaliana* plants transformed with pDOW2771.
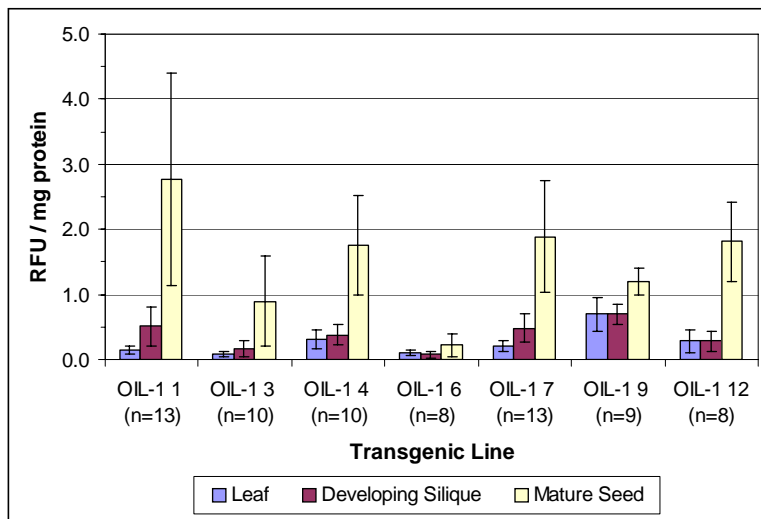

Figure 1-6. Expression of CopGFP coupled to the castor OIL-1 promoter in various tissues of T2 *A. thaliana* plants transformed with pDOW2772.
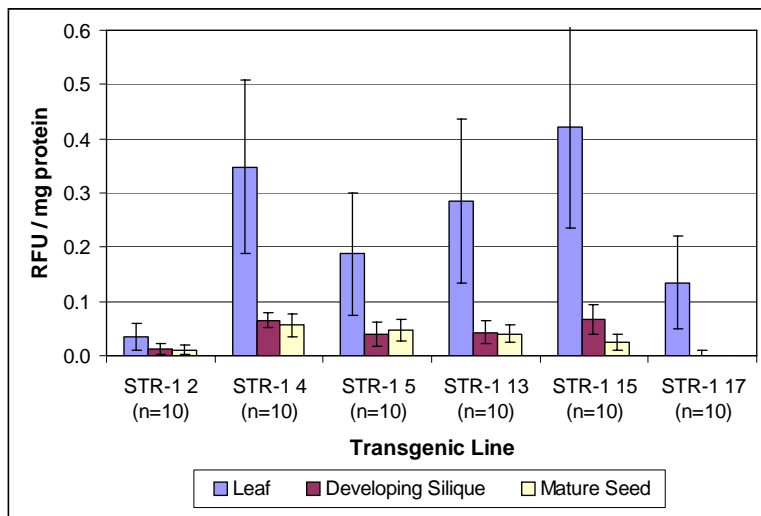

Figure 1-7. Expression of CopGFP coupled to the castor STR-1 promoter in various tissues of T2 *A. thaliana* plants transformed with pDOW2769.

17

References:

Ewing, B. and Green, P. (1998) Basecalling of automated sequencer traces using *phred*. II. Error probabilities. *Genome Res.* **8**, 186-194.

Gish, W. and States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nature Genet.* 3:266-72.

Wu, C.H., Apweiler, R., Bairoch, A., Natale, D.A., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Mazumder, R., O'Donovan, C., Redaschi, N. and Suzek, B. (2006) The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* 34:D187-91. (see also http://www.pir.uniprot.org/index.shtml)

Rose, A.B. (2004) The effect of intron location on intron-mediated enhancement of gene expression in Arabidopsis. *The Plant Journal* 2004 40:744-751.]

Chung, B.Y., Simons, C., Firth, A.E., Brown, C.M. and Hellens, R.P. (2006) Effect of 5'UTR introns on gene expression in *Arabidopsis thaliana*. *BMC Genomics* 19:120

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant Journal* 16:735-743.

USDA/WRRC: Tom McKeon

Because castor oil is an important chemical feedstock for producing a wide array of products, this project was geared to the development of the castor plant as a renewable source of industrially useful fatty acids. Since castor oil is 90% ricinoleate (Figure 1-8), our working hypothesis has been that castor could be engineered to produce other fatty acids bearing polar substituents. The project supported the elucidation of castor oil biosynthesis, particularly in terms of producing alternate fatty acids in castor. Our transformation process encountered difficulty in numbers and stability of tranformants so, as we continued, we used this project to help us develop a more robust transformation system, and we developed an understanding of how castor oil biosynthesis is controlled at the compositional level. This grant supported our effort in analysis of ricin gene expression in developing castor seed, induction of multiple shoots from castor tissue explants for cloning of castor varieties and as an intermediary to high efficiency genetic transformation, and enzymatic conversion of castor oil to enzymatic substrates with potential use as intermediates in synthesis of advanced materials, the mono- and di-ricinolein acylglycerols.

Figure 1-8. Ricinoleic Acid
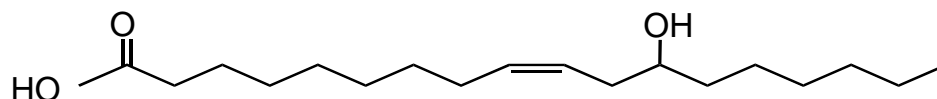
The role of WRRC in this project was involved in Task 1, Plant Science: Castor genomics, biochemistry, and metabolic engineering. The results are in greater detail in a publication resulting from the work. The results obtained in each of the relevant sub-tasks are described.

Gene expression studies in castor

The WRRC role in gene expression studies was limited, primarily focusing on the expression of genes encoding the noxious proteins ricin and 2S albumin, and genes involved in oil biosynthesis. Expression of ricin and 2S albumin coincide with oil biosynthesis, and continue beyond the cessation of oil production (26-40 days after pollination) in the maturing castor seed (Figure 1-9). Genes involved in castor oil production showed different patterns of expression (Figure 1-10), providing potential targets for developing differentially timed promoters.

Figure 1-9. Northern Analysis of 2S Albumin and Ricin Gene Expression in Castor.
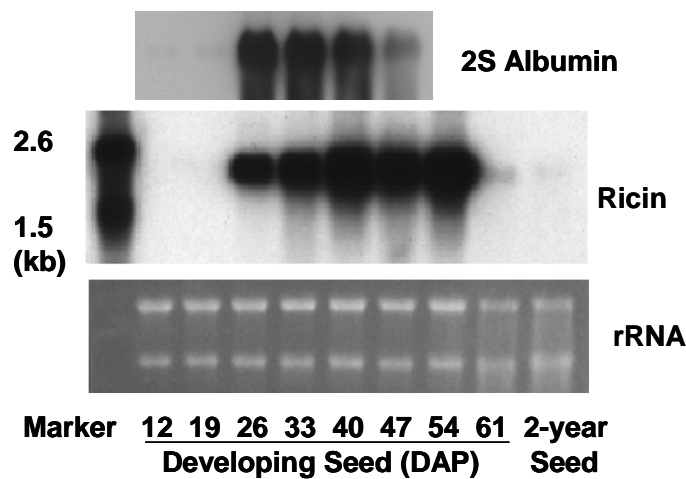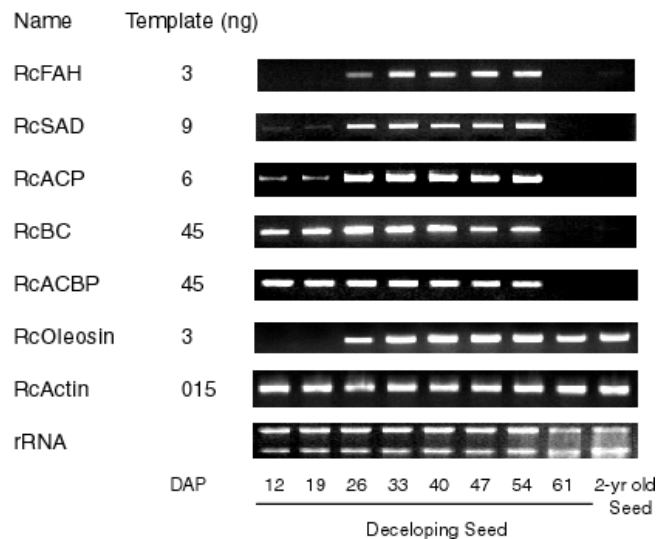

Figure 1-10. Expression of castor genes during seed development. Rc: *Ricinus communis*; FAH: oleoyl hydroxylase; SAD: stearoyl ACP desaturase; ACP: acyl carrier protein;BC: biotin carboxylase; ACBP: acyl-CoA binding protein.

1.4 - Castor Transformation System Development

Castor has proven to be fairly recalcitrant to transformation. The initial system developed has proven to have a very low frequency of success, with considerable conversion to chimeric transformants. As a result, the transformation effort was focused on developing a system that would allow transformation of tissue culture plants. The primary achievement of this work was generation of treatments that allow multiple shoot regeneration from explants derived from different parts of the plant, including shoots and mature embryos (Figure 1-11). Conditions using biolistics to introduce genes into castor tissue were also developed. Although the impacted tissue was transfected and expressed the genes introduced, the tissue became recalcitrant to regeneration. We solved both halves needed to achieve the goal of castor transformation, but were not able to make both halves work together.

Figure 1-11. Plantlets regenerated from castor embryo axes.

Multiple shoot generation using embryo axes

Two clusters from one embryo axes
: at least 5 shoots per cluster

Two clusters from one embryo axes
: at least 5 shoots per cluster

1.5 Development of castor genetic tools

Due to limitations of funding and lack of an efficient transformation system for castor, the effort on this task was minor. One important offshoot from the work on the ricin gene was that we developed a means to differentiate ricin expression from *Ricinus communis* agglutinin (RCA) expression (Figure 1-12). This work is currently being pursued at WRRC and has led to development of a method for detecting contamination of food with crude ricin preparations.

Figure 1-12. Differential detection of ricin and RCA gene expression.
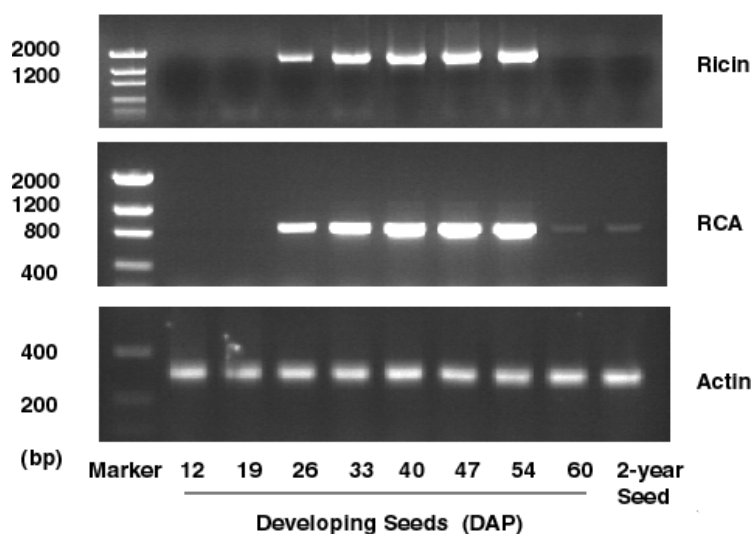
1.6 Castor seed biochemistry studies

Prior to the initiation of this project, the WRRC group had developed chromatographic and enzymatic analyses of castor oil biosynthesis that led to the identification of several enzymes that contribute to the high ricinoleate content of castor oil. The results of this work are summarized in Figure 1-13. As part of the biochemical analysis of how the castor seed directs synthesis of an oil with a high content of hydroxy fatty acids, the WRRC group cloned genes for several enzymes of which two direct ricinoleate incorporation into oil. The first is the diacylglycerol acyltransferase (Type 1 DGAT) which shows a strong preference in acyltransferase activity for diricinolein (Figure 1-14). In a comparison, the DGAT showed a twofold preference for incorporating added acyl-CoA into a diricinoleoyl-diacylglycerol (DG) versus the same enzyme from *Arabidopsis*, a plant which does not make ricinoleate (Figures 1-15 and 1-16). This preference for DG containing hydroxy FAs is the subject of a patent application (serial no. 10/861,616).

The WRRC group cloned genes for a second class of enzymes, the acyl-CoA synthetases (ACS) which comprise a gene family that provides an essential metabolite involved in acyl transfer and beta-oxidation of fatty acids. Two of the cloned genes expressed enzymatically active proteins in a yeast expression system (Figures 1-17 and 1-18). One of these enzymes showed a distinct preference for incorporating oleate into the CoA (Figure 1-19), while the other showed a threefold preference for using ricinoleate versus oleate (Figure 1-20). Together, this ACS and the Type 1 DGAT display a six-fold preference for utilizing ricinoleate versus oleate.

This is similar to the preference observed for castor microsomes incorporating ricinoleate into TG versus oleate.

Figure 1-13. Enzymatic reactions that lead to high ricinoleate oil. Genes for yellow highlighted enzymes have been cloned by WRRC researchers, hydroxylase by Somerville group. We have established the biosynthetic pathway of castor oil and have determined the enzymatic steps driving ricinoleate into castor oil. This information has elucidated the process by which seeds control the fatty acid composition of their oil, especially with regard to incorporating uncommon fatty acids.
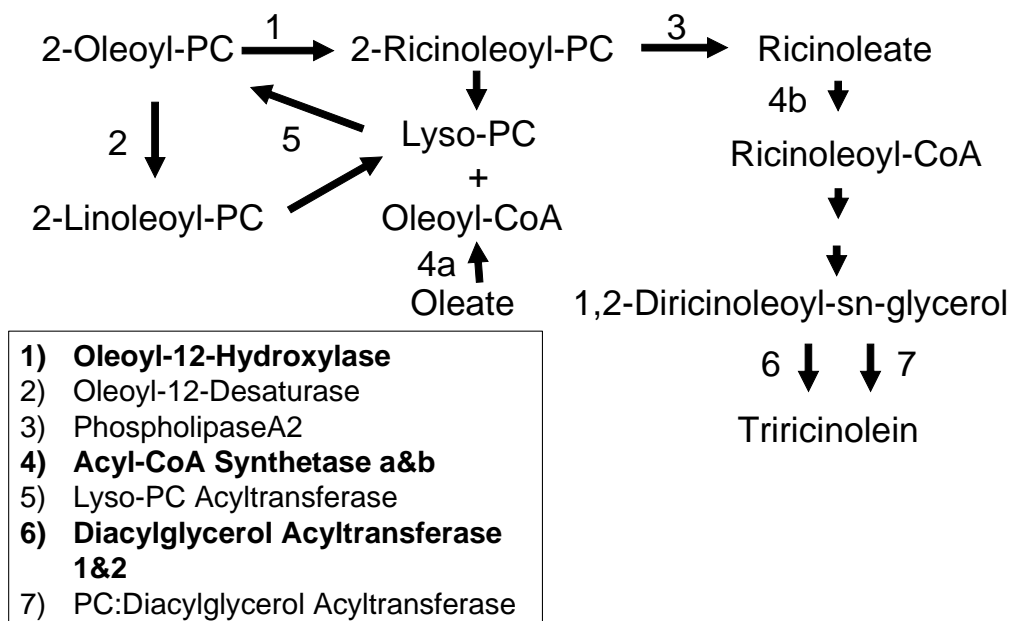
2-Oleoyl-PC $\xrightarrow{1}$ 2-Ricinoleoyl-PC $\xrightarrow{3}$ Ricinoleate

2-Oleoyl-PC → 2 → 2-Linoleoyl-PC

5 → Lyso-PC + Oleoyl-CoA

4a ↑ Oleate

4b ↓ Ricinoleoyl-CoA

↓

1,2-Diricinoleoyl-sn-glycerol

6 ↓   ↓ 7

Triricinolein

1) **Oleoyl-12-Hydroxylase**
2) Oleoyl-12-Desaturase
3) PhospholipaseA2
4) **Acyl-CoA Synthetase a&b**
5) Lyso-PC Acyltransferase
6) **Diacylglycerol Acyltransferase 1&2**
7) PC:Diacylglycerol Acyltransferase

Figure 1-14. Preference of RCDGAT for incorporating acyl group into diricinolein.
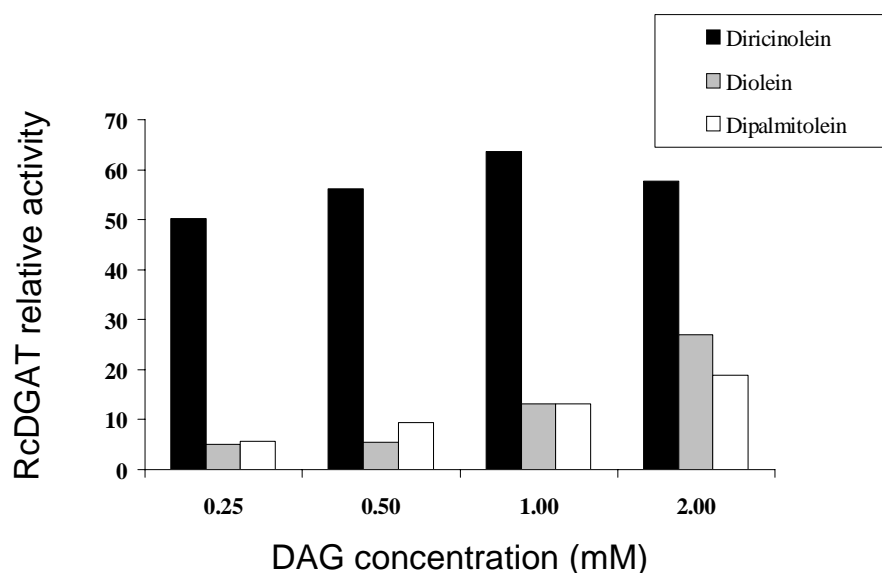

23

Figure 1-15. Activity of castor DGAT and *Arabidopsis* DGAT on dipalmitolein.
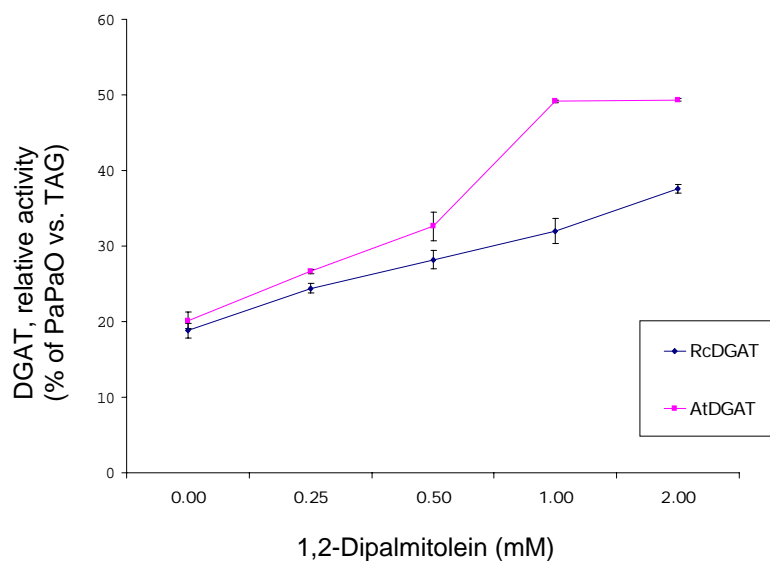

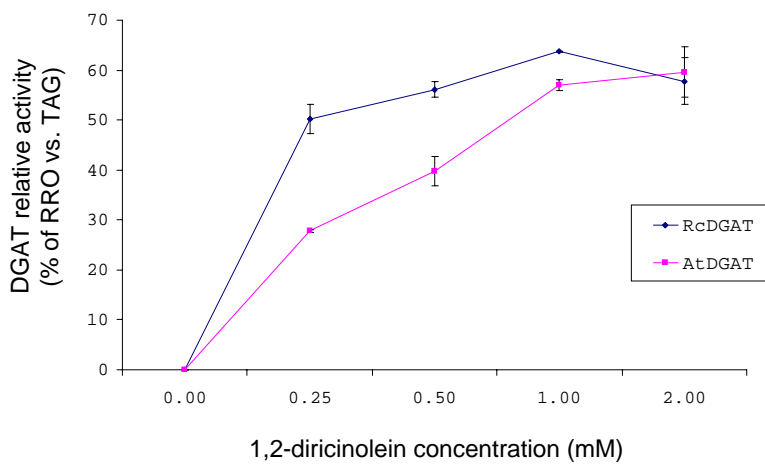**Figure 1-16.** Activity of Castor DGAT and *Arabidopsis* DGAT on diricinolein as acyl-recipient.

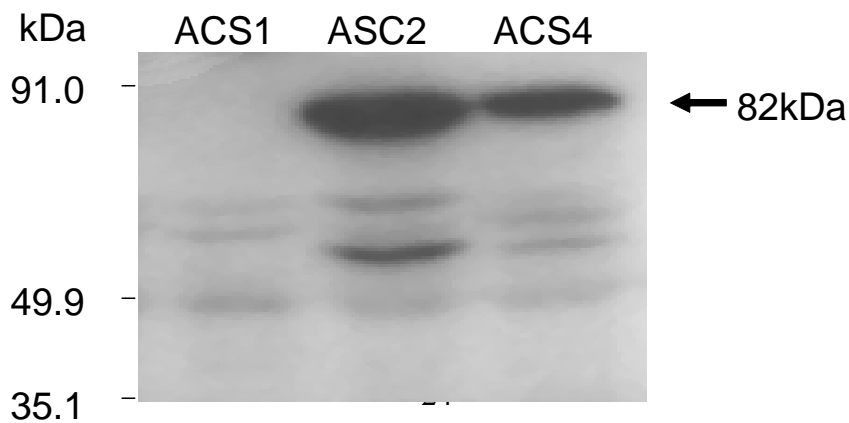**Figure 1-17.** Expression of putative acyl-CoA synthetase genes in a yeast expression system.

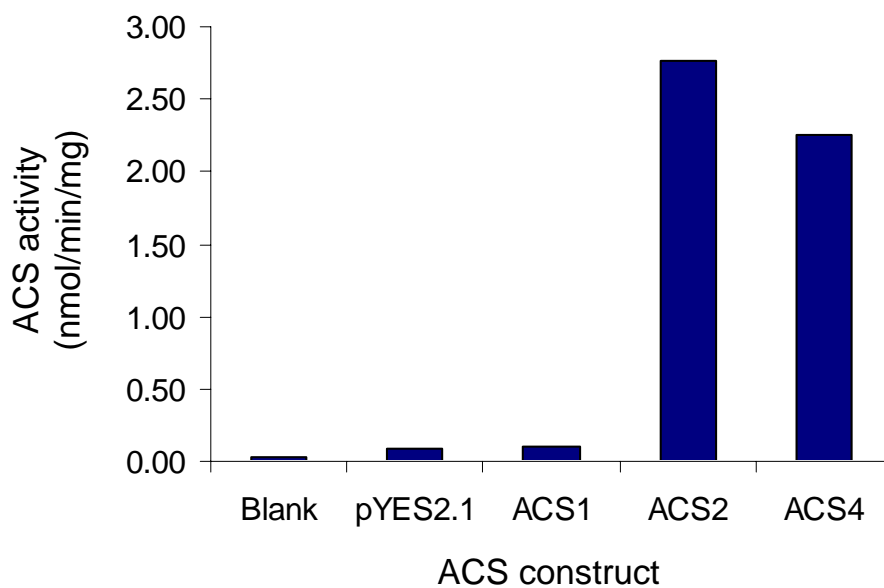**Figure 1-18.** Acyl-CoA Synthetase activity of expressed ACS genes in yeast microsomes.

Figure 1-19. Acyl-CoA Synthetase activity of ACS4 with different fatty acids substrates.
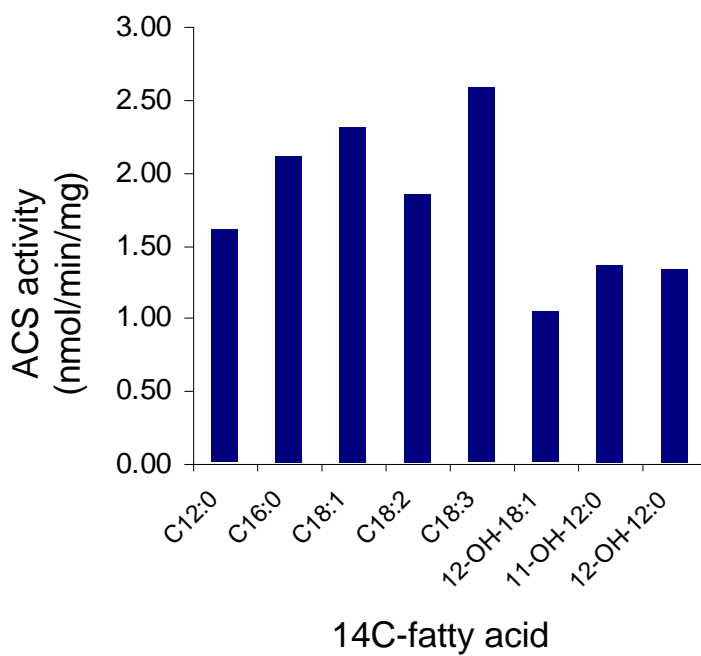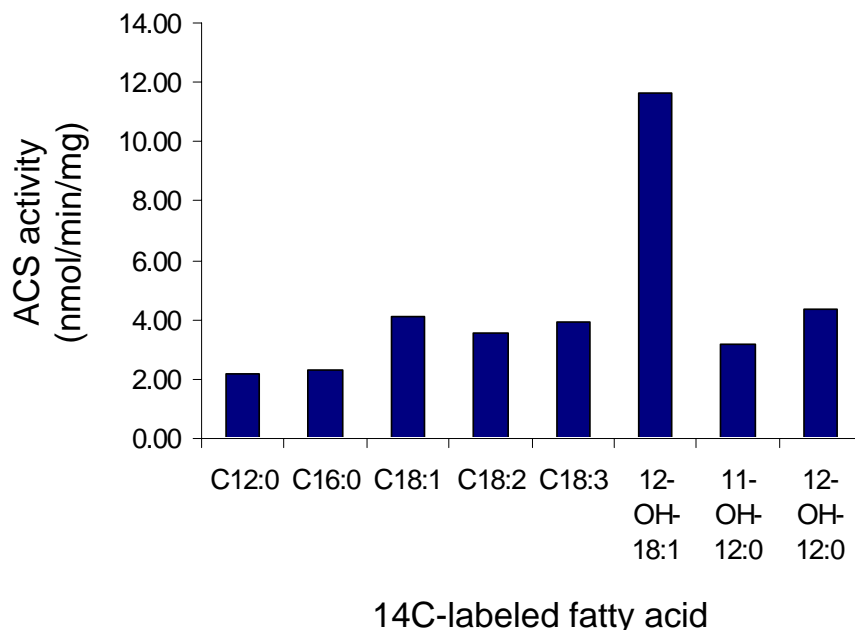
Figure 1-20. Acyl-CoA Synthetase activity of ACS2 with different fatty acids substrates.

1.7 Metabolic engineering of castor:

The hypothesis underlying the original proposal was that the castor plant, which makes an oil containing 90% ricinoleate (Figure 1-8), would be amenable to making oils containing fatty acids with a hydroxyl group in other positions on the FA chain, and would also be able to make oils containing FAs with other polar substituents. Because of the problem with castor transformation, the WRRC group was only able to test for this capability using isolated castor microsomes. One fatty acid class of interest was medium chain terminal hydroxy fatty acids. The castor system was unable to incorporate these FA into the oil (TG) in vitro. Technology for producing novel acyltransferase substrates was developed as part of this task, and included acylglycerols containing 16-OH palmitate, 12-OH-stearate, and lesquerolate.

Task 2 Production: Development of an Industrial Oilseed Crop

Castor Oil Inc, Lee Browning

The second key barrier area in the Technology Roadmap for Plant/Crop-Based Renewable Resources was defined as Production. Production includes the needed improvements to have a robust industrial crop that can be grown under a range of field conditions and have the processing infrastructure needed to be commercially economical.   Task 2 focused on agronomic improvements for castor including harvesting equipment, seed crushing and the identification of new value-added uses for the residual meal.

2.1 Testing and Selection of Castor Lines

There were parts to this goal that were difficult to manage. One part was around ricin content. Ricin has no agreed upon safe level. As far as testing for ricin this is costly and the methodology differs so greatly between labs on testing that we decided not to test for ricin content. Higher oil content was a direct correlation between seed size to oil content the larger the seed, the lower the oil content the smaller the seed the higher the oil content. So we went with the seed size that makes the most sense with harvesting capabilities. The seed size that was most appealing was from 29 grams per 100 seed to 39 grams per 100 seed. Texas castor did not seem to have many pests that affected castor negatively. There were a lot of insects that lived around or on castor plants. There were very few insects that attacked the castor and affected them negatively, at least in Texas. More diseases were found in a wet year than in a dry one but grey capsule mold seemed to be the disease that affected the castor the most. This was not a common occurrence and usually affected smaller seeded castor and tight clusters of seed, maybe due to the compacted racemes of small castor seed. We are still testing drought tolerance from seed population to length of internodes. This concept was difficult to measure. Non dehiscence (castor seeds shattering before harvest) was by far the easiest characteristic to monitor. Crop uniformity was desirable for dwarf plants that could be mechanically harvested and was easy to identify. Higher yields were another characteristic that had many variables to consider, plant population, row spacing, raceme length, raceme number per plant, and fruiting pattern. The smaller the plot the more difficult it was to monitor the variables. On several occasions in the nursery data looked very promising but on a larger scale this data proved to be faulty. Overall we found strong lines that showed most of the favorable characteristics.

In testing and the selection of the castor seed there was very little data to go along with the seed. The task started with visually looking at the seed for consistent seed size. Photographs were taken and 100 seeds weighed. The goal was framed with harvestablity in mind i.e. consistent seed size. There were many obstacles once castor was actually growing in the nursery. Finding reliable workers that would bag the castor racemes in a timely manner was difficult.   Water requirements were different in different plots because of different maturity dates.   There were some castor plants that had such a long maturity time that they never produced any seed. Recording the data was another very large task in the field as in the office as well as trying to decode what information was important verses the information that we had. Depending on the goal all the information was helpful, but knowing when to use it was the next key step. Another obstacle was that for some seed lines there was a lot of information on the lines but very few seeds and in some instances the germination of the seed was zero. These lines are lost forever. We were successful in testing and looking at many castor seed lines but this will always be an ongoing process which goes hand in hand with the breeding process. We

will never be satisfied with the outcome because we want all the great characteristics to be applied to the few lines that we want to develop and all the other characteristics should be for everyone else. Castor seed is a fortress built into one plant with many different varieties each having its own set of plans. Every time one thinks they can exchange one quality for another it actually affects several characteristics which creates it own set of rules. The saying that for every action there is an opposite and equal reaction, is amplified in castor breeding in that one response or action affects many other factors that may or may not be related but it also creates as many answers as it does questions. The castor breeding, because of its complex design, was like a monopoly board game that cannot be finished.   One must be satisfied with the results. A good example of this relates to low ricin castor seed. The low ricin level seed still requires the same precautions as castor seed with high levels of ricin. In my opinion, the only way to eliminate ricin from castor is to genetically modify castor seed to eliminate the ricin.

2.2 Breeding for Improved Castor Lines

COI started with a female which is S-pistil strain.   They were selected for having sex reversal traits based on environmental basis for being either female or monoecious (being both male and female). The environmental characteristics were needed in the very beginning of the growing cycle at the time of germination, but could occur during the growing cycle.   This crop (when growing a pure female line) had to be rouged several times throughout the growing season to insure purity within the line. The environment in which this female line was normally grown was different than our area. In the first year, we had to go to a different climate to make sure we had a pure female line. Once we had enough seed of the female line, we were able to grow several fields of differing crosses to make different hybrids.

Breeding for improved castor lines seemed to be an easy task when looking at it from the basic concept of two genes determines the out come out of four. Castor was so much more complex than your basic fruit fly or sorghum breeding program. The other factor was the environmental conditions that affected the female plants. Since there was not much known information about the different varieties one had to grow them first and note the physical characteristics. But to insure purity, this must be done for at least two years. The racemes on the plants must be bagged in a timely manner or have distance as between different lines to insure purity of the line. Because of the light weight of castor pollen, the distance was measured in miles not feet. When crossing the plants, the planting dates coincided with the first raceme because the first raceme was the largest. The castor fruiting was indeterminate and the yield could vary depending on the season length, watering pattern and the soil pH had a larger factor than we realized. After a crop was planted only 80 miles from Plainview, we found out that in mid-season that the pH was 9 plus. Texas had very wide varying weather patterns from early or late spring planting to early and late fall freezing temperatures. Weather included very windy to extreme temperature changes in one day to on going through out a season. The known facts about Texas weather are there will be no two years alike. The variation has produced castor germination in early March and then the following year the germination in mid May. In a four year breeding program accurate averages to cover all possibilities cannot obtained.   In spite of these difficulties, four or five good high yielding hybrids were identified.

More testing still needs to be done in the field in several areas to include; 1) plant population, 2) row spacing, 3) watering accuracy (both in quantity and timing), 4) optimal planting dates, and 5) harvesting efficiency. Based on our work and experience, harvesting efficiency could change the yields by 15 to 50 percent.

2.3 Improve Harvesting Technology

Harvest technology for castor was a multi-problem task. The impediments were numerous, like oil content, soft meat of castor seed, delicate or brittle seed coat, light bushel weight and the potential for machinery that might be available. COI built a harvester that would remove only the capsule and hull from the plant and not take the whole plant through the combine.   In the course of the study, some farmers claimed that they could harvest castor seed by cutting the stalk at ground level and separating the castor seed from the stalk stem and leaves using a conventional combine. COI tried this approach with terrible results. This was an unmanageable technique due to the large proportion of stalk stem and leaves.   After spending four days setting and resetting the combine, COI went back to the converted combine noting the changes that could be made. Every year of the study, other weak points were found with the combine or the header and the necessary changes were made. This goal could be further improved by modifying a header to remove the castor seed and the hull but not hulling them. This modification might even speed up the harvest, but then stationary hullers would be needed to hull the castor seed.

COI received claims of being able to harvest castor using conventional combines but since we received the seed and they were harvesting over one hundred miles away it was very impractical to leave the receiving station to witness the harvest. We took notes on settings and tried to duplicate these with our combines with very poor results. There was far too much foreign material going thru the combine to do a accurate job of separation.   Not only did we have a lot of trash in the seed bin but we also lost a large portion of seed on the ground through poor separation in the combine. After studying the problem, we determined we should be able to use a modified header to harvest the seed and capsules from the plant.   Then we would use a conventional combine, not trying to remove the capsule from the seed, and then hulling the seed using a stationary huller to remove the capsules from the seed.   This should not only improve efficiency but also aid in gathering more seed before weather related problems cause damage or loss of the crop. This technique should make the free fatty acid content go down as well.   The broken seed should have less time to react with the oxygen before oil extraction, increasing the oil recovery. Unfortunately, the capsule weight contributed 25 percent of the total weight when combined.

2.4 Improved Oil Extraction Process

This was a bigger task than what was hypothesized. COI believed that there would be enough prior knowledge on oil extraction in general and it could be applied to castor seed. This was not the case. The task was initiated by attending Texas A&M to take a short course on oil extraction. It turned out to be on refining oil. The course was a good course but it was hard to follow without knowing the extraction part first.   Two weeks later I took a course on oil extraction and they told me that some one had given incorrect dates and these courses were inverted. After taking the oil extraction short course, the process started to make sense. Of the speakers that were there, none of them has had any experience with castor oil extraction. Most of the people there wanted to know either about a specialty crop or they were there to learn about soybean oil extraction. As the course went on, I was not getting the needed information. They told me about a pilot plant where we could crush castor seed. We set a date to crush and took about six tons of seed to crush. We tried several methods and made several attempts to see what would be the most cost effective method.   We also looked at this from an energy

return stand point. On a small scale it was not cost effective to use chemical extraction. We tried both hexane and acetone as solvents to remove oil from the castor cake. Both solvents have pro's and con's. The acetone removed the oil the best but would probably take the most energy to refine or distill the acetone. The hexane also removed the oil and was the easiest to refine with the least amount of energy. There were several obstacles but the goal was achieved.

Once at the pilot plant at TAMU, we met for a plan of action. New issues came up due to the ricin and potential safety issues. TAMU needed more time to research the hazards associated with oil extraction of castor. Five tons of seed had been transported to crush and we wanted to leave the castor seed there but they would not allow the seed to be unloaded. Their fears were numerous and no one was willing to work with the seed until all objections had been satisfied. One month later, the seed was delivered for a second time to test extraction techniques. Upon meeting a second time, everyone but me took all precautionary measures, goggles, disposable coveralls, fine particle masks, and rubber or leather gloves. As the day progressed, everyone discarded everything but the gloves. Castor oil extraction was done on several occasions. The first lesson learned, was that oil cannot be separated from the meal with out binding the protein other wise we made castor butter. This made for a very difficult filtering process. We tried an extruder with the thought not cooking the castor. Pressure and a little steam would recover a little oil off the cage and then solvent extract the rest. There are several heads to the extruder and it was determined the square head with adjustments is the best for castor.

The next problem that occurred was fine particles clogged the screens flooding the beds making this technique un-useable. The protein had to be bound by cooking the seed. We found there are two effective ways to extract the oil. The first technique included cooking the seed then going thru a hard press and the residual oil in the castor cake is 5% or less. For small quantities this was the most economical way, using the least amount of energy. The second way was to cook the seed, go thru a soft press removing 17 to 25 % of the oil and then go thru a solvent extraction process to remove the balance of the oil. This was a more economical way for larger quantities of seed.

The next obstacle to overcome was the filtering and refining of the oil. TAMU did not have the time so we went out looking for others to help. ACH had a pilot refining facility in Memphis TN. The biggest problem with castor oil was the filtration of the oil due to the viscous nature and large amount of fines in the oil.   The seed may not have been cooked long enough to bind the protein. The oil was filtered using the correct filtering aid. Many days were spent in the lab trying different clays and temperatures determining the best for the desired result which is US #1 castor oil. The specifications were optimized and many obstacles over come like removing green color out of the castor oil.

The next and larger obstacle included how to add value to the residual castor cake. Extensive testing is underway in many areas using castor cake as fire ant control, potting soil mixes, fungicides for citrus nematode control in strawberries as well as tobacco and as a fertilizer in vegetable and flower gardens. Several areas have been successful but then labeling must be addressed.   Extracting the protein in a new way was tried but failed due to the residual oil left in the cake.   If the residual oil was one percent or less it could be done cost effectively but currently residual oil is five percent. Maybe once crush volumes are enough to justify a solvent extraction unit it will be feasible to extract the protein.
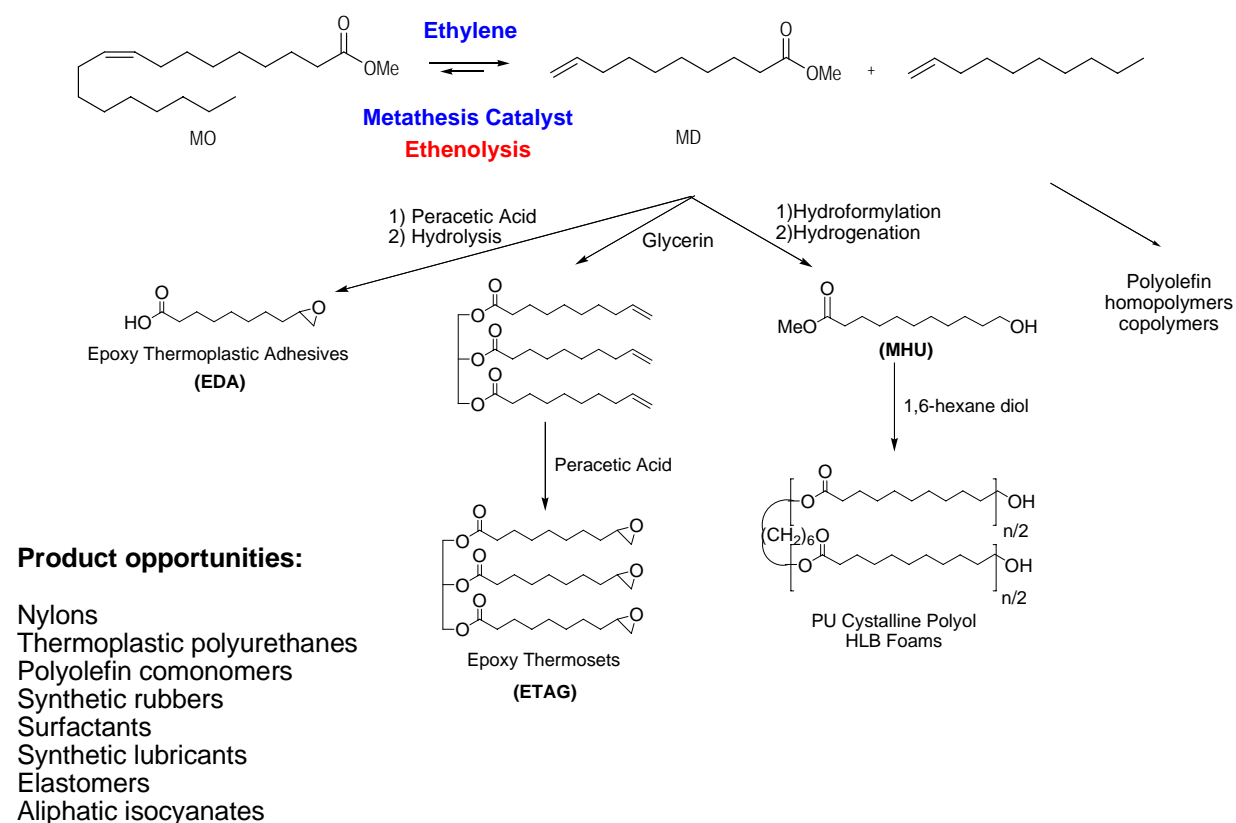
Task 3: Catalyst Discovery

Project Activity:

Hypotheses:

Ethenolysis of fatty acid methyl esters, FAME's, derived from seed oils leads to α,ω-functionalized derivatives which will serve as valuable, renewable feed-stocks for the chemical industry. The cascade of materials which would result from this transformation and the potential application opportunities these materials would participate in is shown in Figure 3-1. The key enabling technology to the entire process lies in the catalytic performance for the metathesis transformation.

Figure 3-1. Cascade of Materials and Product Opportunities

Product opportunities:

Nylons
Thermoplastic polyurethanes
Polyolefin comonomers
Synthetic rubbers
Surfactants
Synthetic lubricants
Elastomers
Aliphatic isocyanates

While olefin metathesis as a catalytic transformation is well know, the catalysts were limited in applicability to the transformation of hydrocarbon substrates due to their intolerance toward polar functional groups. The relatively recent discovery and commercial availability of functional group tolerant metathesis catalysts by Grubbs *et al.* made them potential candidates for ethenolysis of FAME's.

The objective of this project was to evaluate and develop the Grubbs family of catalysts for the ethenolysis of FAME's and to optimize process conditions for this conversion. Hence, catalyst development and catalyst/process evaluation were pursued on parallel paths. Our

target turnover for the Grubbs catalyst was estimated to be ca. 75,000 based upon assumptions catalyst cost and no catalyst recovery and recycling.

Infrastructure Development: Early work on this project involved establishing infrastructure so that the process and catalyst development could take place. This involved the construction and commissioning of a bank of glass reactors with limited pressure capabilities as well as the construction and commissioning of metal reactors capable of operation at elevated ethylene pressures (i.e. > 60 psig). In addition, analytical methods were developed so that the reactions could be quantitatively evaluated.

Process Chemistry: A series of known and potential impurities in our methyl oleate feed were examined for their impact on catalyst performance. Included in this series were methanol, ethyl vinyl ketone, 1-pentene-3-ol, water, methyl linoleate, and methyl linolenate. These materials were examined to not only understand the feed purity necessary for obtaining the maximum catalyst performance but also to identify the primary feed poisons/inhibitors and their role in affecting catalyst activity. All of these material resulted in a decrease in catalyst performance and controlling their levels prior to reaction will be necessary to maximizing catalyst performance. Their impact in order of increasing effect on catalyst performance is: methyl linoleate, methyl linolenate (minor impact) < ethyl vinyl ketone < 1-penten-3-ol, water (significant impact).

The impact of hydroperoxide impurities on the bis(tricyclohexyl)benzylidene ruthenium dichloride (Grubbs 1) was examined. Levels of <0.2 meq/kg hydroperoxides were necessary to maximize conversions in batch reactions, and no significant activity was observed at > 3 meq/kg. The impact of hydroperoxides on the [tricyclohexylphosphine[1,3-bis(2,4,6-trimethylphenyl)-4,5-dihydroimidazol-2-ylidene] benzylidene ruthenium dichloride] (Grubbs 2) was examined as this complex is reported to be more tolerant of polar impurities. Although overall turnovers were lower than with the Grubbs 1 catalyst, this system was much more tolerant of peroxides showing activity in the presence of >100 meq/kg hydroperoxides.

Evaluation of alumina activation procedures was accomplished to ensure the complete removal of hydroperoxides. The goal was to generate a highly purified methyl oleate free of hydroperoxides to measure the maximum possible turnovers for the Grubbs 1 catalyst in batch ethenolysis reactions. Alumina was treated at 480 °C under vacuum for 14 h prior to use, and the treated methyl oleate was reacted at methyl oleate/catalyst ratios of 17100, 51593, and 103,000:1. The resulting turnovers observed after >18 h run time were 10700, 12833, and 16069 respectively. These results were not significantly different from alumina activated between 200 -300 °C, and so alumina activations can be carried out at these lower temperatures. It should be noted that these activity levels are well below what we believe is needed for commercial viability.

Kinetic Model for Grubbs Catalysts: Investigation of the kinetics of each of the elementary steps (Figure 3-2) of the ethenolysis of methyl oleate (MO) and the impact of process variables on the kinetics of this reaction with Grubbs 1 was undertaken. Decene-1 (DC) and methyl decenoate (MD) coupling, as well as methyl oleate homometathesis, have been examined both at high catalyst loadings (to identify equilibrium constants) and at lower loadings (to examine detailed kinetics). In addition, the impact of catalyst concentration, temperature, pressure,

cis/trans olefin content, and product inhibition was also examined. This data was compiled to develop both a semi-empirical model for process development and a mechanistic kinetic model.

Figure 3-2. Elementary Steps in the Direct Ethenolysis of Methyl Oleate
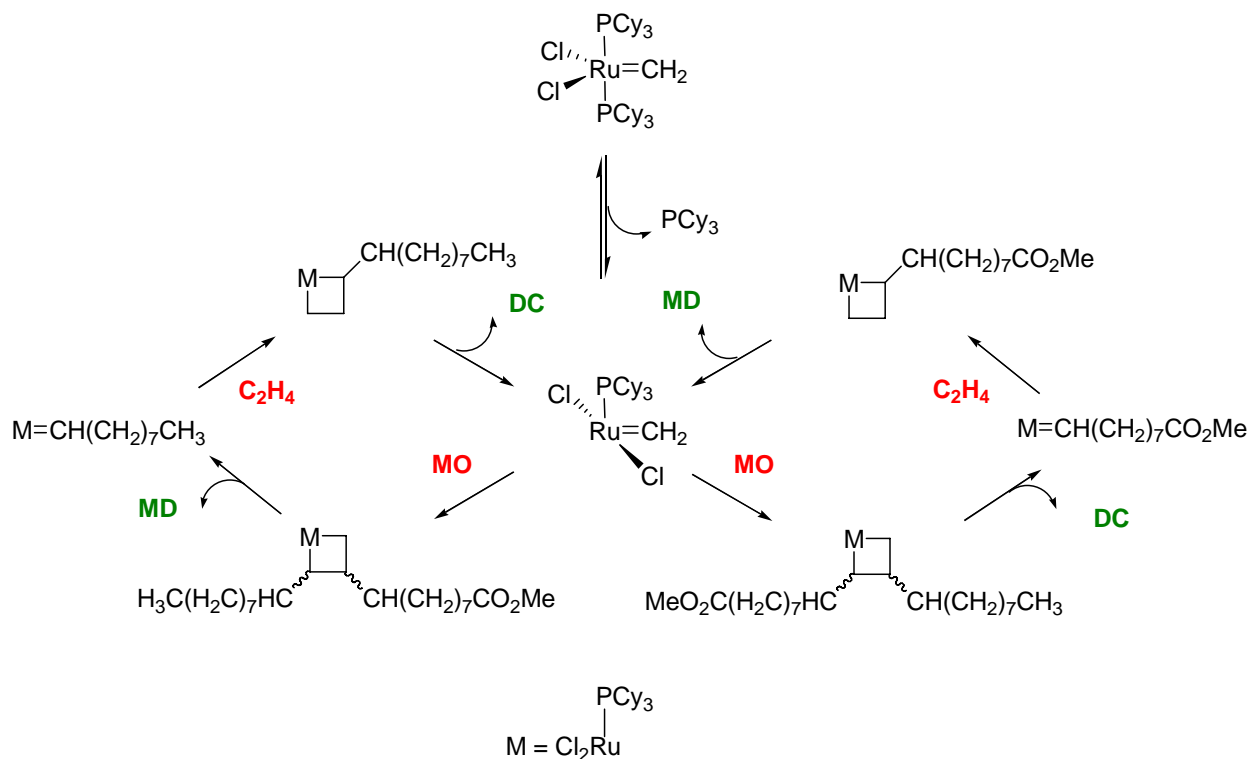
Routine catalyst turnovers under optimized feed and process conditions were found to be ca. 12,000 for direct ethenolysis of methyl oleate using the Grubbs 1 catalyst. This is far short of the target turnover for the Grubbs catalyst of ca. 75,000. Kinetic modeling demonstrates that the low activity results from two phenomena: (1) product inhibition and (2) catalyst deactivation. Product inhibition can be rationalized since the products, being alpha-olefins are sterically less encumbered (i.e. more reactive) than the internal olefin starting material. As a result, the catalyst performs many degenerate (non-productive) turnovers as the concentration of products increases. In addition the catalyst was found to undergo a decay reaction that is directly related to increasing ethylene concentration. This is a rather unfortunate phenomenon as increasing the ethylene concentration would otherwise be desirable to drive the reaction towards products according to fundamental equilibrium considerations.

The kinetic investigation of the Grubbs II catalyst systems revealed that the inefficiency of the Grubbs II catalyst for ethenolysis is related to the bias for alkylidene over methylidene resulting in a low methylidene concentration which rapidly diminishes upon formation of product. As a result there is a pronounced preference for homometathesis with the Grubbs 2 catalyst. Of the commercial catalysts, Grubbs 1 was determined to be the better option.

As indicated above, results from kinetic/molecular modeling and lab reactor runs demonstrated that, even in the absence of catalyst decomposition, the effective maximum
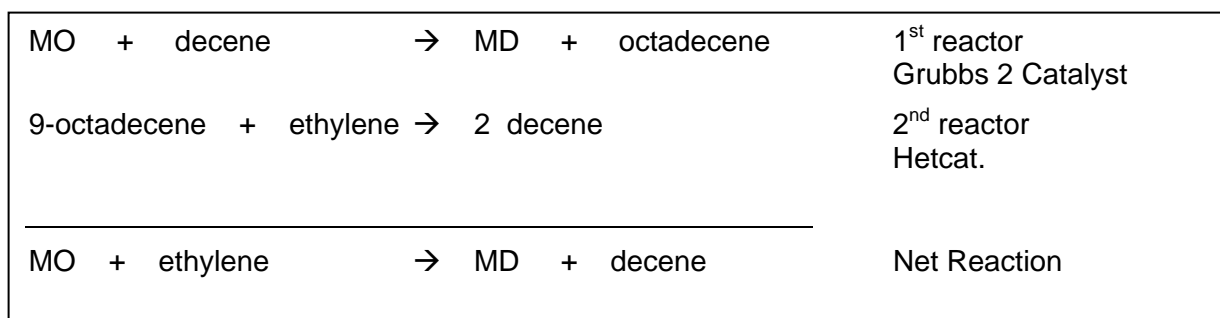
conversion for methyl oleate ethenolysis decreases with decreasing catalyst loading due to the impact of product inhibition, thereby limiting total turnovers below our target. In a **successful process, therefore, use of Grubbs 1 catalyst will require (1) limited conversion be targeted to maximize this initial catalyst productivity, and (2) a method for catalyst separation/recycle.** To this end, catalyst separation strategies of selective membranes, ionic liquids, and polar/apolar biphasic catalysis were reviewed for applicability in the ethenolysis of fatty acid esters. Selective membrane separation was viewed as the best option.

Process Development: Our approach to membrane separation involved permeating a portion of the products/reactants through an appropriate membrane, while retaining the active metathesis catalyst on the reaction side. A number of literature references indicated that this type of separation has been achieved with a Wilkinson hydrogenation catalyst and a hydrophobic membrane. Initial proof-of-concept experiments were performed with a flat plate membrane in a 3" diameter cell. A PVDF RO (reverse osmosis) membrane (MPF-50 MWCO of 700 from Pall Corp) was evaluated in the small lab separator at 4 atmospheres (trans membrane pressure). Operating in a filtration mode, the permeate contained 8.9 ppm Ru while the retentate contained 61 ppm Ru. This represents and 85% retention of catalyst.

Efforts to recover catalyst using ceramic membranes manufactured by Pall Exekia were also explored. A fully instrumented experimental apparatus was assembled to test tubular ceramic membranes at the high tangential velocity (5 m/s) that they require. A heat exchanger in the recirculating loop allows the hydrocarbon solutions to be tested at higher temperatures which reduce liquid viscosity. Experiments were first conducted with purified methyl oleate, then a Grubbs I catalyst solution was added to determine the selective permeability of these membranes for different chemical species. A ceramic membrane rated at 5 nm passed no methyl oleate at 5 bar mean transmembrane pressure (TMP), while a 50-nm membrane produced a steady-state permeate flux of 48 L/m^{2}hbar. Subsequently, a 20-nm membrane was tested first with methyl oleate alone and then with catalyst added. Permeate flux for methyl oleate alone was lower but showed less decay yielding a steady-state value of 32 L/m^{2}hbar. Sufficient catalyst solution was added to achieve ruthenium (Ru) concentration of 20 ppm in methyl oleate. The resulting initial flux of permeate, 23 L/m^{2}hbar, was good, and it declined slowly as concentration of the catalyst in retentate increased. Visual observation and color suggested that significant but incomplete separation of the catalyst was accomplished. Analytical testing showed a 79% retention of ruthenium was achieved in this membrane separation study.

As indicated above, the Grubbs 1 catalyst was found to undergo a decay reaction that is directly related to increasing ethylene concentration. This fact limits the amount of ethylene that can be used to achieve high conversions of methyl oleate. A dual catalyst/dual reactor process option (i.e. indirect ethenolysis) was considered to circumvent this problem. The concept is best described by the chemical conversion steps shown in Figure 3-3.

Figure 3-3. Indirect Ethenolysis

MO + decene	→ MD + octadecene	1st reactor Grubbs 2 Catalyst
9-octadecene + ethylene →	2 decene	2nd reactor Hetcat.
MO + ethylene	→ MD + decene	Net Reaction

MO = Methyl Oleate,
MD = Methyl Decenoate

The indirect ethenolysis process utilizes the advantages of each type of catalyst while avoiding their weaknesses. The first reactor avoids the use of ethylene and this avoids more rapid catalyst decay for Grubbs 1 catalyst. The Grubbs 2 catalyst on the other hand, is very well suited for the cross metathesis which does not use ethylene. In the second reactor, the ethenolysis of octadecene involves no polar functional groups such that traditional catalysts can be used and an over-pressure of ethylene can be employed to drive the reaction. **The catalyst turnover for the 1st reactor of this new process was found to be 70,000 on a consistent basis. This approaches the targeted catalyst performance we originally established for commercial viability.**

Catalyst Development: As indicated above, the Grubbs 1 catalyst was found to undergo a decay reaction that is directly related to increasing ethylene concentration. It was anticipated that if we could elucidate the mechanism of decomposition, it might be possible to design a catalyst that would circumvent that decomposition route. This would also allow for the use of higher ethylene concentrations

To address the mechanism of catalyst deactivation, the Grubbs I catalyst was treated with ethylene at 60 psig for 5 days to ensure decomposition of the methylidene intermediates, and these decomposition products were analyzed by NMR spectroscopy, mass specrometry, and XPS. The Grubbs I catalyst shows the presence of $Cy_3PMe^+Cl^-$ as a primary decomposition product. This has been subsequently verified by synthesis of this phosphonium salt by known methods. Decomposition of Grubbs I under isotopically labeled ethylene, C_2D_4, results in the incorporation of deuterium into the methyl group, pointing to the carbene moiety itself as its source for this carbon.

Catalyst development efforts were focused primarily on (1) synthesizing new catalysts to screen for increased activity and lifetime; (2) measuring lifetime of synthesized catalysts; (3) utilizing molecular modeling to identify catalysts which will limit product inhibition; and (4) evaluating process options for catalyst separation/recycle. In order to elucidate the structure/activity relationships in the Grubbs family of catalysts a campaign of 1680 experiments was completed in the combinatorial screening reactors. Focus areas and results are described below.

   a) Process variable optimization of 12 metathesis catalysts was accomplished examining temperature, pressure, olefin/catalyst concentration, and solvent. The optimum conditions for each of these systems were identified
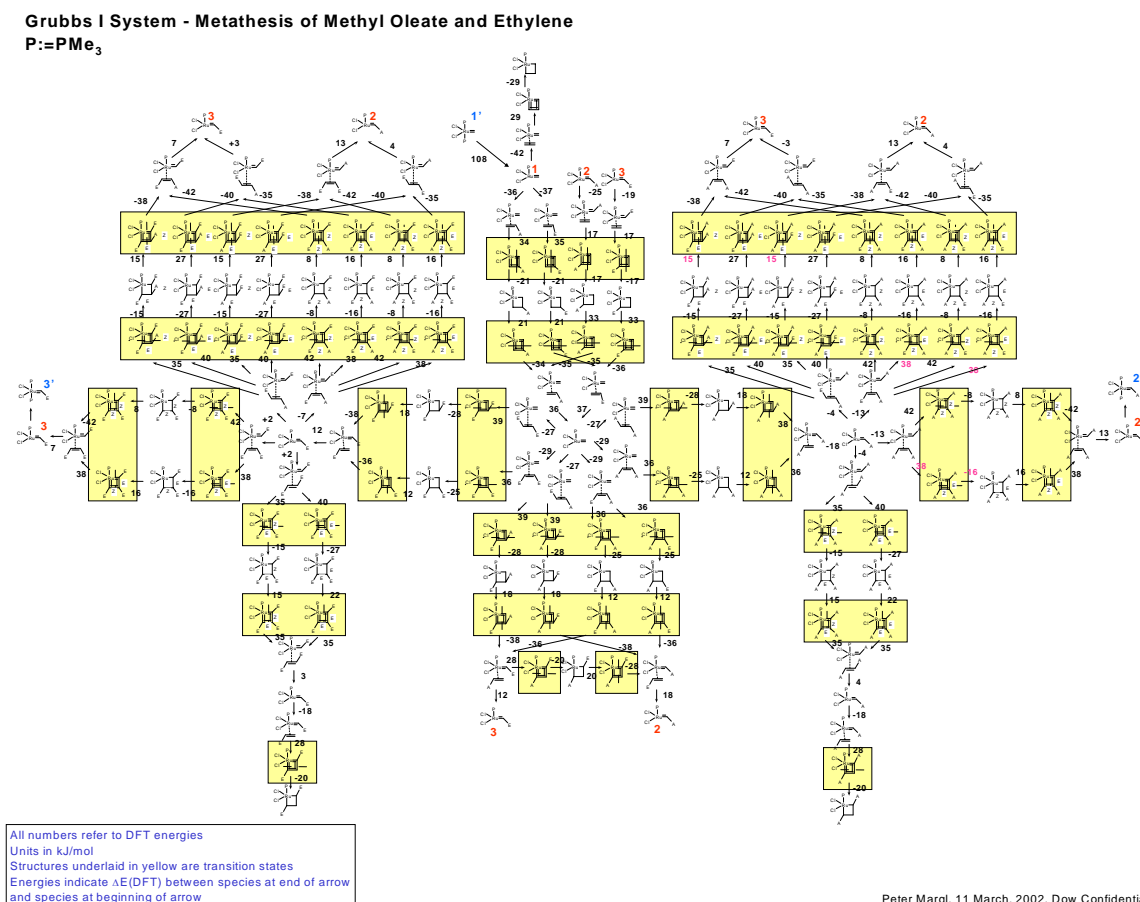
b) A series of 192 *in situ* prepared catalysts were examined using the X_2GIC and X_2GIIC catalysts and a series of 32 phosphines, phosphites, nitriles, pyridines, ethers, and alcohols. A number of improvements were identified.

c) A collection of 32 activators were examined in an effort to further raise catalyst productivity. Up to a 20% increase in catalyst productivity was observed for several catalyst/activator combinations.

d) Using the optimized catalysts/process variables identified in the earlier experiments, a series of experiments on the ethenolysis of methyl ricinoleate and methyl ricinoleate acetate were undertaken. For methyl ricinoleate, a 7% conversion to the desired terminal olefins was obtained using the Grubbs 1 catalyst at 60 psig ethylene, room temperature, and an olefin/catalyst ratio of 4140:1. This was increased through our combinatorial screening .

Catalyst lifetime and turnover measurements were conducted for the previously prepared complexes under continuous conditions for ethenolysis using the model reaction, ethenolysis of cis-2-butene to give propylene. The impact of the ethylene/2-butene ratio, the total flow of each gas, the reactor residence time, and the catalyst concentration has been elucidated for both the Grubbs I/II catalysts resulting in the demonstration of > 16,000 turnovers at 40 °C for Grubbs I. At 40 °C, lifetimes ranged between 2 to 8.5 h for the cross-metathesis reaction, with the highest stability found for the dibromide derivative of Grubbs 1. Interestingly, the dibromide and diiodide versions of Grubbs 2 exhibited the shortest lifetime.

New complexes were prepared for a second round of high throughput evaluation. A total of 17 systems were evaluated in a combinatorial chemistry campaign (> 1300 experiments) and a number of these catalysts were further evaluated in our continuous lifetime reactor systems. Unfortunately, no systems and/or reaction conditions were identified which demonstrated improved performance over the Grubbs 1 catalyst system or our previously identified diiodide-Grubbs 2 and Hoveyda 2 complexes. In addition, the stabilities of a number of these complexes were evaluated in the continuous lifetime reactors and found to have equivalent or lower lifetimes compared to the Grubbs 1 catalyst.

Throughout the course of catalyst development, molecular modeling was used extensively to aid and guide the bench chemists in their pursuits. Ground-state thermodynamics were calculated for each of the elementary steps in the catalytic steps cycle for the commercial Grubbs catalysts. Density functional theory at the B3LYP/LACVP** level was used to predict the thermodynamics and kinetics for the entire oleo metathesis cycle with the PMe_3 /Grubbs 1 system. A detailed diagram is given below (Figure 3-4, must be expanded to be readable). The results showed a rather shallow energy change through the course of the catalytic cycle for both catalysts. Comparing PMe_3 data to PCy_3 data showed that sterics is a major driver for phosphorus binding to the active site, which is unsurprising. The computational data matches all experimentally known facts. Olefin binding energetics and methylidene/alkylidene equilibrium energetics seem the main keys to manipulating catalyst behavior. Candidate catalysts were screen for a similar profile using molecular modeling.

Figure 3-4. Energetics and thermodynamics of the PMe$_3$-GrubbsI system calculated at the B3LYP/LACVP** level of theory. Energies in kJ/mol.

Grubbs I System - Metathesis of Methyl Oleate and Ethylene
P:=PMe$_3$

All numbers refer to DFT energies
Units in kJ/mol
Structures underlaid in yellow are transition states
Energies indicate ΔE(DFT) between species at end of arrow
and species at beginning of arrow

Peter Margl, 11 March, 2002, Dow Confidential

In addition to ground state energetics, transition state energetics were calculated for key elementary steps including that for catalyst decomposition. Molecular modeling anticipated the decomposition pathway which subsequently was demonstrated experimentally.

Molecular modeling was used extensively in the high throughput (combinatorial) experiments as well. Prior to the high throughput runs, molecular modeling was used to ensure that a diverse set of ligands were chosen in the screen. Molecular modeling was also used to analyze the results of a combinatorial screen by establishing QSAR equations (quantitative structure/activity relationships) to predict selectivity and turnovers for metathesis catalysts based on high throughput data and developing new catalyst motifs/directions that could give enhanced productivity.

Redirected Catalyst Development: During the course of the project intellectual property issues arose due to the partnership agreement formed between Materia and Cargill to develop Grubbs catalyst for the conversion of seed oils to chemical feedstocks. At this point, no more Materia catalyst was commercially available to Dow and a research license was not granted. This situation, coupled with our inability to improve Grubbs catalytic performance for the direct ethenolysis of methyl oleate, caused a shift of our research direction away from these systems toward the discovery of non-Ru based olefin metathesis catalysts.

Task 4 Utilization: Life Cycle Analysis, Modeling and Market Research

In this task, the intermediates made from current seed oils were modeled for future development of new transgenic seed crops. The life cycle analysis aided in understanding the economic and environmental impact of a new plastic from a renewable resource. Early stage market research was completed by NBID, an external to Dow company that specializes in early stage product exploration.

4.1 Modeling Studies

Dow compared the sustainability of flexible foam polyols made through conventional petrochemical routes to a product with the same performance and attributes but made largely from renewable materials (castor oil or soy oil) by doing a partial LCA that follows closely the ISO standards.   Sub task 4.1 was included in the full life cycle analysis. The life cycle work was not just for castor but for also for a soy polyol as well as comparison to the traditional chemical route. The results will be reported in 4.2.

The functional unit was 1 kg of un-blown polymer solution, capable of being used to make a flexible polyurethane foam identical to current petrochemical-derived materials. The solution is made by polymerizing propylene oxide (PO) or a seed oil derivative with an initiator, made from glycerin and ethylene oxide (EO). EO is a petrochemical used in all three routes, so the product is not 100% bio-based, but has identical mechanical properties to conventional foam. Three primary routes to this product were compared. Petrochemical foam polyol solution, as described by the APME eco-profile, was used. The European industry average data was chosen for the reference case in this study. Soy-oil derived foam polyol solution was the second route and castor-oil derived foam polyol solution was the third.

4.2 Life Cycle Analysis

Three basic routes were compared. Figures 4-1 to 4-5 present the calculated life cycle inventories (energy, mass & water) and impacts (greenhouse gas and acid gas emissions) for the three compared routes to make a functionally identical polymer solution for flexible foam polyurethanes: a petrochemical route (APME), soy oil route, and a castor oil route, based on irrigated farming in Texas and production of useful fertilizer and oil from castor seeds (no low value or useless co-product).

Figures 4-1 and 4-2. Gross energy intensity and raw material unit ratios for three polyol options
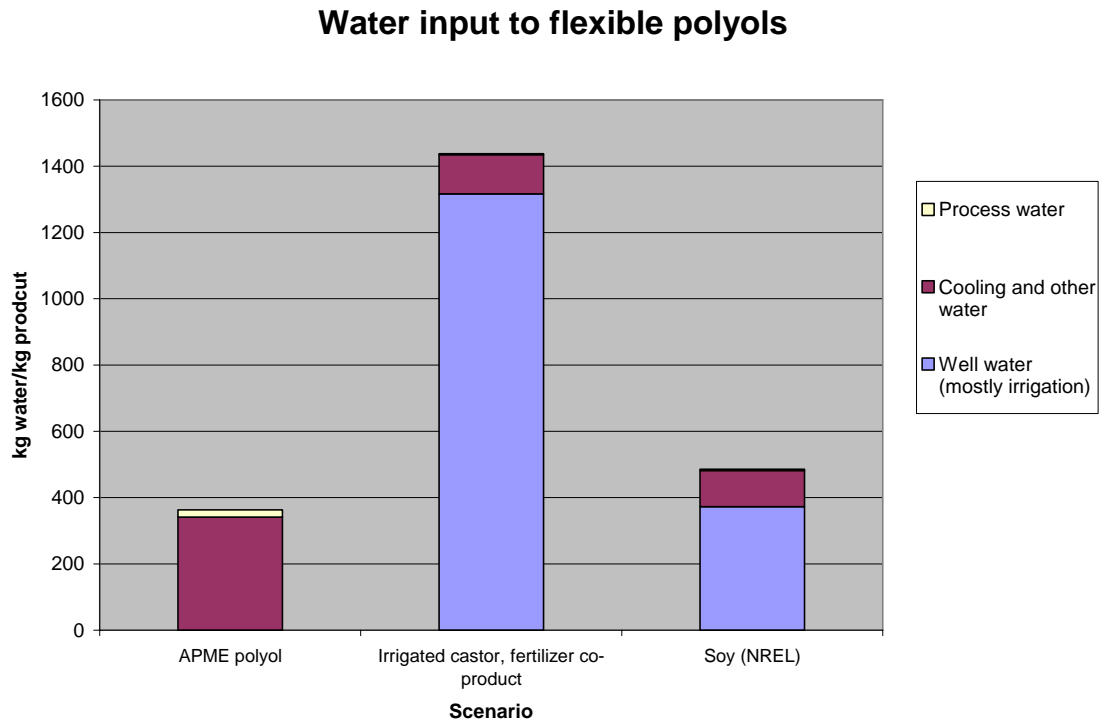
Gross Energy Intensity of Flexible Polyols
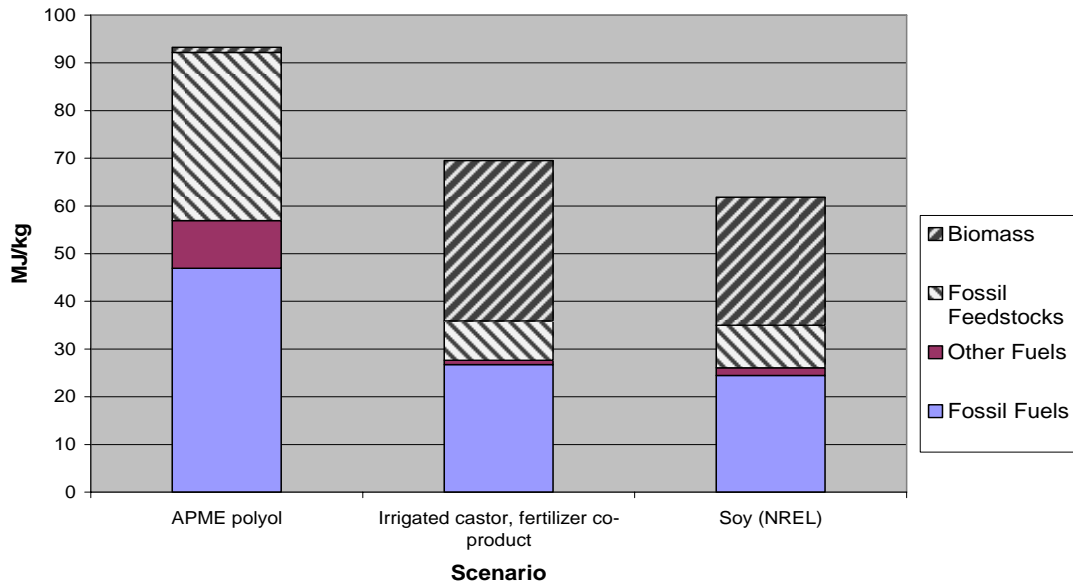
Raw Material Unit Ratios for Polyol options
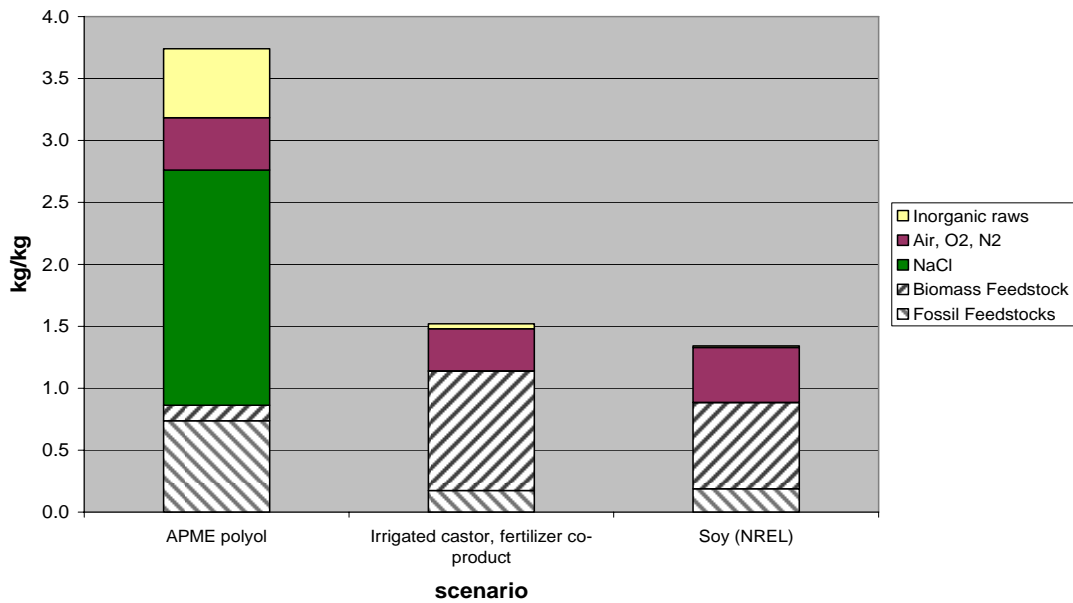
Figures 4-3 and 4-4. Water intensity and greenhouse gas emissions for three polyol options.

Water input to flexible polyols

Greenhouse Gas Emissions for Polyols
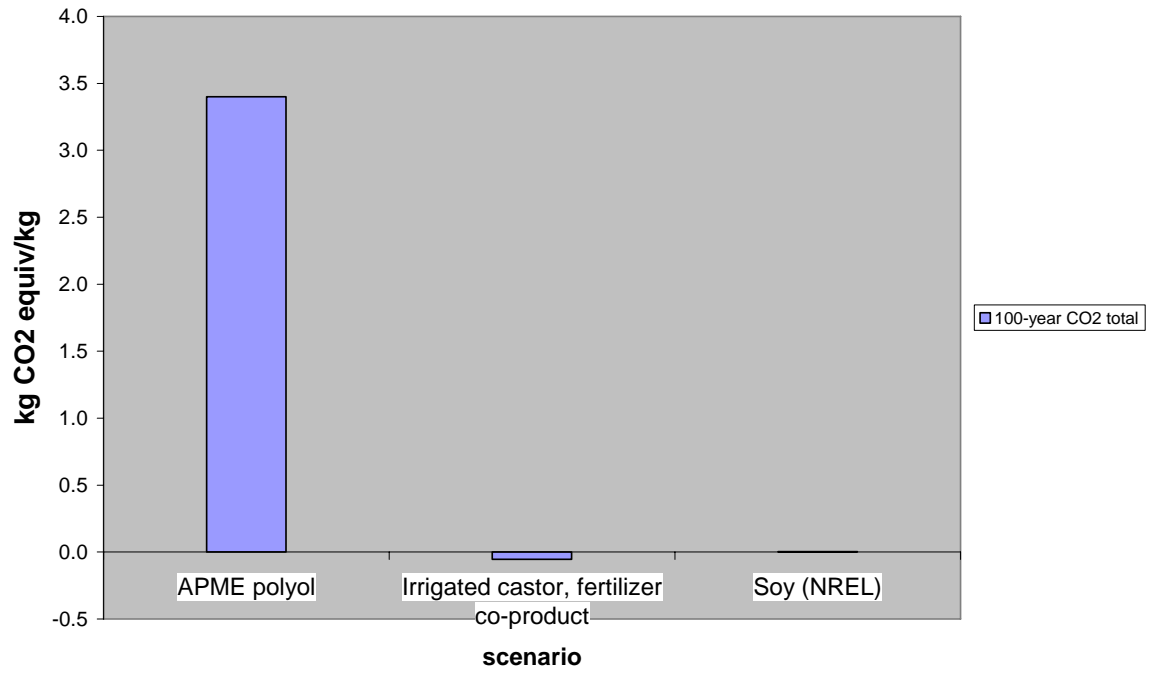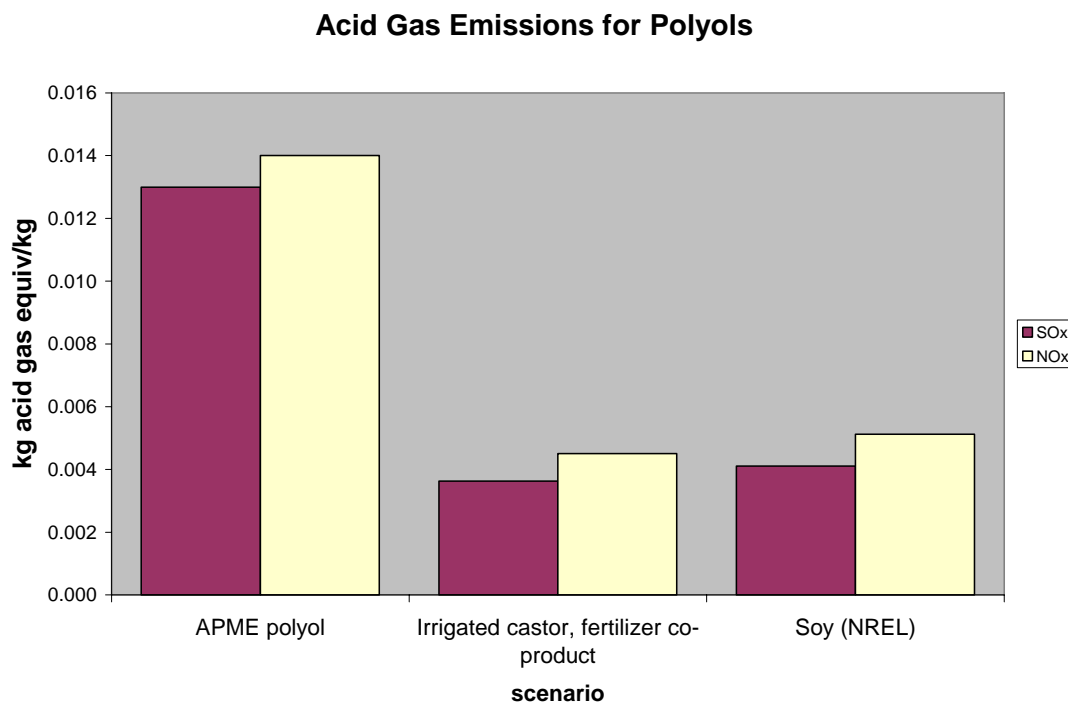
Figure 4-5. Acid gas emissions for three polyol options.

Acid Gas Emissions for Polyols

The partial life cycle calculations demonstrate some of the significant potential benefits for bio-based polyols. Both seed oil polyols have lower gross energy (including biomass) than the European average. More importantly, polyols made from soy used only 41% of the fossil resources as the European average, and polyols from castor used only 43%. Use of either seed oil would decrease use of fossil resources. For this cradle-to-gate study (which neglects the use and disposal of the polyol) there were essentially zero net greenhouse gas emissions for polyol made from either oil. The seed oil polyols also have a reduced regional air impact, with 32% and 37% of the SO_x and NO_x emitted from the soy route compared to the European average, and 28% and 32% for castor.

Seed-oil polyols decouple polyol production from chlorine-based chemistry, as shown in Figure 4-2 by the lack of NaCl as a raw material unit ratio as compared to the European average. Some of the current petrochemical routes do not used chlorine.   Although the safe use of chlorine is certainly possible, its use requires a relatively large capital infrastructure, and chlorinated organic by-products can be difficult or expensive to destroy. If one faces capacity limits in chlorine supply or chlorinated organic destruction, then non-chlorine chemistries can be more economical.

Water use for seed oil polyols is 134% (soy) to 396% (castor) of the European average. Most of the increase is due to water used in irrigation. It is possibly not of the same source and quality as the water used in the chemical production plant.

43

Ecological footprint of the routes

The LCA results for the different routes and scenarios were described above primarily through a series of graphs. They can also be presented as numbers in a table, or by using an "ecological footprint". An example of an ecological footprint is well-known today from work at BASF[1], although it can trace its origins to the "eco-compass" concept created by Dow[2]. Figure 4-6 below is an adaptation of this concept for this study. The figure is a "radar plot". The five axes represent five key sustainability metrics that come from the LCA calculations: gross energy intensity, fossil resource use, mass intensity, total water use, and gas emissions. For each axis, the worst option of all the options evaluated gets a score of 1.0; the values on that axis for the other options are normalized to a number between 0 and 1.0.. For example, the total water use of polyols made from irrigated Texas castor with no co-products was 1438 kg/kg. For the soy-based polyol, the total water was 485 kg/kg, for a normalized score of 0.34.   The normalization of gas emissions was more complicated. Each emission was normalized first to the worst to put the values in the same numerical order of magnitude, then they were added with an assumed weighting (50% from normalized greenhouse gas emissions, 25% each from normalized SO_x and NO_x emissions) and renormalized.

The information content of this one figure is very high – and it helps to view it in color! But many of the insights from Figures 4-1 to 4-5 are possible to draw from this one figure. The decreased mass intensity, fossil resource use, and gas emissions for all the various bio-based scenarios are clearly significantly less than the petrochemical route (APME).   Higher water use – primarily for irrigation – is one of the environmental costs to trade-off with the benefits in the other dimensions. The differences in overall energy intensity are not as great among the options as the differences on the other axes, since the energy dimension does include biomass energy.

---

1. Shonnard, D. R., A. Kicherer and P. Saling, Industrial Applications Using BASF Eco-Efficiency Analysis: Perspectives on Green Engineering Principles, *Environmental Science & Technology*, **37** (23), 5340-5348, 2003
2. Fussler, Claude, with Peter James, Driving Eco-Innovation: A breakthrough discipline for innovation and sustainability, Pitman Publishing, London, 1996, ISBN 0-273-62207-2

Figure 4-6. Ecological footprint for polyol routes – petrochemical, soy, castor.

Ecological Footprint for Polyols
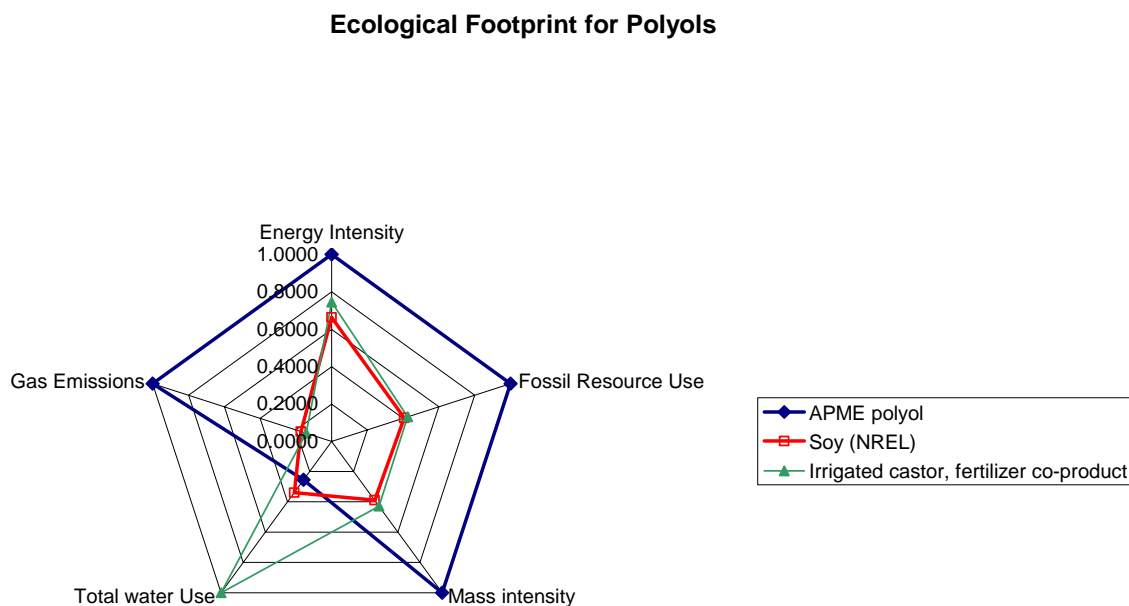
Quantitative impact of co-product assumptions

An essential feature of LCAs conducted according to ISO standards is consideration of the impact of key assumptions on the results. The results described above were based on several key assumptions about castor co-products and their value and also castor farm inputs.

The "base case" assumed all non-oil parts of the seed are useful as an organic fertilizer and thus share equally the per kg burdens from farming and seed processing. This assumption was made as it would increase the revenue stream to the farmer and would make castor a more likely commodity crop in the future. But it is not the case today. Currently, only the oil has significant value – the meal and hulls do not (allergens in the meal and toxic ricin in the hulls hamper the use of these, especially as feed, although these can be detoxified by steam). The impact of this assumption is explored in the following graphs. These are similar to the ones above, and for reference include the three bars shown in the previous graphs: the APME data on the far left, Dow soy polyol on the far right, and the base case castor scenario (irrigated farms in Texas, where the fertilizer co-product includes hulls), in the column adjacent to the soy data. The two added bars are polyols made from castor from irrigated Texas farms, but where the useful (steamed) meal co-product does NOT contain the hulls (the center column), and the scenario using irrigated Texas castor with NO byproduct value (the column adjacent to the APME column). This last scenario is the closest to the situation today. The two added scenarios increase the proportion of the farm and processing inputs from 1.0 in the base case to 1.33 in the case without hulls and 2.22 in the case without meal. They also use the current

hexane extraction process to recover the castor oil and steam in the meal co-product case to detoxify it. The inputs & emissions were estimated using engineering principles from a published flowsheet[3].

Figure 4-7. Impact of co-products on gross energy.

Gross Energy Intensity of Flexible Polyols
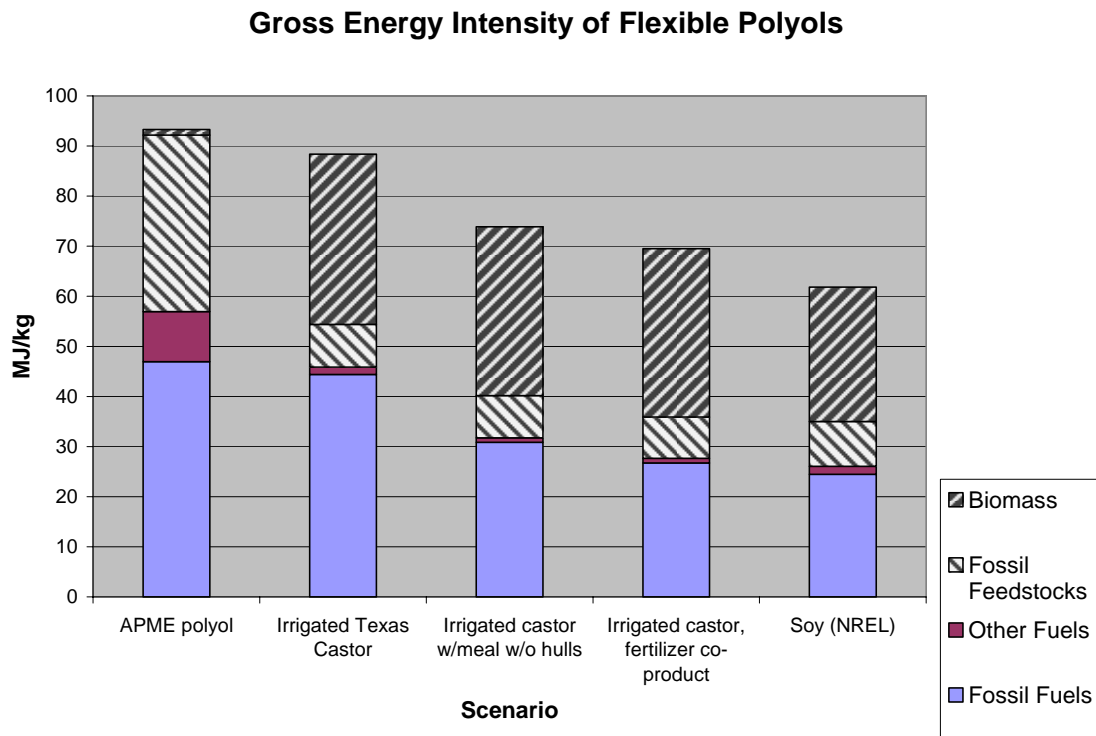
3. Weiss, E.A., Oilseed processing, in *Castor, Sesame and Safflower* pp 761-793, 1971; available at http://www.wsu.edu/~gmhyde/433_web_pages/433Oil-web-pages/castor/castor-oil2.html via the Jayant Oil Mills web page

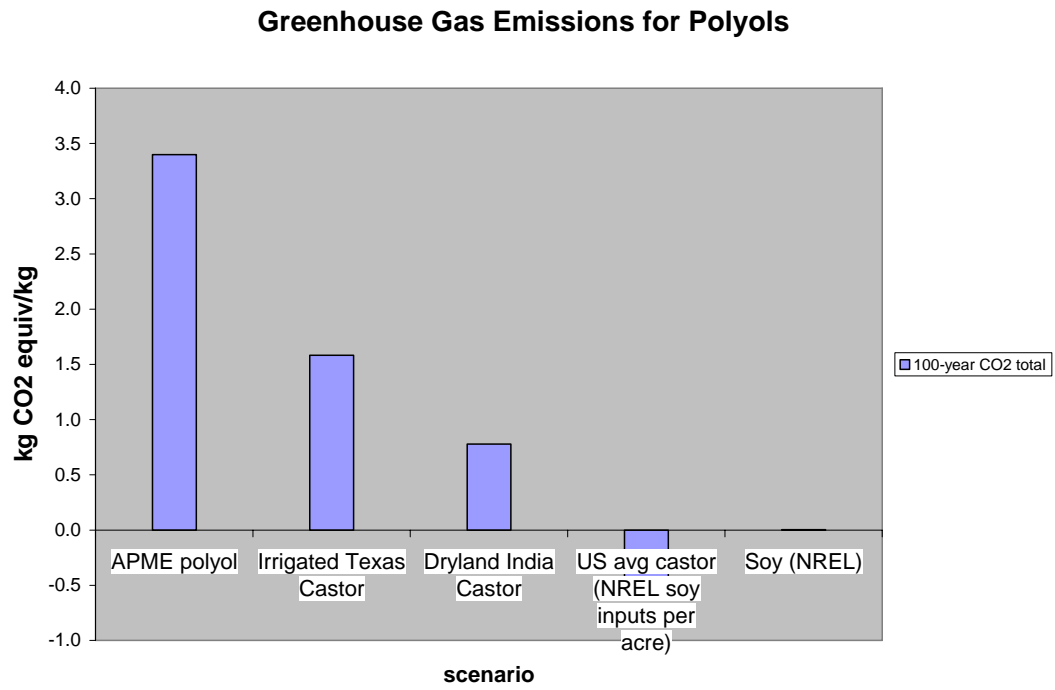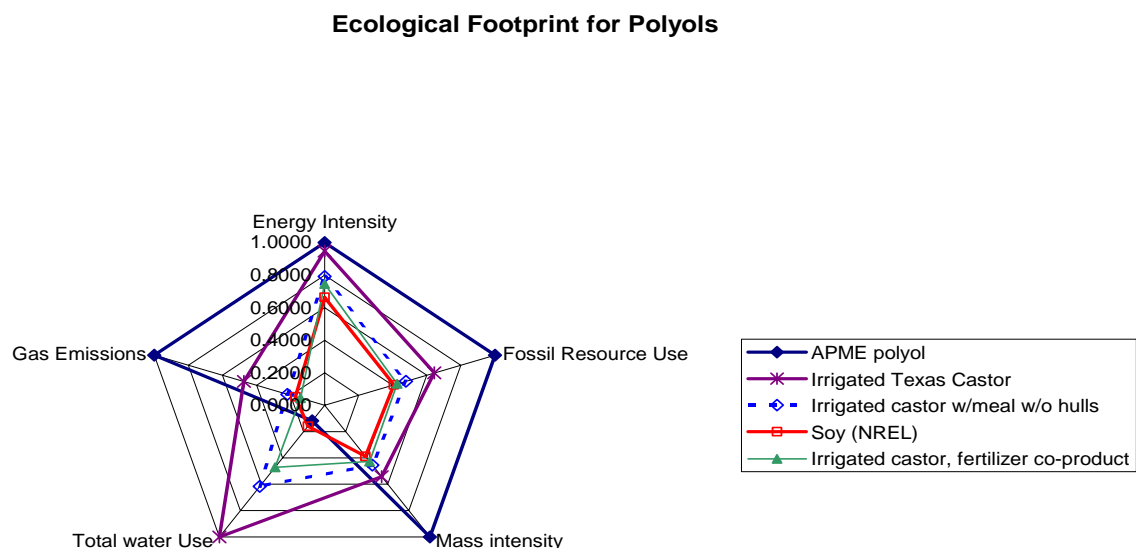**Figure 4-8.** Impact of co-products on greenhouse gasses.

Greenhouse Gas Emissions for Polyols

Figure 4-9. Ecological footprint for polyols with different castor co-product assumptions.
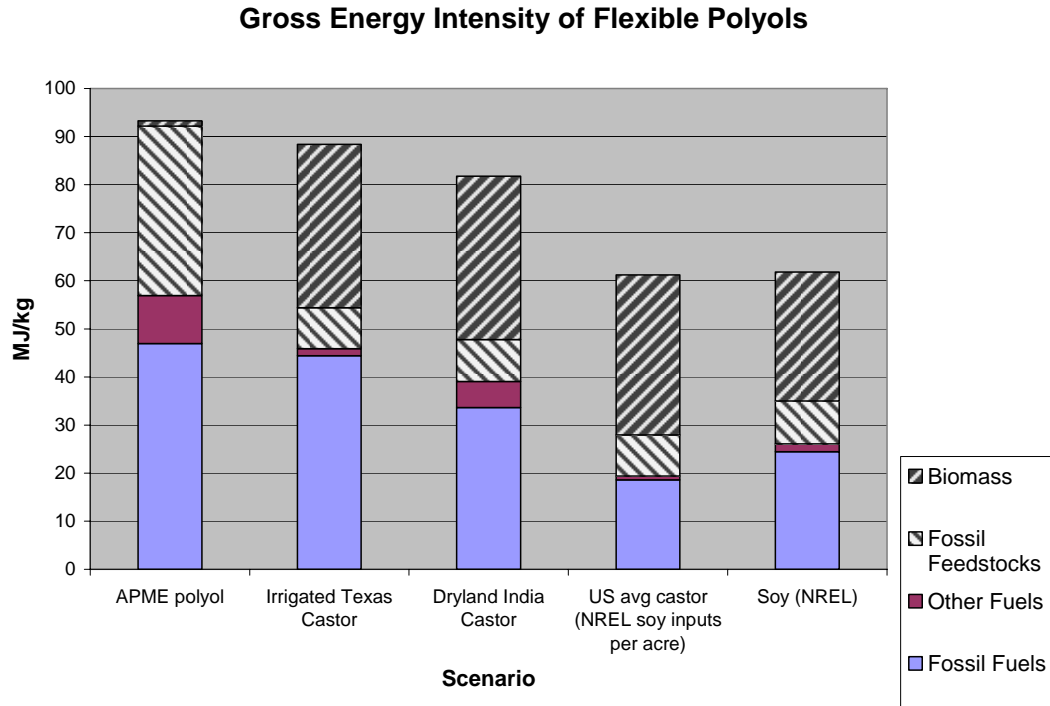
Ecological Footprint for Polyols

These graphs clearly show the tremendous importance of the assumptions about co-products. For all the metrics, the case with meal but without hulls yielded results between the base case (fertilizer including hulls) and the case with no useful by-product, so the following discussion will describe the results going from the one extreme (all products useful – oil and fertilizer) to the other (no useful by-products). The gross energy of the castor polyol increased from 75% to 95% of the APME European average, with the fossil resource use increasing from 43% to 64%. Water use increased from 396% to 840%. Greenhouse gas emissions increased from -2% to 47% of the average. Acid gases increased from 28% to 44% for SO_x and from 32% to 52% for NO_x. These are all significant impacts. In all cases, the inputs to the farm per acre planted and the many of factory inputs per mass of seeds are unchanged (there are differences for co-product processing), so the total burden to the environment is nearly the same. The major change is in the sharing of those burdens ranging from complete (base case) to none (no meal – the oil carries it all). There are many ways to do this split ("allocation") in LCA, including by mass (as done here), economic value, energy content, or by changing the system boundaries. The mass basis used here is acceptable, as long as it is clearly stated and the impact of the assumption is shown. The work also shows the inherent benefit to producing useful products rather than waste streams – one gets more benefit for a given quantity of environmental burden.

Quantitative impact of castor farming assumptions

The other key assumption was the model for castor farming. Castor is not yet a mature commercial crop in the US, unlike soybeans or corn. Comprehensive field data do not exist for castor farming. The "base case" for this analysis used farming inputs from Castor Oil, Inc., but with no useful co-products. This is a different reference case for the castor polyol than the one used previously, which did include useful co-products, and was chosen to more clearly show the impact of the farming assumptions. The impact of the farming assumption is explored in the following graphs. These are similar to the ones above, and for reference include the three columns from previous graphs: the APME data on the far left, Dow soy polyol on the far right, and the new reference case castor scenario (irrigated farms in Texas, where the meal and hulls are NOT useful products), in the column adjacent to the APME data. The two added bars are: polyols made from castor from dry-land Indian (Andhra Pradesh) farms, including transport to US Gulf Coast (the center column); and US farming, assuming the same farm inputs per acre as per NREL's soy data (the center-right column) but with the yield of seed per acre being that of castor rather than soy (castor seed has about twice the mass yield per acre as soy beans) This scenario is meant as a "best guess" for castor farming possible in other, less irrigated (more natural rainfall), sections of the country.

Figures 4-10 and 4-11. Impact of castor farming model on gross energy and greenhouse gas intensity.

Gross Energy Intensity of Flexible Polyols

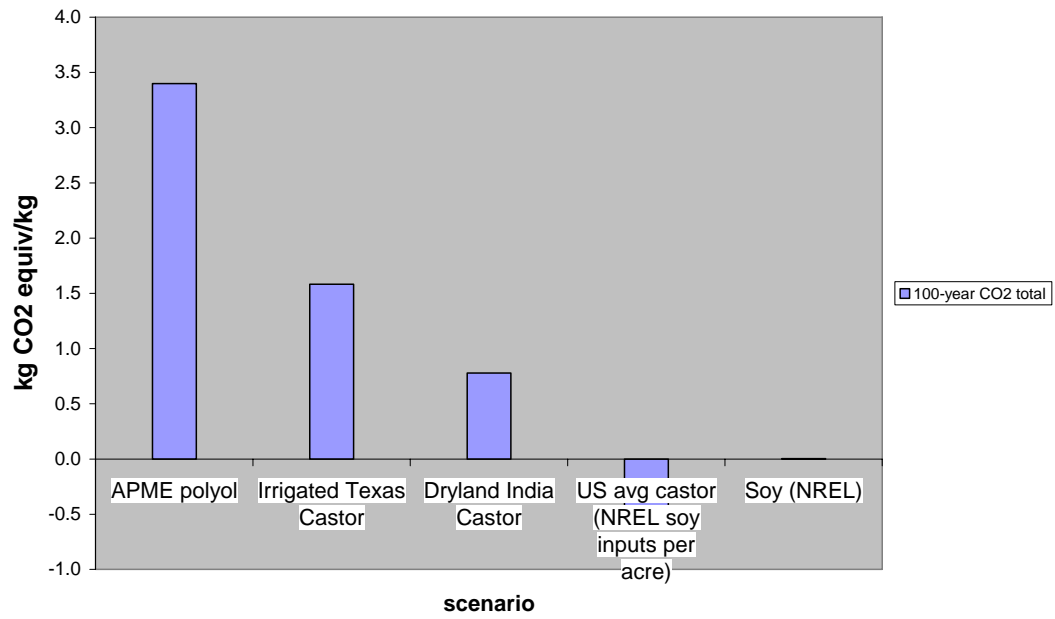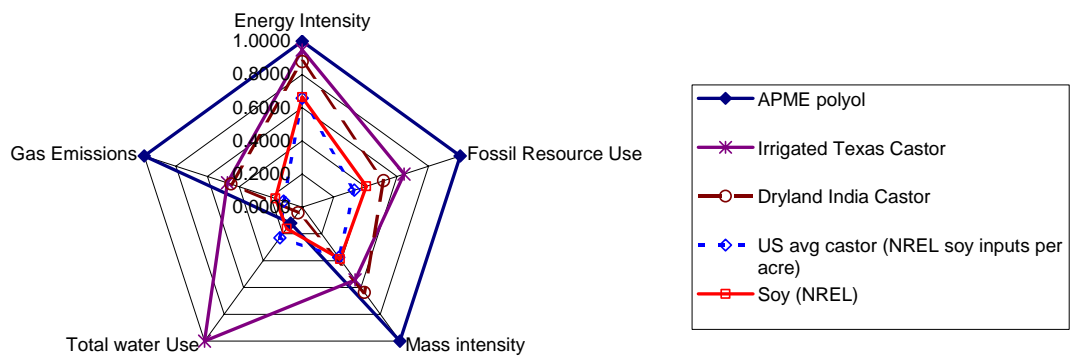## Greenhouse Gas Emissions for Polyols

Figure 4-12. Impact of castor farming model on the ecological footprint for polyols.

Ecological Footprint for Polyols

These graphs clearly show that the choice of farm model is of tremendous importance. The magnitude of the possible benefits for polyols from castor depends on the success and system of farming. Fossil resource use ranges from 33% to 64% of the APME European average data, greenhouse gas emissions range from -13% to 47%, acid gas emissions range from 33% to 52%, and water use ranges from 35% to 840% higher. The assumption about farm inputs changes many impacts by more than two-fold. The impact of the farm model impact is less, but still significant, if co-products are assumed to be useful.

The farming system in India is much different than that of the US, primarily as it is dry-land farming. Two tractor passes were done on the fields (first plough, then harrow plus sow), but harvesting and other work was done by hand or animals (the human & animal input was modeled as a biomass energy input (this assumes they are 100% efficient vegetarians), but only for the work performed, and not other living). The yield was about half that expected in the US. Even though the Indian data had lower mass and energy inputs per acre than the US, the lower yield led to very comparable product gross energy and higher mass inputs on a per-mass-polyol basis.

The major difference between the Texas and US (i.e., soy) models for castor farming is the energy required for irrigation. The total energy input to US soy farming is ~3 MJ/kg-bean, but for Texas castor it is about 13.3 MJ/kg-seed, of which 9.4 MJ/kg is the pumping energy for irrigation. This is a very large burden, which makes farming in other regions – if it can be done with the same inputs as soy beans – much more beneficial. The possible impacts, such as with regard to CO_2 sequestration or biodiversity, for converting non-farm land into farmland was not included in this analysis. It was assumed that unused land was supporting native grasses and thus participates in the natural carbon cycle.

Quantitative impact of soy farming assumptions

An additional source of information on soy cultivation was available but not selected as the base case for this or the prior studies of soy polyol LCA. It is attributed to Olivier Muller at PricewaterhouseCoopers (PWC) and was obtained from Barbara Lippiatt at the US NIST (National Institute of Standards & Technology)[4], as part of the "BEES for USDA" program. The data were also used and described by J. Pollack of Omni Tech International in a life cycle comparison of soy polyols[5]. The PWC description of soybean cultivation has a lower primary energy input per kg of soybeans than NREL's analysis (2.16 vs. 3.14 MJ/kg), but much higher field emissions of N_2O (3.9 g/kg vs. 0). NREL felt that the N_2O emissions data for soybeans were too widely variable to know the "true" value, and were likely not distinguishable from fallow land. PWC used the IPCC methods to estimate N_2O emissions from soy fields. The Omni Tech work also differs from the NREL approach in how to assign the CO_2 uptake from air. NREL does this based on the carbon content of the oil used, and does not take credit for CO_2 used in roots, stalks, stems, and meal. NREL assumed that biomass used in food, crop residue & fuel are merely different pathways for the same net-zero CO_2 flux between plants and air. It also carries an implicit assumption that land not used for soy agriculture would be used to grow biomass of some sort, with a similar carbon flux. Omni Tech assigned CO_2 uptake during plant

4. E-mail received from Barbara Lippiatt, NIST, 14 April 2004
5. Pollack, J., Soy vs. Petro Polyols, A Life Cycle Comparisoin, *GPEC 2004 Paper Abstract "37*, Omni Tech International, Inc, Midland, MI, 2004

growth, and also (apparently) during use of the oil in the polyol. This overstates the biomass CO_2 in the polyol compared to the NREL approach.

The NREL model was used as the base case for soy farming because it is better documented and their assumptions seem justifiable. But the choice of soy farming model has a significant impact on the soy polyol LCA, especially the greenhouse gasses. This is shown in the following figures, the ecological footprint and the greenhouse gas bar graph. The NREL model used much more irrigation water, but produced much less N_2O so had less greenhouse gas emissions. The other metrics for the polyols based on the two farming models are very similar.

Figure 4-13. Impact of soy farming model ecological footprint of polyols.

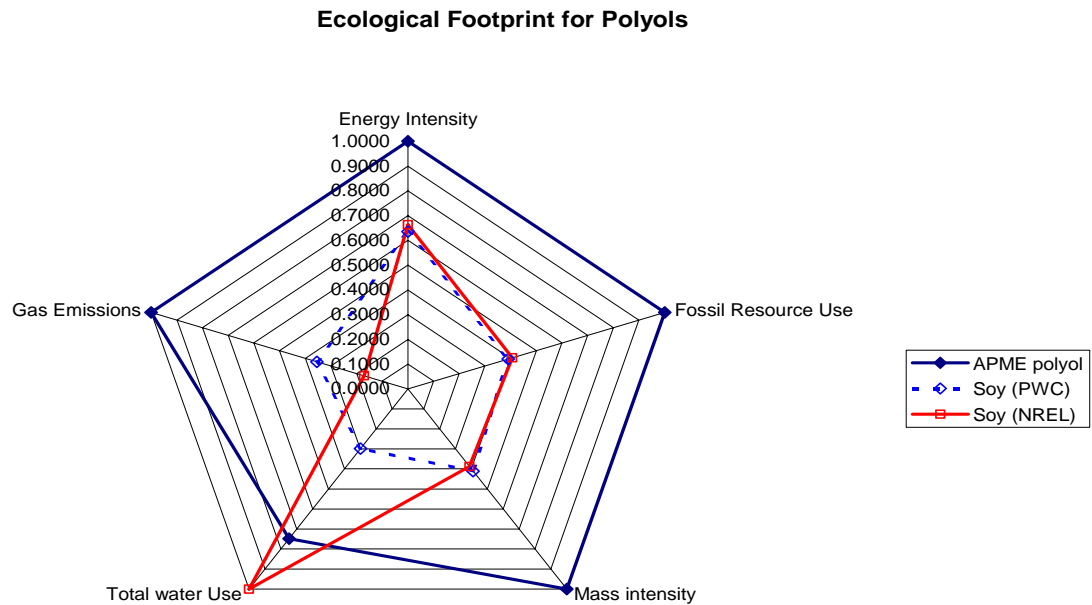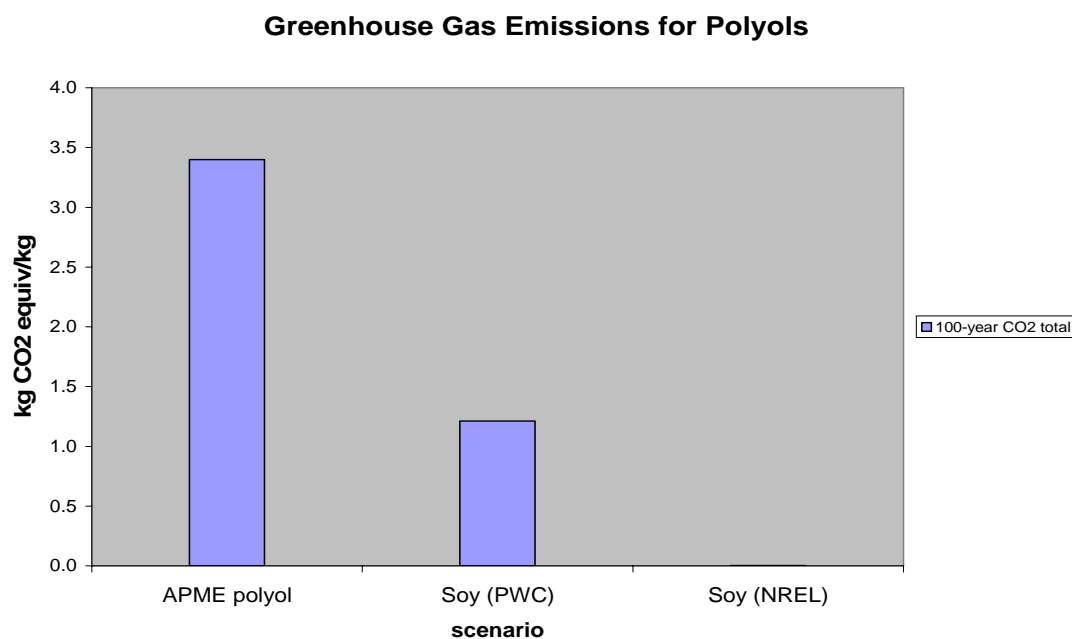
Ecological Footprint for Polyols

Figure 4-14. Impact of soy farming model on greenhouse gas emissions of polyols.

The reports by PWC and NIST do not clearly define the method use to calculate the "cradle to gate" LCI data or methods for delivered fuels and power, nor the manufacturing process used to convert the soy oil into polyol. They calculated a gross energy of 12 MJ/kg for a soy polyol, of which 11.4 was for non-renewable energy inputs and 0.5 MJ/kg for fossil feedstocks. This is much lower than the soy case calculated here, which had (excluding biomass energy) 34.74MJ/kg gross energy, of which 24.4 MJ/kg were for fuel use, production & transport and 8.9 MJ/kg was for fossil feedstocks. The first key difference is that the Dow soy polyol includes ethylene oxide (EO) as part of the polymer and the Omni Tech polyol does not. EO contributes 8.6 MJ/kg of feedstock energy and 5.4 MJ/kg of other fossil and non-fossil energy to the Dow polyol. The next key difference is that the Dow polyol includes hydroformylation as a chemical process step, and the synthesis gas used for this adds another 6.6 MJ/kg. The Omni tech polyol does not appear to include this step or materials. Subtracting these raw materials from the Dow polyol, and increasing the unit ratio of soy oil to polymer in the Dow material from 0.678 to 1.0, produces a same-basis estimate of soy polyol gross energy of 15.4 MJ/kg. This is 28% higher than the Omni Tech calculation, which is due to the lower estimate of soy farm inputs by Omni Tech. The polymer Omni Tech modeled (100% soy based) is intended for carpet backing. This LCA is a comparison of directly equivalent materials for flexible foams, so cannot be compared with the Omni Tech work.

Sensitivity analysis for major energy inputs

Another aspect of an LCA done to ISO standards is a sensitivity analysis for the impact of variations in the input parameters on the calculated results. The parameters selected for this analysis were determined by using the Boustead Model to calculate the contribution of the

different inputs to the gross energy of polyol made from irrigated Texas castor, with fertilizer as the useful co-products. The top five inputs, and their contribution as a percentage of the total gross energy, were:

- Natural gas (Texas (ERCOT) source and production data) used for irrigation pumps – 41.0%
- Ethylene oxide production (Dow data – top operation code 2122) – 18.8%
- Fertilizers (both ammonium phosphate (Dow (Vink) data – top operation 1301) and urea (Boustead Model material operation code 3669) – 13.9%
- Electricity (Texas (ERCOT) grid average) used in seed & oil processing – 14.1%
- Steam (Boustead Model materials processing operation code 1) used for energy in the polyol plant – 6.1%

These five inputs accounted for 94% of the gross energy of the castor polyol. Each of these inputs was individually varied by $\pm$ 20%, the results recalculated, and tabulated as a percentage change from the base case for five key metrics: gross energy, fossil resource use, reduction in greenhouse gas emissions, total inorganic raw materials used, total water used. The crop yield was also varied by 20% around a base of 1862 lb oil per acre (4138 lb seeds per acre; 4638 kg seeds/hectare).   These calculations are summarized in the table below. The impact on the metric and calculations other than those listed was typically <5%.

Table 4-1. Sensitivity of key metrics to $\pm$20% variations in largest inputs.

Input variable	Gross energy	Fossil use	GHG reduction	Inorganic raws	water use
Crop yield	3.8%	6.7%	9.9%	9.0%	17.1%
Irrigation energy (natural gas)	2.7%	4.8%	5.3%	0.0%	0.0%
EO mass in polymer	3.3%	6.1%	4.2%	10.9%	0.3%
Fertilizer use	1.0%	1.4%	1.5%	9.0%	0.0%
Electricity use in seed oil production	0.8%	1.3%	1.7%	0.0%	0.0%
Steam use in polymer plant	1.7%	2.8%	3.3%	0.0%	0.0%

The usually small impact of the $\pm$20% input changes shows the robustness of the results to errors in details of process knowledge. The sensitivity results can also be put into perspective by doing a similar calculation for the impact of the assumptions described earlier using a series of graphs. The table below shows the absolute value of the changes. The first line compares polyols from Texas irrigated castor with no useful products to that from the castor co-produced in Texas with fertilizer. The second line compares the best farm model (soy inputs) to the worst (irrigated Texas), both with no co-products.

Table 4-2. Impact of assumptions on key sensitivity metrics.

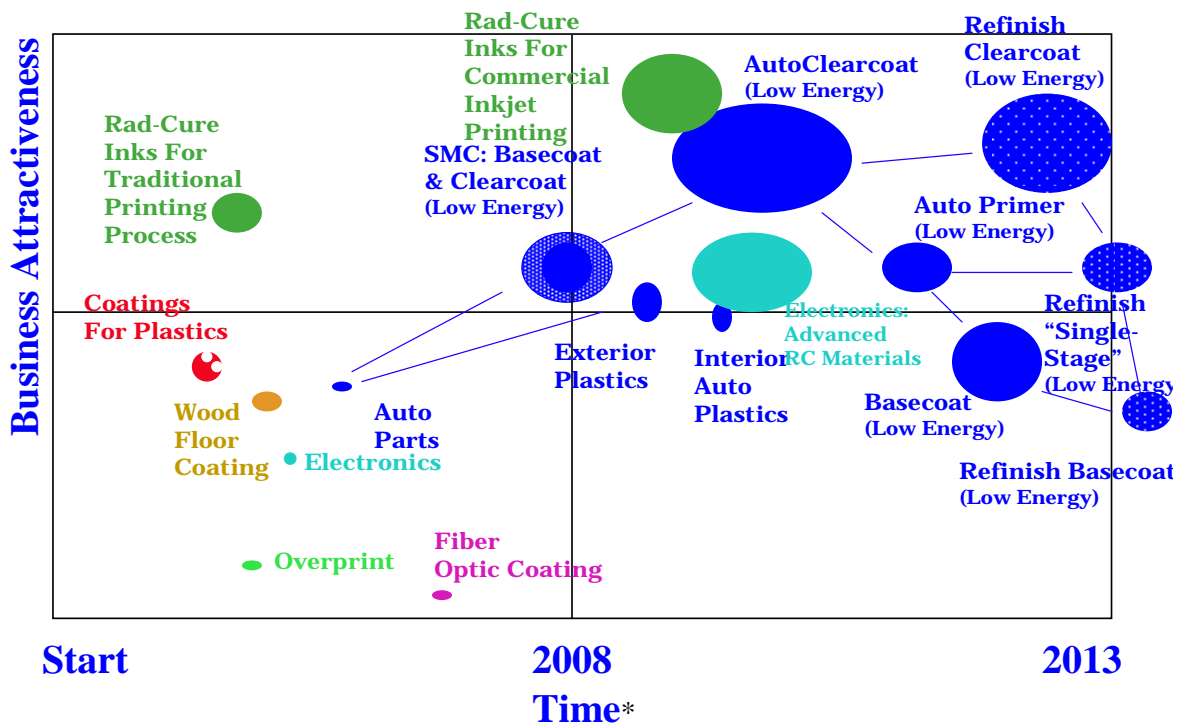| Input variable	Gross energy	Fossil use	GHG reduction	Inorganic raws	water use
No useful co-products	27.1%	51.3%	47.3%	121%	112%
"best" farm (soy inputs, castor output)	25.1%	35.9%	47.1%	89.8%	448%

These results confirm the much larger impact of the co-product and farm input assumptions on the LCA results than variations in input parameters.

4.3 Market Research for Alkenoates

NBIB Associates was contracted to explore the commercial viability of a new oleochemical-based monomer, methyl decenoate, and the co-product decene. In fields experiencing strong discontinuities, they looked for new business opportunities in terms of market openings and needs, what was required for competitive advantage and a business model for capturing value.

Starting with 100 million pound volume for combined product opportunities, the high growth area of radiation cured products was identified as higher value, performance driven market. Radiation cured (rad-cure) coatings were growing in volume due to 1) increased productivity, 2) no emissions or fire concerns and 3) with sensitive substrates there could be total cure temperature control which is ideal for heat sensitive surfaces.   The performance properties in seed oil oligomers (weatherability, flowability and good reactivity in rad-cure systems) were valued in rad-cure materials for coatings and inks. The co-product decene, was also valuable as a main ingredient in high performance lubricant market.

Leading applications include overprint varnishes, printing inks, wood and furniture coatings, plastics coatings, automotive coatings, electronic and optical fiber coatings. The following bubble chart outlines the potential in rad-cure opportunities.

Business Attractiveness

Rad-Cure Inks For Commercial Inkjet Printing

AutoClearcoat (Low Energy)

Refinish Clearcoat (Low Energy)

Rad-Cure Inks For Traditional Printing Process

SMC: Basecoat & Clearcoat (Low Energy)

Auto Primer (Low Energy)

Coatings For Plastics

Refinish "Single-Stage" (Low Energy)

Wood Floor Coating

Auto Parts

Electronics

Exterior Plastics

Interior Auto Plastics

Electronics: Advanced RC Materials

Basecoat (Low Energy)

Refinish Basecoat (Low Energy)

Overprint

Fiber Optic Coating

Start   2008   2013

Time*

* = Year of First Commercial Sale
Bubble size = relative potential

Publications

WRRC/USDA

Publications supported in part by grant:

McKeon, T. A., He, X., Turner, C., Chen, G. Q. and Lin, J. T. (2003) Altering seed oil composition for food and non-food uses. *Essential Fatty Acids and Eicosanoids: invited papers from 5th International Congress.* Editors, Yung-Sheng Huang, Shing-Jong Lin, and Po-Chao Huang. pp. 50–52. (Proceedings Paper)

McKeon, T.A. (2003) Genetically modified crops for industrial products and processes and their effects on human health. *Trends in Food Science and Technology 14*, 229-241.

Lin, J. T., Turner, C., Liao, L. P. and McKeon, T. A. (2003) Identification and quantification of the molecular species of acylglycerols in castor oil by HPLC using ELSD. *J. Liquid Chromatography & Related Technologies 26*, 773–780.

Turner, C., He, X., Nguyen, T., Lin, J. T., Wong, R. Y., Lundin, R. E., Harden, L. and McKeon, T. A. (2003) Lipase-catalyzed methanolysis of triricinolein in organic solvent to produce 1,2(2,3)-diricinolein. *Lipids 38,* 1197-1206.

He, X., Turner, C., Chen, G. Q., Lin, J. T., McKeon, T. A. (2004) Cloning and Characterization of a cDNA Encoding Diacylglycerol Acyltransferase from Castor Bean. *Lipids*, 39, 311-318.

Turner, C., King, J.W. and McKeon, T. (2004) Selected uses of enzymes with critical fluids in analytical chemistry. *J. AOAC International* 87, 797-810 (review).

Chen G.Q., He X., Liao L.P., McKeon T.A. (2004) 2S albumin gene expression in castor plant (Ricinus communis L.) *JAOCS,* 81, 867-872.

He, X., Chen, G. Q., Lin, J. T., McKeon, T. A. (2004) Regulation of diacylglycerol acyltransferase in developing seeds of castor. Lipids 39, 865-871.

Chen, Q. H., He, X. and McKeon, T. A. (2005) A Simple and Sensitive Assay for Distinguishing the Expression of Ricin and *Ricinus communis* Agglutinin Genes in Developing Castor Seed (*Ricinus communis* L.) J. Agric Food Chem. 53, 2358-2361.

Lin, J. T., Wani, S., He, X., Nguyen, T. and McKeon, T. A. (2005) Incorporation of laurate and hydroxylaurate into phosphatidylcholines and acylglycerols in castor microsomes. J. Am. Oil Chemists' Society 82, 495-499.

McKeon, T.A. (2005) Genetic modification of seed oils for industrial applications. In *Industrial Uses Of Vegetable Oils* (S.Z. Erhan, ed) pp. 1-13.

Turner, C., Wani, S., Wong, R., Lin, J.T. and McKeon, T. (2006) Lipase-catalyzed esterification of 2-monoricinolein for 1,2(2,3)-diricinolein synthesis. Lipids 41, 77-83.

He, X., Chen, G. Q., Lin, J. T., McKeon, T. A. (2006) Diacylglycerol acyltransferase activity and triacylglycerol synthesis in germinating castor seed cotyledons. Lipids 41, 281-285.

In addition to publications listed, there are 5 publications in preparation, one on the cloning of genes and characterization of castor acyl-CoA synthetases and 4 on preparation of novel acylglycerols for use as acyltransferase substrates.

Dow Chemical

Stability of the First-Generation Grubbs Metathesis Catalyst in a Continuous Flow Reactor
Zenon Lysenko, Bob R. Maughon, Tezi Mokhtar-Zadeh, Michael L. Tulchinsky

Submitted for publication to the Journal of Organometallic Chemistry.

Renewable Monomer Feedstocks via Olefin Metathesis: Fundamental Mechanistic Studies of Methyl Oleate Ethenolysis with the First-Generation Grubbs Catalyst
Burdett, K. A.; Harris, L. D.; Margl, P.; Maughon, B. R.; Mokhtar-Zadeh, T.; Saucier, P. C.; Wasserman, E. P.
Organometallics; **2004**; *23(9)*; 2027-2047.

Patent Applications

Dow Chemical

> 61071A   Filing   2/27/2002
>
> METATHESIS OF UNSATURATED FATTY ACID ESTERS OR UNSATURATED FATTY ACIDS WITH LOWER OLEFINS   Burdett Kenneth A., Maughon Robert R., Morrison Donald Lee, Newman Thomas H., Rand Cynthia L., Wasserman Eric P.
>
> > US   Publication   3/31/2005

> 61829A   Filing   4/17/2003
>
> INTEGRATED CHEMICAL PROCESSES FOR INDUSTRIAL UTILIZATION OF SEED OILS   Bicerano Jozef, Burdett Kenneth A., Christenson Christopher P., Cummins Clark H., Dettloff Marvin L., Lysenko Zenon, Maher John Michael, Maughon Robert R., Schrock Alan K., Thomas Pulikkottil J., Varijian R, White Jerry E.
>
> > US   Publication   7/14/2005

> 62198A   AR   Filing   9/26/2003
>
> STABILIZATION OF OLEFIN METATHESIS PRODUCT MIXTURES Au-Yeung Patrick H., Burdett Kenneth A., Maughon Robert R.
>
> > US   Publication   7/27/2006

> 62937A   Filing   9/14/2004
>
> AN IMPROVED PROCESS FOR THE SYNTHESIS OF UNSATURATED ALCOHOLS   Burdett Kenneth A., Lysenko Zenon, Maughon Robert R.

> 62937B
>
> PROCESS FOR THE SYNTHESIS OF UNSATURATED ALCOHOLS Burdett Kenneth A., Lysenko Zenon, Maughon Robert R.
>
> > US   Publication   4/14/2005

63545A   WO   Publication   5/4/2006

MEMBRANE SEPARATION OF A METATHESIS REACTION MIXTURE   Burdett Kenneth A., Collins Ray M., Maughon Robert R., Tulchinsky Michael Leo

64186   Filing   6/6/2005

METATHESIS PROCESS FOR PREPARING AN ALPLHA, OMEGA - FUNCTIONALIZED OLEFIN Burdett Kenneth A., Mokhtarzadeh Morteza

64186A   WO   Filing   6/2/2006   METATHESIS PROCESS FOR PREPARING AN ALPLHA, OMEGA -FUNCTIONALIZED OLEFIN Burdett Kenneth A., Mokhtarzadeh Morteza, Timmers Francis J.

The following application was just recently prepared and submitted internally:

"Seed-preferred Gene Promoters from the Castor Plant, *Ricinus communis*" Roessler, P., Rasochova, L., Lee, V.

USDA/WRRC

Diacylglycerol Acyltransferase and Its Use to Preferentially Incorporate Fatty Acids into Diacylglycerol.   T. McKeon, X. He, J.T. Lin. Reported June 3, 2004, DOE "S" Number: 10/861/616