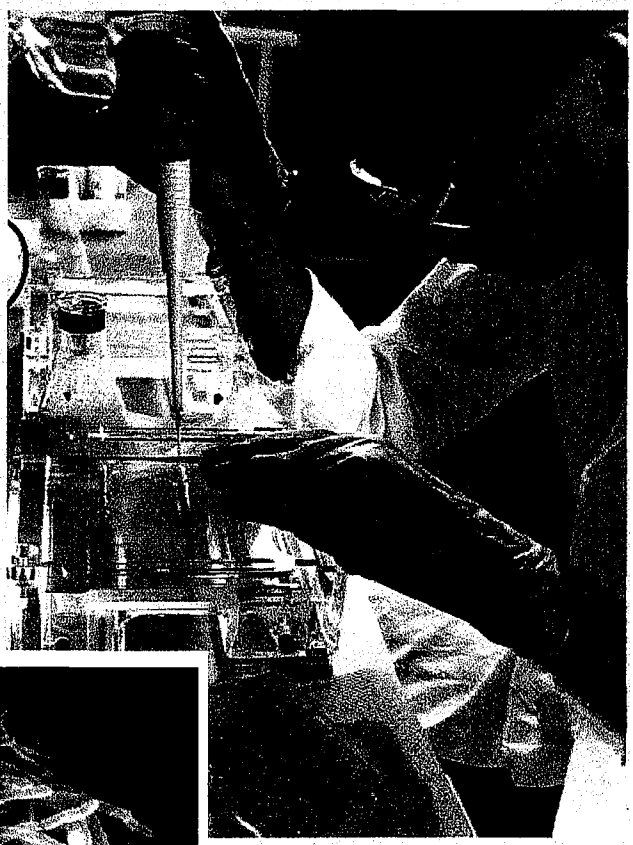
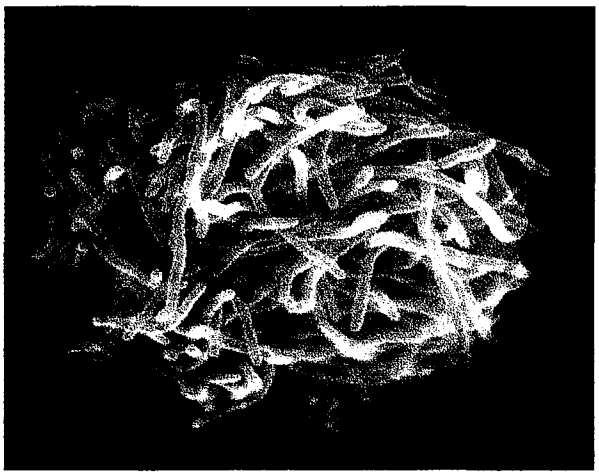
Biological and Chemical Technologies Research at OIT

Annual Summary Report FY 1997





Office of Industrial Technologies
Office of Energy Efficiency and Renewable Energy
U.S. Department of Energy

March, 1998

The Biological and Chemical Technologies Research Program is managed under the direction of Department of Energy Headquarters Program Management with technical support from the National Renewable Energy Laboratory and Oak Ridge National Laboratory.

The Biological and Chemical Technologies Research Program focuses on resolving the technical barriers that impede the use of biological systems in chemical production processes, and the barriers that restrict the design of new chemical and biological catalysts.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

This document was edited and prepared by the National Renewable Energy Laboratory (Ed. G. Petersen)

Abstract

The annual summary report presents the fiscal year (FY) 1997 research activities and accomplishments for the United States Department of Energy (DOE) Biological and Chemical Technologies Research (BCTR) Program. This BCTR program resides within the Office of Industrial Technologies (OIT) of the Office of Energy Efficiency and Renewable Energy (EE). The annual summary report for 1997 (ASR 97) contains the following: program description (including BCTR program mission statement, historical background, relevance, goals and objectives); program structure and organization; selected technical and programmatic highlights for 1997; detailed descriptions of individual projects; and a listing of program output, including a bibliography of published work, patents, and awards arising from work supported by the program.

Foreword

For further information about the Biological and Chemical Technologies Research Program, contact:

Mr. David J. Boron
U.S. Department of Energy
1000 Independence Ave, S. W.
Forrestal Building, MS EE-232, 5F-043
Washington, D.C. 20585

(202) 586-0080
FAX (202) 586-3180
david.boron@hq.doe.gov

Biological and Chemical Technologies Research Program

Annual Summary Report
FY 1997

Table of Contents

1.0 Executive Summary 1

2.0 FY 1997 Technical and Programmatic Achievements 5
- 2.1 Programmatic Achievements 5
  - 2.1.1 Collaborations and Industrial Support 5
  - 2.1.2 Solicitation in support of OIT's Chemicals Team and Technology Vision 2020. 5
- 2.2 Technical and Scientific Highlights 6
  - 2.2.1 California Institute of Technology (Frances Arnold) 6
  - 2.2.2 University of Pittsburgh, Department of Chemistry (David Beratan) . . . 6
  - 2.2.3 Oak Ridge National Laboratory (Brian H. Davison) 6
  - 2.2.4. Massachusetts Institute of Technology (Alexander M. Klibanov) 8
  - 2.2.5 California Institute of Technology (William A. Goddard, III) 8
  - 2.2.6 Lawrence Berkeley Laboratory (C. Judson King) 9
  - 2.2.7 National Renewable Energy Laboratory (Robert Meglen and Gene Petersen) 10
  - 2.2.8 Idaho National Engineering and Environmental Laboratory (R.D. Rogers) 10
  - 2.2.9 Howard University (Arthur Williams) 11

3.0 Industries of the Future 13
- 3.1 Chemicals Industry 13
- 3.2 Agriculture or Plant/Crop Biobased Renewables 15
- 3.2 Forest Products and Petroleum Refining 15

4.0 Biological and Chemical Technologies Research Program 17
- 4.1 Programmatic Relevancy 17
  - 4.1.1 OIT's Industries of the Future 17
  - 4.1.2 Relationship to EPAct and the Department of Energy 19
- 4.2 BCTR Program Overview 19
  - 4.2.1 Program Goal and Objectives 20
  - 4.2.2 Strategy 21
  - 4.2.3 Metrics 22
- 4.3 BCTR Program Structure 22
  - 4.3.1 Chemical Approaches 23
  - 4.3.2 Biological Approaches 23

5.0 Project Descriptions 25

6.0 Program Output 93
6.1 Publications in Professional Journals, Books and Symposium Chapters, and DOE Technical Publications 93
6.2 Patent Awards/Filings, and Licenses 96

1.0 Executive Summary

The Biological Chemical Technologies Research (BCTR) program concluded another successful year of research in FY 1997. In FY 1997, the BCTR program activities received strong industrial support, contributed to the economy through job formation, and achieved scientific and professional success. Corporate support is reflected in EPAct-related cost-sharing of about $1.6 million and leveraged contributions amounting to over $2.5 million (Table 1.1). Overall, industry provided about 0.8 dollar for every federal R&D dollar invested through the program. Research activities involved 27 industrial partners as well as leveraging work with other federal agencies, industry, and other institutions (Tables 1.2 and 1.3). Additionally, BCTR supported leveraged the equivalent of about seven (7) full-time equivalents from industry (Table 1.1). One measure of the value of federal funding is the number of students being trained in high-skill and eventually high-paying jobs. The program involved such training for over 13 students and during FY97, five students and postdoctoral fellows found employment in industry or academics (Table 1.4). Another measure of program quality is the publication record of the participants and the development of technology transfer elements such as patents and licenses. Program participants published 43 peer-reviewed articles and documents, filed four (4) new patents, and have established one licensing agreement (Table 1.5).

Section 2.0 provides a brief description of the significant findings and results of work performed by program participants. Technical and scientific highlights include development of new ways to design biocataylsts, the development of new biocatalyst systetms, and CRADA work on novel bioprocessing techniques.

Section 3.0 describes the integration of the BCTR program into the Industries of the Future effort by OIT. The application of the program efforts has typically emphasized work in the Chemical and Allied Products and Petroleum Refining industries. However, the work is clearly applicable to the existing Forest Products industry and the emerging vision area of Agriculture. Tables to illustrate that are provided in section 4.1.1.

Section 4.0 outlines the mission of the DOE BCTR Program, which is to provide evidence of the technical and economic feasibility of advanced chemical and biological concepts that can improve the energy utilization, operation efficiency, and environmental soundness of U.S. industry process operations. Programmatic relevancy to these industry visions and an overview of the program are provided. The program strives to combine science with technology to develop novel and revolutionary process concepts; to improve conventional process approaches; and, if the economics and market needs warrant, to facilitate the introduction of these technological advances in the chemicals, petroleum refining, other renewable, and emerging biotechnology industries.

Section 4.0 also outlines the management of the program which resides with the BCTR program manager in the Office of Industrial Technologies (OIT). Program industrial focus is provided primarily through input from industrial forums and partners and OIT roadmapping in the chemical and agriculture industry vision areas. In addition, technical field assistance is provided in all areas of R&D by DOE national laboratories.

Section 5.0 details the progress and future work of each of the projects. Section 6.0 provides lists of published papers, awards, patents, and filings.

Table 1.1 OIT's BCTR Program Industrial Cost-share Information

Project with Abbreviated Title and Performing Organization[a]	Program Element[b]	FY 97 Funds	Cost Share by Category[c] in ($000)		IND FTE
			C	L	
Enz. Catalysts for Biotech Based Chem Industry (CIT)	BDV	137	193.3	95.4	0.75
Theory of Biocatalysis (U. Pitt.)	BDS	0	75	150	
Adv. Bioreactor Concepts for Non-Traditional Bioprocessing (ORNL and subcontracts)	BPS	828	356	1138	3.5
Biocatalysis by Design (Simulation center)[d] (CIT)	BDS	348	348	750	1.5
Separations by Reversible Chemical Association for Biobased Chemicals (LBL)	BPS	180	150	250	0.5
Strategies for Biocatalyst Design in Extreme Conditions for Industrial Applications (NREL and subcontracts)	BDS	642	260	200	0.25
Biological separation of phosphate from ores (INEEL)	BPS	0	266		1
TOTALS BY CATEGORY		**2135**	**1648.3**	**2583**	**7.5**

a. May include industrial cost-sharing by subcontractors
b. Key for Program: **BDS** - Biocatalyst Design; **BDV** - Biocatalyst Development; **BPS** - Bioprocess Systems
c. Cost sharing is defined in two categories. C - These are contributions, monetary or equipment, in-kind, or other forms of assistance directly contributing to completion of the statement of work of the project and conform to EPact 92. L or leveraged contributions indicate work supported from other federal agencies or industrial sources which do not directly support the completion of the specific BCTR/OIT projects but which impact the performance of the work or are complementary to the OIT/BCTR effort.
d. The Materials and Molecular Simulations Center at California Institute of Technology (CIT) has a formal agreement with a number of companies but since the center is also funded by DoD, NASA, and NSF, the industry cost-share for DOE/BCTR is considered to be equal to the DOE contribution.

Table 1.2 Listing of Industrial Partners and BCTR supported institutions

	Industrial Partners	BCTR supported Institution (either directly or via subcontract with a National Laboratory)
1	Altus Biologics	ORNL, U. Iowa, U.C. California Berkeley, Massachussets Institute of Technology
2	Ashai Chemical	California Institute of Technology
3	Avery-Dennison	California Institute of Technology
4	Biotechnology Research and Development Corporation	California Institute of Technology
5	BP Chemical	California Institute of Technology
6	Cargill	LBNL
7	Cleveland Cliffs Iron Company	INEEL
8	Chevron Petroleum Technology	California Institute of Technology
9	Chevron Research	California Institute of Technology
10	Cray Research	University of Pittsburgh
11	Diversa	North Carolina State University and University of Georgia
12	Dow Chemical	ORNL and California Institute of Technology
13	DuPont Chemical	ORNL and INEEL
14	Eastman Chemicals	ORNL (19th Symposium)
15	Eli Lilly	California Institute of Technology
16	Enzymol International	U. Iowa and U.C. California Berkeley
17	EnzyMed, Inc.	U. Iowa, U.C. California Berkeley
18	Exxon Research	California Institute of Technology
19	Genencor	U. Iowa and U.C. California Berkeley
20	J.R. Simplot Company	INEEL
21	Kraft Foods	U. Iowa and U.C. California Berkeley
22	Owens-Corning	California Institute of Technology
23	Oxyrase Inc.	ORNL
24	Procter and Gamble	California Institute of Technology
25	Saudi Aramco	California Institute of Technology

	Industrial Partners	BCTR supported Institution (either directly or via subcontract with a National Laboratory)
26	Seiko-Epson	California Institute of Technology
27	Thermogen	California Institute of Technology

Table 1.3 Listing of Leverage Contributors

Companies, Institutions or Funding Agencies
Advanced Separations Technologies, Inc. Altus Biologics, Inc. Army Research Office DOE/ER EMSP Program DOE/EE Office of Transportation Technologies Dow Chemical DuPont Chemical Michigan Biotechnology Institute National Aeronautics and Space Administration National Science Foundation Office of Naval Research

Table 1.4 Educational and Jobs Impact

Students Trained in High-skilled Jobs	Job Placement
13.4	**5** - 2 to industry and 3 to academia

Table 1.5 Publications, Patents, and Licenses

Total Publications by Program Participants	Patents - Filed or Awarded Licenses
43	**Four** (4) patents filed **One** (1) license negotiated

2.0 FY 1997 Technical and Programmatic Achievements

2.1 Programmatic Achievements

2.1.1 Collaborations and Industrial Support

In FY 1997, the BCTR program activities received strong industrial support and contributed to the economy through job formation. Corporate support is reflected in EPAct-related cost-sharing of about $1.6 million and leveraged contributions amounting to over $4.2 million (Table 1.1). Overall, industry provided about 0.8 dollar for every federal R&D dollar invested through the program. Research activities involved 27 industrial partners as well as leveraging work with other federal agencies, industry, and other institutions (Tables 1.2 and 1.3). Additionally, BCTR supported leveraged the equivalent of about seven (7) full-time equivalents from industry (Table 1.1). One measure of the value of federal funding is the number of students being trained in high-skill and eventually high-paying jobs. The program involved such training for over 13 students and during FY97, five students and postdoctoral fellows found employment in industry or academics (Table 1.4). Another measure of program quality is the publication record of the participants and the development of technology transfer elements such as patents and licenses. Program participants published 43 peer-reviewed articles and documents, filed four (4) new patents, and have established one licensing agreement (Table 1.5).

2.1.2 Solicitation in support of OIT's Chemicals Team and Technology Vision 2020.

This year a focused solicitation was made to universities, industry, and the national laboratories to solicit new ideas and approaches to the development of biocatalysts suitable for use in the extreme environments of the chemical processing industry. The solicitation yielded four high-quality projects that were started in FY 1997 using the subcontracting functions at two DOE laboratories. The following lists the projects that were selected by a peer review process largely composed of industrial reviewers.

Investigator(s) and Institution Laboratory managing the subcontract is listed following the name of the awardee	Title of Effort
Michael Adams (U. Georgia) ORNL Robert Kelly (N. Carolina State U) NREL	*Synthetic Biocatalysis by Hyperthermophilic Enzymes*
Doug Clark (U. Cal. Berkeley) ORNL Jonathan Dordick (U. Iowa) ORNL	*Activation and Stabilization of Enzymes for use in Organic Solvents*
Alan Russell (U. Pittsburgh) NREL	*Efficient Biocatalytic Synthesis of Polyesters from Bulk Monomers*
Frances Arnold (Caltech) NREL	In vitro *Evolution of Oxidation Catalysts: Towards Efficient Biochemical Production of Commodity Chemicals*

All work began about the end of FY 1997 or the beginning of FY 1998 and are limited to about 15 months work, although each project was defined as a 2 year effort.

2.2 Technical and Scientific Highlights

2.2.1 **California Institute of Technology** (Frances Arnold)

A major technical advance was made to assist in the development of new biocatalyst for use in the chemical processing industry. During this past year we filed a utility patent application and prepared a publication on a novel method for *in vitro* recombination of mutated genes. This recombination method provides for the rapid accumulation of beneficial mutations and removal of neutral and deleterious mutations from a population of parent genes, enhancing the rate of evolution (paper submitted to *Nucleic Acids Research*).

2.2.2 **University of Pittsburgh, Department of Chemistry** (David Beratan)

A major chemical reaction in the chemical processing industry involves oxidation and reduction of chemical species. The catalysis roadmap[1] for the Chemical Technology Vision 2020 points out that selective oxidation is the number one area of interest by industry. The major, critical element of these reactions is the transfer of electrons. If enzymes are to be used in the chemical processing industry, a mechanistic and experimental understanding of electron transfer must be achieved. In addition, the site architecture of these potential catalyst materials must be better known. The following highlight some of the strong work that has occurred there over the past few years in both electron transfer and site architecture.

- We expanded our studies of two-electron transfer reactions to known reactions in solution, and used the formalism to treat three-state electron transfer processes. This will assist us in computing the rates of two-electron transfers as a function of free energy in stepwise and concerted regimes.

- We completed an extensive evaluation of DNA electron transfer experiments using our new theories. This will allow estimates of electron transfer rates in experimental DNA systems such as those constructed by the Meade, Barton, and Harriman groups.

- We completed our development of electron transfer contact maps which allowed us to develop a multiple pathway representation of tunneling interactions in proteins using the concept of electron transfer contact maps.

- We completed our analysis of inverse design strategies to help design new materials.

2.2.3 **Oak Ridge National Laboratory** (Brian H. Davison)

Bioprocesses and Biotechnology are among the three major technology areas in the Technology Vision 2020 recommendation for New Chemical Sciences and Engineering Technology. The OIT chemical solicitation specifically called out this area as a focus area for new proposals. Initial inputs on roadmapping suggest three major targets of this focus: New biocatalysts (enzymes or microbes), developing bioprocess engineering principles, and using

---

[1]"Catalyst Technology Roadmap Report" Built upon Technology Vision 2020: The U.S. Chemical Industry. Workshop held March 20-21, 1997 at the American Chemical Society Headquarters, Washington, D.C. and sponsored by the Council for Chemical Research, U.S. Department of Energy, and the American Chemical Society.

biologically based feedstocks[2]. Work at ORNL has centered on further development of unique biocatalyst systems, viz. those that operate in extreme environments such as gas phase biocatalysis and nonaqueous bioprocessing. In addition, efforts involving advanced bioreactor design have been undertaken via a CRADA with Dow Chemical Company in an examination of potential applications in converting waste streams into value-added products.

- ORNL also contributed to the development of a document entitled "Biotechnology for National Industrial Needs, part II: Development of Biotechnology Capabilities in Extreme Environments for the Chemical and Related Products Industry." We also assisted David Boron and Gene Petersen in the development of a companion roadmap for this area. The use of enzymes in organic media for chemical production offers several advantages over traditional heterogeneous catalysis, including greater substrate solubility, elimination of undesirable side reactions with water, thermodynamic equilibria more favorable than those in water, and easier and less expensive product recovery.

- A cohesive nomenclature and methodology for classifying nontraditional biocatalytic processes was developed and published as a topical paper in *Biotechnology Progress*, **13**: 512-518 (1997). Such categorization provides a means by which to classify new work within the nontraditional arena, as well as to broaden or refine current research. Furthermore, a technology roadmap was developed which outlines nontraditional areas and their associated development issues which still require examination, in terms of both bridging and fundamental research, before these techniques will be adopted by the private sector.

- Gas-phase transesterification reactions were demonstrated in both batch and continuous reactor configurations for production of organic compounds such as butyl acetate and isoamyl acetate. Short-term trials have demonstrated the feasibility of this unique nonaqueous approach for chemical synthesis which is particularly amenable for situations in which the substrates are sparingly soluble in water. Effects of pH, humidity, and synthesis type on activity were measured.

- Immobilization and reactor design were investigated for the use of high-efficiency immobilized enzymes for dehalogenation in a CRADA with Dow Chemical. Dehalogenation addresses a direct industrial need to transform a current by-product waste of organic synthesis and allow it to be recycled in the process. The basic organism has been cultured for enzyme production at high yield and at 70-L scale. Dow Chemical has developed a greatly improved genetically modified organism which is being used at ORNL. ORNL has successfully immobilized the active enzyme into a novel porous glass matrix. The enzyme retained 10-20% of the original activity. We are continuing to try and increase the enzyme loading and activity.

- The 19th Symposium on Biotechnology for Fuels and Chemicals was partially supported by OIT. Industry leveraged DOE's contribution by a greater than 2:1 margin.

- ORNL also monitored an ongoing subcontract with the Massachusetts Institute of Technology in Rational Enhancement of Enzyme Performance in Organic Solvents.

---

[2] Workshop held at the 19th Symposium on Biotechnology for Fuels and Chemicals, Colorado Springs Colorado, May, 1997 sponsored by the Council for Chemical Research and supported by DOE.

2.2.4. **Massachusetts Institute of Technology** (Alexander M. Klibanov)

The basis for this work is described in the heading for the highlight for ORNL, 2.2.3. This work is supported by DOE through a subcontract with ORNL.

- We have discovered that the catalytic activity of cross-linked enzyme crystals suspended in organic solvents can be dramatically increased by adding appropriate "organic buffers" (i.e., mixtures of an acid and its conjugate base). This phenomenon has been rationalized mechanistically.

- We have directly demonstrated by means of Fourier Transform InfraRed (FTIR) spectroscopy that while many neat organic solvents do not denature enzymes, mixtures of the same solvents with water do. This explains why enzymes are catalytically active in neat organic solvents (and in water) but not in aqueous-organic mixtures.

- We have developed a simple thermodynamic methodology that can explain the enzymatic substrate specificity in organic solvents even if the enzyme structure is not known.

- We have explained, using FTIR structural analysis, why subtilisin is catalytically active when dissolved in some organic solvents but not in others.

2.2.5 **California Institute of Technology** (William A. Goddard, III)

The focus of this work is to develop an atomistic simulations tool useful in optimizing biotechnology processes, in particular biocatalysis. These are specific components of the Technology Vision 2020 Technical Recommendations in New Chemical Sciences and Engineering Technology (*under Chemical Synthesis, Biotechnology, and Computation*). Consequently the focus was on mechanisms in biotechnology processes.

The following are topical areas that were evaluated and investigated throughout the past year. References and outcomes are noted where specific output resulted in publishable information or information useful to industry. Some are related to biotechnology even though they are not specifically employing biological systems.

<u>Process Simulations</u> – Mario Blanco. We have continued our industrial collaborations in the areas of process simulation. These include:

Scale Control : *Chevron Petroleum Technology Company*. Work was designed to develop a new scale dissolver formulation.
Polymer Compatibilizers and Silane Coupling Agents: *Owens Coming*. Work effort involved developing new software for compatibility predictions that save laboratory efforts.
Wear Inhibitors: *Chevron Research Technology Center*. Work involves developing a self-assembling monolayer model for metal wear inhibition by dithiophosphates.
Gasoline Reformer Catalysts: *Exxon Research*. Work included examining new catalysts for gasoline reformulation.
Water-Resistant Adhesives: *Avery Dennison*. Identified a mechanism for outstanding water resistance in new acrylic modified adhesives

<u>Mechanism of Uncoating of the Rhinovirus</u> – Nagarajan Vaidehi. We have published this work in the *Proceedings of the National Academy of Sciences* USA, **94**, pp 2466-2471 (1997).

Domain Motions in Phospho Glycerate Kinase (PGK) – Nagarajan Vaidehi The movie of the dynamics can be found at http://www.wag.caltech.edu

A Second Generation Force Field for Nucleic Acid Simulations – Ken A. Brameld & Siddharth Dasgupta. We have published the work. Brameld, K.A., Dasgupta, S., Goddard, W.A. III. "Distance Dependent Hydrogen Bond Potentials for Nucleic Acid Base Pairs from *ab initio* Quantum Mechanical Calculations (LMP2/cc-pVTZ)." *J. Phys. Chem B.*, **101**, 4851-4859.

Investigation of the Mechanism of Family 18 and Family 19 Chitinases – Ken. A. Brameld, William A. Goddard, III. Three papers were submitted for publication: a) Brameld, K. A.; Shrader, W. D.; Imperiali, B.; Goddard, W. A., III. "Substrate Assistance in the Mechanism of Family 18 Chitinases: Theoretical Studies of Potential Intermediates and Inhibitors." *J. Mol. Biol.*; b) Brameld, K. A.; Goddard, W. A., III. "The D-Boat Substrate Distortion Mechanism for Family 18 Chitinases." *J. Am. Chem. Soc.*; and c) Brameld, K. A.; Goddard, W. A., III. "The Role of Enzyme Distortion in the Single-Displacement Mechanism of Family 19 Chitinases." *Proc. Natl. Acad. Sci. USA.*

Resolving the Levinthal Paradox – Practical Protein Folding – Derek A. Debe. A paper was submitted for publication in *Nature*, D. A. Debe, M. J. Carlson, and W. A. Goddard III, Resolving the Levinthal Paradox: Practical Protein Folding.

The Early Stages of the Protein Folding Process – Ruth Ann Bertsch, Nagarajan Vaidehi, Yao-Chun Peng, Sunney I. Chan, & William A. Goddard, III. We have predicted that the folding time scales in our simulations would be in the nanoseconds scale for 100% folding yield. This paper is being communicated to *Protein Science*.

Sidechain Rotamer Predictions in Folded Proteins – Matt J. Carlson, Derek Debe, William A. Goddard III. Our rotamer prediction method compares favorably with other methods in the literature. The strength of the method lies in its good performance with a smaller rotamer set, about half the number of rotamers of other current methods. This is achieved with our custom force field used in the energy evaluation. Thus, our method has the capability to operate on a large number of candidate protein folds to better fit within our hierarchical framework.

2.2.6 Lawrence Berkeley Laboratory (C. Judson King)

Recovery of Carboxylic Acids

Organic acids currently represent a large market in the food and beverage industry. They also represent a potentially large market in the commodity chemical industry. Recovery of these compounds tends to be costly. Understanding the underlying separation principles will allow industry to employ new routes to these compounds, including fermentative processes.

Regeneration of Adsorbents - One approach to lowering recovery costs is to focus or concentrate the compounds of interest so that adsorbent recovery processes will be more efficient. We have found that leaching an organic/adduct solution in a fixed bed concentrates a carboxylate-adduct ahead of the organic adduct solution. This results in a higher concentration of product and hence lowers cost of recovery.

Sustained Adsorption Capacity - We have confirmed a prediction that the presence of carbon dioxide during adsorption of a carboxylic acid onto a basic sorbent such as Dowex MWA- I serves to buffer the pH and thereby enable a high degree of recovery of the carboxylic acid.

Removal of carbon dioxide then enables regeneration of the sorbent. Similar effects have been found by others in solvent extraction processes.

Fractionation among Solutes - Stillage from fermentation ethanol plants contains both lactic and succinic acids and glycerol. We have found that a sorbent, Dowex MWA-I, fractionates effectively between the carboxylic acids and glycerol. One potential means of separating between lactic and succinic acids is to utilize the much lower solubility of succinic acid for selective precipitation. We found that the solubility of succinic acid in ketone solvents is only slightly influenced by the simultaneous presence of lactic acid.

Adsorption of Poly -OH Compounds onto Activated Carbons - We discovered an effect that affords what should be an economical regeneration process. We have found that regeneration of glycols, glycerol, sugars and sugar alcohols from activated carbons is incomplete in the presence of water, but is complete following prior removal of co-adsorbed water. Furthermore, there is a substantial focusing effect for fixed-bed leaching of glycols with acetone, and of glycerol with methanol.

2.2.7 **National Renewable Energy Laboratory** (Robert Meglen and Gene Petersen)

The goals for FY 1997 were to further refine and develop the structural database for quantitative structure activity modeling, to select additional model systems, and to develop and test several predictive models for enzyme activity.

- QSAR modeling has been extended to green fluorescence protein (GFP). X-ray crystal structures for modeling 13 published mutants and combined with published fluorescence spectra.

- Molecular modeling (Insight/Discover/MSI) has also been used to generate 413 synthetic crystal structures on *de novo* GFP mutants.

- Initial exploratory studies have been initiated to combine empirical data from infrared spectroscopy to the prediction of secondary structure such as - helix and β-pleated sheets.

2.2.8 **Idaho National Engineering and Environmental Laboratory** (R.D. Rogers)

A cost analysis on the production of a micro-encapsulated, biologically mediated, slow release phosphate fertilizer (biobead) was prepared. The study validated previous assumptions that the biobead concept is economically viable with attractive environmental benefits. This is documented and details the development and preliminary testing of the biobead. Preliminary agreements are being reviewed with phosphate fertilizer producers for the development and licensing of the biobead technology.

- A new cost analysis on the production of a pelleted microencapsulated, biologically mediated, slow release phosphate fertilizer (dry pellet) was prepared. The study validated previous assumptions obtained from production of a water-based biobead that the concept is economically viable with attractive environmental benefits.

- Component parts of the pellet were selected and tested for comparability.

- Growth studies conducted to evaluated manufactured biopellets showed that plant's response to the pellets was comparable to that of phosphate fertilizer.
- Preliminary agreements are being reviewed with phosphate fertilizer producers for the development and licensing of the biotechnology.

2.2.9 **Howard University** (Arthur Williams)

This report covers 2 years of work. The initial set of highlights covers the overall results.

- Generated several PPO recombinant clones of *C. versicolor* by means of cDNA approach.
- Assessed the expression of the PPO gene in bacterial cells as well as a unique yeast variant.
- Quantified PPO production via antibody probing.
- Enhanced PPO production via cultural manipulations.
- Detected other enzymatic activities associated with the PPO complex.
- Initiated the utilization of PPO as a biocatalyst in the processing of lignocellulose.

Additional work that was conducted earlier included the following:

- Marketing of the mass production of PPO via genetic engineering and cultural manipulations.
- Resolving other enzymatic activities associated with the PPO complex.
- Utilizing multiple activities of PPO complex in the bioprocessing of lignocellulose.
- Demonstrating other industrial applications of PPO.
- Generating a patent for PPO proteins.

3.0 Industries of the Future

The Office of Industrial Technologies' (OIT) Industries of the Future (IOF) represents a new approach to aligning federal investments in technology research, development, and deployment with the needs and expectations of technology users in the private sector. It aims to shrink the time required to move innovative industrial technologies from concept to marketplace. A central theme of the Industries of the Future process is helping manufacturers reduce costs and improve efficiency.

Based largely on energy consumption and waste generation, OIT has identified seven industries for its visioning process: chemicals, agriculture, glass, metal casting, steel, aluminum, and forest products. Formerly petroleum was another IOF. The Biological and Chemical Technologies Research (BCTR) Program pursues research that directly supports two of the seven industry visions, chemicals and agriculture. Technology development could also be applied to forest products. While petroleum is no longer a vision area, BCTR work does involve several areas of petrochemical production. The following is an overview of the vision and needs of the industries that the program directly supports.

3.1 Chemicals Industry

The chemical industry's vision of the future is one that provides the building blocks for a modern society and leads the world in the manufacturing and sale of chemicals and related products to meet the need of their customers. The relevance of the chemical industry in the U.S. economy and world marketplace and of issues that confront the industry's ability to maintain its competitiveness is synopsized below:

According to the *Technology Vision 2020* 1996 Report,[3] the chemical industry must accomplish five broad goals over the next 25 years. To meet this challenge, it must:

- Improve operations, with a focus on better management of the supply chain
- Improve efficiency in the use of raw materials, the reuse of recycled materials, and the generation and use of energy
- Continue to play a leadership role in balancing environmental and economic considerations
- Aggressively commit to longer term investment in R&D
- Balance investments in technology by leveraging the capabilities of government, academe, and the chemical industry as a whole through targeted collaborative efforts in R&D.

To meet its goals, the U.S. chemical industry should accomplish the following:

- Generate and use new knowledge

---

[3] *Technology Vision 2020, The U. S. Chemical Industry*, ACS, AICHE, CMA, CCR, SOCMA, (1996). Copies of this report can be obtained from American Chemical Society, Department of Government Relations and Science Policy, 1155 16th Street, NW, Washington DC 20036, (202) 452-8917, http://www.acs.org.

- Capitalize on information technology
- Encourage the elimination of barriers
- Work to improve the legislative and regulatory climate
- Improve logistics efficiencies
- Increase agility in manufacturing
- Harmonize standards
- Create momentum for partnering
- Encourage educational improvements

The chemical industry has recommended four technology areas with which to meet its vision:

Manufacturing and Operations This involves the operation of facilities and equipment vital to this industry.

Supply Chain Management This industry provides the building block materials for all U.S. manufacturing. Managing the supply chain involves all of the planning and operational decision making involved in the movement, handling, and storage of all materials purchased, processed, or distributed by chemical manufacturers. As such, it involves not only the manufacturers, but also all of their trading partners.

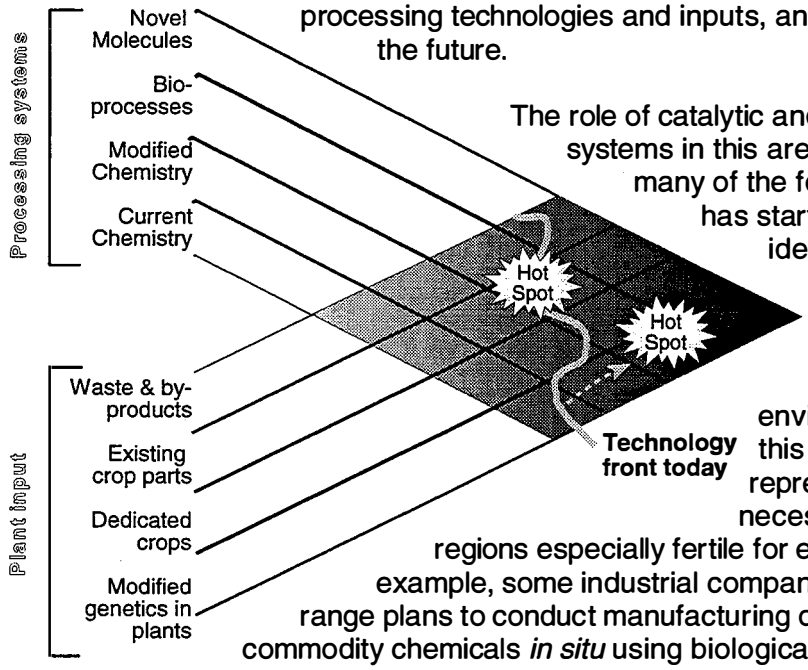
Information Systems The chemical industry has a future vision of a highly efficient and automated manufacturing operation and supply chain. To achieve this vision, data and information must flow in a seamless fashion along the whole supply chain from raw materials suppliers through all the steps of manufacturing operations to the customer.

New Chemical Sciences and Engineering Technology This is the major area where industry has determined that DOE's OIT support can have an impact. Chemical science and engineering technology drives new product and process development. Recently, biotechnology has begun to provide alternatives to traditional chemistry, and the application of physical sciences in chemical processes and materials production is advancing the chemical and materials sciences. In a similar manner, measurement and computational technologies have emerged as critical enabling technologies, significantly influencing the conduct and capability of chemical sciences and engineering technology. The intersection of chemistry, engineering, physics, biology, computational technologies, and measurement is creating a new technology paradigm, resulting in increasing complexity and skills required to develop new products and processes. Technical recommendations include the following areas – chemical synthesis, bioprocesses and biotechnology, and materials technology. Enabling technologies include process science and engineering, chemical measurement, and computational technologies. For example in the area of catalysis alone, there is significant growth predicted. Studies have shown a 5.2% annual increase in demand between 1995 and 2000 to a total value of nearly $3 billion and a 1.5% increase in volume over that span to 5.94 billion lb (C&EN, Nov 4, 1996, p. 28). Catalyst manufacturing is about a $10 billion market in refining, polymerization, chemicals and environment. One hundred firms, half of which are U.S. based, compete in catalyst manufacture. Our American manufacturing industry consumes about 45-50% of worldwide

catalyst production (ATP Focused Program Competition 98-05-Catalysis and Biocatalysis Technologies, December, 1997)

3.2 Agriculture or Plant/Crop Biobased Renewables

The production of agrichemicals accounts for a significant portion of the total volume of chemicals produced in the United States. Biobased chemicals production is being vigorously pursued by DOE and the U.S. Department of Agriculture (USDA). The agriculture sector has a market value of about $960 billion (or 16% of the U.S. GDP), provides 23 million jobs, and consumes about 20% of all energy sources. A vision document has been produced and published[4]. The vision calls for directional targets to chart successful progress that include 1) achieving at least a 10% market share of chemical building blocks derived from plant-based renewables by 2020; 2) developing production chain platforms that employ renewable-based feedstocks, and 3) developing partnerships that dissolve the traditional barriers between food, feed, fiber, and manufacturing from biobased products. A matrix was developed that allows for evaluation of opportunities and priority areas for R&D. The matrix employs the intersection of processing systems with plant/crop inputs to establish technology fronts that estimates where we are today, demonstrates the simultaneous developments between processing technologies and inputs, and shows progress towards the future.



The role of catalytic and biocatalytic processing systems in this arena is likely to be evident in many of the focus areas. Roadmapping has started and obvious needs identified include the development of new biocatalyst systems and catalytic systems that work in aqueous environments. "Hot spots" (in this figure these are merely representative and not necessarily actual) would indicate regions especially fertile for enabling-type research. For example, some industrial companies have announced long-range plans to conduct manufacturing of fine, speciality, and commodity chemicals *in situ* using biological plants. Obviously the enabling technologies provided by OIT will have application in this industrial sector when applied to moving the technology front towards a potential hot spot opportunity.

3.2 Forest Products and Petroleum Refining

The forest products industry has identified six major focus areas in their Agenda 2020 which will establish the future path of the industry. They are sustainable forest management, environmental performance, improved capital effectiveness, energy performance, recycling, and

[4] Plant/Crop-Based Renewable Resources 2020: A vision to enhance U.S. economic security through renewable plant/crop based resource use. DOE/GO-10098-385, January, 1998. Available by calling (800)-363-EREC.

sensors and controls. In previous years, BCTR projects included efforts to convert renewable biomass feedstocks into chemicals and to model catalyst/feedstock interactions and complex fluid flows. These are the kinds of technologies employed by the pulp and paper industry in their Kaymr digestors and recovery boilers.

As a result of the former OIT Refinery of the Future initiative, the refining industry provided a listing of the core technology areas in which it encouraged DOE to become involved. These areas have definite impacts on process efficiency, capital investment strategies, and improving pollution prevention. Improved technologies within the following seven areas will lead to reduced energy usage, better strategies for capital investments that include energy efficient operation, and better strategies to prevent pollution. While the OIT efforts through the BCTR program have had only modest incursions into this industry, the potential research areas are outlined below to provide some understanding of this industrial sector's technology needs.

The seven areas include catalysis, hydrogen generation, advanced process/controls, environmental technologies, plant reliability, computational fluid dynamics for reacting multiphase flows, and separations.

4.0 Biological and Chemical Technologies Research Program

The emphasis of the BCTR program has always been toward the Chemical and Allied Products (SIC 28) and Petroleum Refining (SIC 29) industries and has included the use of renewables as feedstock materials in the chemicals industry. This latter portion of the emphasis ties in well to the Agricultural and Forest Products efforts within OIT. The program has employed industrial review panels to provide guidance relevant to the needs of industry and in specific solicitation efforts. The BCTR program fully supports the visioning process, and has established an R&D agenda responsive to those needs. Section 4.1 shows how the BCTR R&D agenda matches with the applicable technology needs identified in section 3.0 by each of these industries. While this is a more comprehensive analysis, the actual work performed in FY 1997 focuses on those areas highlighted in bold print or an asterisk.

4.1 Programmatic Relevancy

4.1.1 OIT's Industries of the Future

INDUSTRY NEEDS (FROM VISIONS AND ROADMAPS)	OIT'S BCTR PROGRAM/PROJECT ACTIVITIES*
Chemical Industry of the Future: New Chemical Sciences and Engineering Technology	
Synthesis - Using computational techniques to guide synthesis by theory and molecular modeling (The roadmap on catalytic syntheses suggested further refinement to four areas 1. Enabling catalyst design, 2. Techniques for high-throughput synthesis and analysis of catalysts, 3. Better in-situ catalyst characterization, and 4. Catalysts with specific site architecture.)	**Computer-Aided Catalyst design (CACD) to enable design of chemical and biological catalysts** and development of biomimetic catalysts which have specific site architecture. Tools to characterize catalysts.
Bioprocess and Biotechnology - Biocatalysts for alternate biochemical routes and enhanced biochemical engineering capabilities/applied engineering skills	**Biocatalyst development for aqueous and nonaqueous systems, advanced bioreactors, and bioprocessing systems for high yields and productivities***
Materials Technology - Prediction of materials properties from the molecular to the macroscopic level including structure property relationships and computational techniques	**CACD for** chemical **and biological materials***
Process Science and Engineering - Reactor design principles/tools for new, emerging process industries and design principles/tools focused on source waste reduction	**Advanced bioreactor design for aqueous and nonaqueous high yield/productivity systems.*** Process simulation models that identify waste streams from unit operations.

INDUSTRY NEEDS (FROM VISIONS AND ROADMAPS)	OIT'S BCTR PROGRAM/PROJECT ACTIVITIES*
Computational Technologies - Development of computational fluid dynamics tools, user interfaces for molecular modeling, design, kinetics, and thermodynamic properties	**Multiphase bioreactor predictive models**, new computer codes for commercial software and user interfaces for design of catalysts (DeCAL, CEMBOS, Cell Multipole Method, etc.)
Agriculture or Plant/Crop-based Renewables Industry	
Synthesis of novel molecules	**CACD for** chemical **and biological materials*** Catalytic antibodies.
Bioprocesses	**Biocatalyst development for aqueous and nonaqueous systems, advanced bioreactors, and bioprocessing systems for high yields and productivities***
Modified or current chemistry	Process simulation models that identify waste streams from unit operations.
Forest Products	
Improved capital effectiveness	**Multiphase bioreactor predictive models.** Production of low-cost fungal pulping and bleaching enzymes
Sensors and controls	
Environmental performance	**Biocatalyst development for aqueous and nonaqueous systems, advanced bioreactors, and bioprocessing systems for high yields and productivities*** Process simulation models that identify waste streams from unit operations
Petroleum Refining[5]	
Tools to design improved catalysts and modeling of catalyst structure	**Computer-Aided Catalyst design (CACD) to enable design of chemical and biological catalysts*** and development of biomimetic catalysts which have specific site architecture. Tools to characterize catalysts.

[5]While this industrial sector does not have a vision or any roadmaps, it has identified some key areas for R&D development. BCTR project technologies that are related to these key areas are identified. Note these have much in common with the chemical industry. While projects were ongoing in these areas, their focus was not for the refining industry.

INDUSTRY NEEDS (FROM VISIONS AND ROADMAPS)	OIT'S BCTR PROGRAM/PROJECT ACTIVITIES*
Computational fluid dynamics for multiphase flows	Analogous modeling with bioreactor systems which possess three-phase dynamics. Models of zeolite catalyst systems for fluidized catalytic crackers (FCC)kinetics, diffusion, and fluid dynamics.
Hydrogen production technologies	Novel microbial systems for production of bioplastics can be adapted to produce hydrogen.

* Active efforts in FY 1997 are in bold or marked with an asterisk. Others items noted have been part of OIT efforts in the past

4.1.2 Relationship to EPAct and the Department of Energy

The long-term research efforts of the OIT BCTR projects have been centered in six areas: CD - Catalyst Design; PD - Process Design; PDV - Process Development; BD - Biocatalyst Development; BM - Biocatalyst Modification; BPS - Bioprocess Design/Systems. For FY 1997 the areas receiving emphasis were limited to Biocatalyst Development and Modification, and Bioprocess Design/Systems. These R&D activities support EPAct Sections 2103 (Pulp and Paper) via BD and BPS; 2107 (Improving Efficiency in Energy-Intensive Industries) via all six areas; 2108 (Energy Efficient Environmental Program) via CD, PD, PDV; and 2202 (National Advanced Manufacturing Technologies Initiative), via all six areas. Collectively, these research efforts respond to the Administration's focus on **Global Climate Change** and **Pollution Prevention** through the development of high-productivity, low-waste-generation chemical and biological technologies.

The Department of Energy The Department has five core business areas: Industrial Competitiveness, Energy Resources, National Security, Science and Technology, and Environmental Quality. Of these, the Program directly supports two, Industrial Competitiveness and Science and Technology, through its R&D efforts that address feedstock and energy flexibility and advanced energy and process technology development and deployment. In addition, the Program is responsive to the core values of the Department, particularly in regard to being *Customer-Oriented* (through collaborative research with XX U.S. companies representing the Petroleum Refining, Chemical and Allied Products, and Renewables Industries); in being *Committed to Excellence* (a number of BCTR researches have received meritorious awards for their scientific contributions); and to being a team and advocating teamwork (the Program is a collaborative R&D effort involving over XX U.S. industries, 7 national laboratories, over 10 universities, various government agencies, and a number of DOE offices).

4.2 BCTR Program Overview

The mission of the DOE BCTR Program is *to provide evidence of the technical and economic feasibility of advanced chemical and biological concepts that can improve the energy utilization, operation efficiency, and environmental soundness of U.S. industry process operations as defined in industry crafted vision and roadmap documents. The program combines science with technology* to develop novel and revolutionary process concepts; to improve conventional

process approaches; and, if the economics and market needs warrant, to facilitate the **introduction** of these technological advances in the chemicals, petroleum refining, other renewable, and emerging biotechnology industries. Figure 4.1 describes the role of the program in bridging science and technology to industry and other DOE programs to develop useful products and processes.

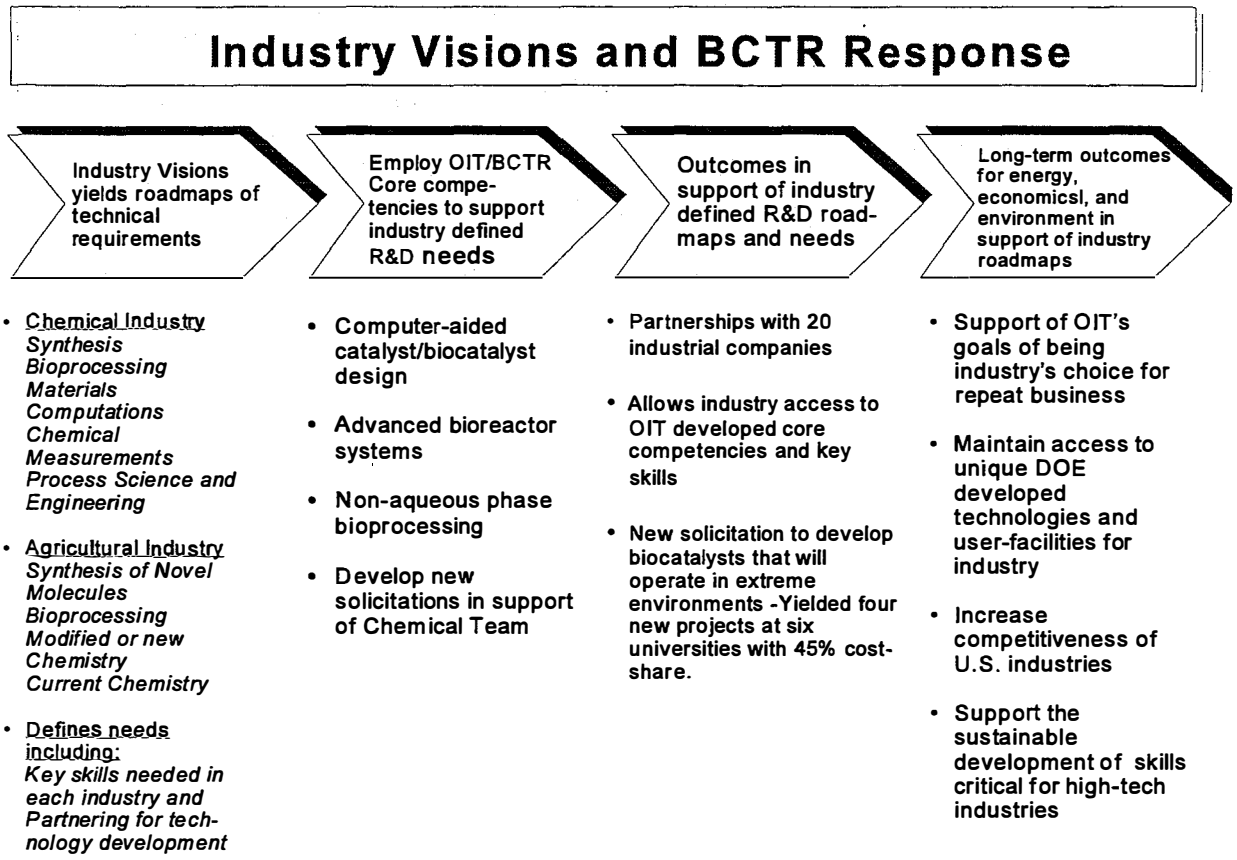


Figure 4.1 Industry Visions and BCTR Response

4.2.1 Program Goal and Objectives

The overall goal is **to stimulate and nurture the development, testing, and commercial deployment of advanced chemical and biological technologies** that offer the potential to reduce energy use, enhance economics, and mitigate environmental impact within the industries providing visions and roadmaps for use by DOE and other federal funding agencies.

The current program objectives involved in reaching this goal are as follows:

1. Develop the technology base required to overcome limitations in the use of biotechnology in the chemical and agriculture industries.

2. Pursue development of enabling technologies that offer broad application and use in support of the development of advanced chemical and biological processes.

3. Maintain partnering capabilities to enhance the move and development of technologies toward commercial deployment.

4.2.2 Strategy

The strategy to meet these goals and objectives involves being responsive to the industry visions and roadmaps. OIT and this program will serve as a catalyst and an integrator of

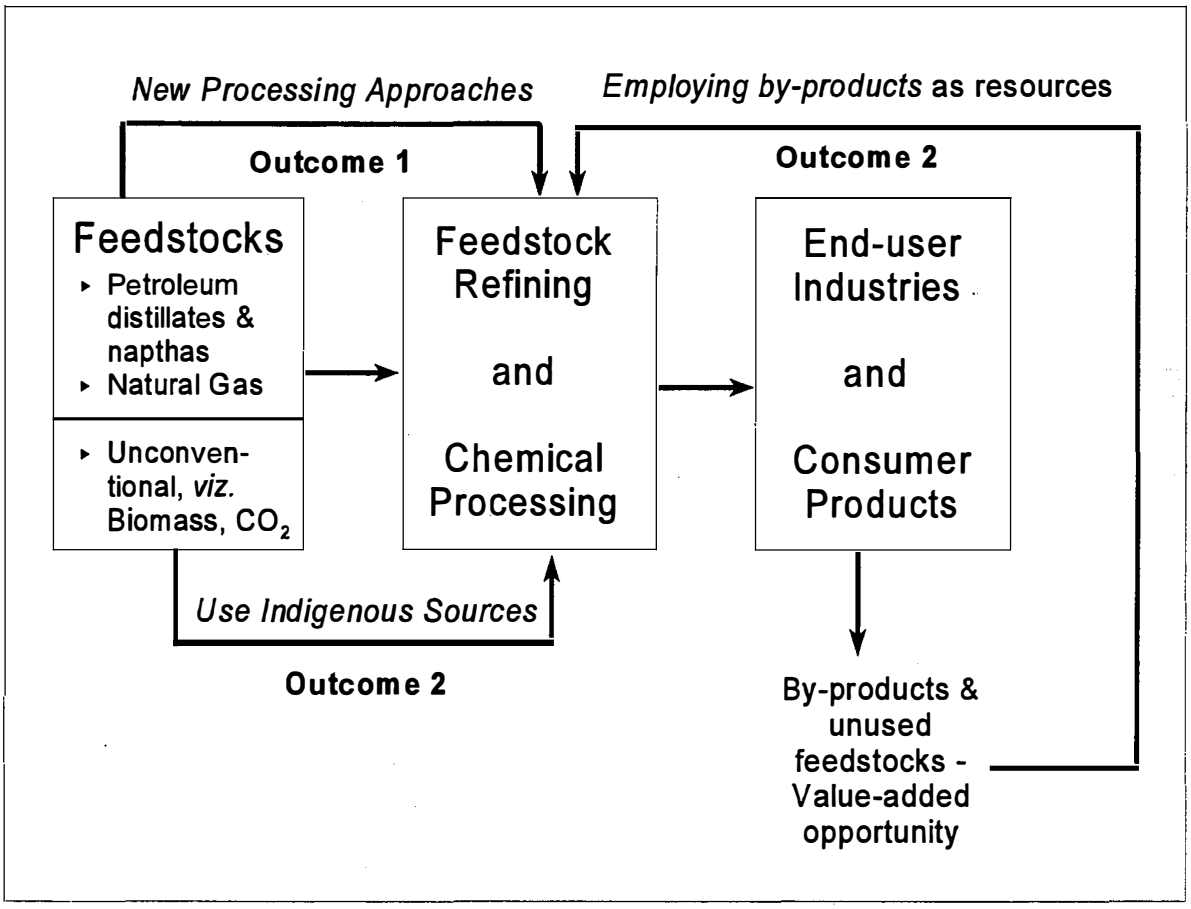


Figure 4.2 BCTR strategic approach to meeting quality metrics

capabilities within the DOE system for use by industry. By meeting these program goals and objectives, two general outcomes are expected and depicted in Figure 4.2 and are based on the energy-supply/demand issues of process operations. Outcome 1 - **Boost Domestic Energy Supply** by replacing petroleum feedstocks with indigenous resources and waste material from the industrial, commercial, utility, and residential sectors. Outcome 2 - **Cut Energy Demand** by developing technologies that reduce energy demands for industrial unit operations.

4.2.3 Metrics

The petroleum refinery industry processes about 22 quads of petroleum-based feedstocks for the transportation industry and about 3.0 quads of petroleum-based feedstocks for the chemicals industry, accounting for about 26% of all energy forms in the U.S. Through the strategy outlined above, the program has identified an accumulated energy savings or petroleum displacement potential of about 2.1 quads by the year 2020. This target assumes a fully funded R&D agenda and is gauged by the primary energy use that could be impacted by successfully deploying innovations developed within the program. Of this total energy saving, 1.8 quads have been identified for the chemicals industry and 0.3 quads in the petroleum industry. Many of these savings are reflected in the agricultural industry sector as well, and new energy saving options may also become available from this sector once new efforts are underway.

Timely technology development, transfer, and industrial implementation are necessary to achieve the projected 2.1 quads energy savings impact. This energy savings is envisioned to occur through improvements in conventional processing (1.1 quad reduction in energy demand via Outcome 2) and introducing alternative feedstocks, including recycled waste material (1.0 quad boost in energy supply via Outcome 1). Combined, these two outcomes for the strategy provide a rational approach for optimum resource use combined with energy efficient, cost-effective, and environmentally sound process technologies. In this regard, three cornerstone reports by Pacific Northwest Laboratories and Battelle Columbus Memorial Institute identify a potential saving of about 1.1 quads in process and feedstock energy by improving catalyst and process efficiencies. The three reports, *Brief Characterizations of the Top 50 U.S. Commodity Chemicals, Characterizations of the Top 12 U.S. Commodity Polymers,* and *The Top 50 Commodity Chemicals: Impact of Catalytic Process Limitations on Energy, Environment, and Economics* provide extensive information on the key processes used in the U.S. Chemical Industry. Opportunities to achieve the remaining 1.0 quad were identified in a study conducted by the Alternative Feedstocks Program.

4.3 BCTR Program Structure

The program is structured to address the industry needs discussed in Section 2.0. This structure emphasizes two core technological approaches, chemical and biological, as outlined below:

Biological and Chemical Technologies	
Chemical Approaches	**Biological Approaches**
Catalyst Design	Biocatalyst Development
Process Design	Biocatalyst Modification
Process Development	Bioprocess Design/Systems

For FY 1997 the BCTR effort has focused on the biological approaches since these seem to be germane to both the chemicals and agricultural industry sector visions. Chemical approaches are being approached through the DOE solicitation. Although the DOE chemical solicitation does offer work in bioprocessing, the technical R&D issues were not well-defined. Through a roadmapping exercise conducted by industry, academia, and DOE a specific area was targeted. This target area involved the development of biocatalyst systems that could operate in the extreme environments typically found in the chemical processing industry. The intent is to target enabling capabilities that would allow these novel biocatalytic systems to be employed by not only the chemical industry but also the agricultural sector. Below is a brief description of the chemical approaches although they were not pursued in FY 1997 and the biological approaches including some milestones to mark progress.

4.3.1 Chemical Approaches

Catalyst Design - Objective – Create tools to aid in the design of chemical catalysts for chemicals and materials production. Goal – Chemical Computer aided catalyst design (CACD) first-generation model development by 1996 (*completed*) and commercial-model applications by 2005.

Process Design - Objective – Create process model(s) for intelligent chemical processing of feedstock and waste streams. Goal – First-generation model by 2000 and commercial model applications by 2003.

Process Development - Objective – Exploratory development of advanced chemical/ petrochemical process systems to meet needs identified by these industries. Goal – Complete systems/process analysis of at least one advanced process per year.

4.3.2 Biological Approaches

Biocatalyst Design - Objective – Create tools to aid in the design of biochemical catalysts for chemicals and materials production. Goal – Third-generation Biological CACD tools/models developed by 1998 and additional commercial model applications by 2000.

Biocatalyst Development - Objective – Identify and characterize at least one new biocatalyst per year. Goal – Adapt five (5) biocatalysts for commercial application by 2000.

Bioprocess Development - Objective – Exploratory development of advanced biochemical process systems to meet needs identified by these industries. Goal – Complete systems/process analysis of at least one advanced process per year.

5.0 Project Descriptions

1. Project Title Synthetic Biocatalysis by Hyperthermophilic Enzymes
Principal Investigator(s) Michael Adams and Robert Kelly
Project Sites University of Georgia and North Carolina State University

Description:

This work focuses on strategic opportunities to catalyze synthetic reactions with enzymes from hyperthermophiles, which are organisms that grow near and above 100°C. Using several thermophilic enzymes which have recently been purified in our laboratories, a series of model synthetic reactions will be investigated which make strategic use of the remarkable stability and activity of these enzymes at extreme temperatures. The results will be essential in defining the areas of opportunity for enzymes from these unique microorganims in the chemical process industry. This is a joint effort between two departments: The Department of Biochemistry and Molecular Biology at the University of Georgia and the Department of Chemical Engineering at North Carolina State University. Enzymes discovered or evaluated in one laboratory will be analyzed for catalytic activity in the other laboratory. The laboratories will also be able to develop techniques and procedures in concert which should speed development time of potentially useful enzymes.

The enzyme systems employed will be those that are hydrolytic in aqueous phase. Activity in the reverse direction is expected to yield novel chemical synthetic routes of interest to the chemical process industry. The enzymes have been obtained from the hyperthermophile, *Pyrococcus furiosus* (Pf) that grows optimally at 100°C. Kilogram quantities can be grown in our laboratories from which we have isolated 20 different enzyme types. These will be used as basis for our studies along with additional enzymes provided by Diversa, Inc. There is an extensive range of chemical processes to which such enzymes have relevance, including potential formation of C-N and C-O bonds which are of interest to the synthetic chemical industry.

1996-1997 Accomplishments:

New start in FY 1997

1998 Planned Activities:

Tasks will occur simultaneously at the two laboratories. Activities are outlined below, some are actually joint between the two laboratories, others are specific. These are the tasks designed for a 15 month effort.

Task 1: Obtain recombinant and native forms of up to three proteases from *Pyrococcus furiosus* and prepare solvent-soluble forms for homogeneous catalysis following protocols developed by the U. Georgia group for hyperthermophilic hydrogenases. Several thermostable lipases will also be prepared for evaluation as homogeneous catalyts. These lipases will be provided by the industrial enzyme company, Diversa, Inc.

Task 2: Evaluate catalytic properties of chemically modified proteases and lipases in solvent and aqueous solutions.

Task 3: Evaluate catalytic properties of several *Pyrococcus furiosus* exo-acting glycosidases for function in organic solvents for transglycosylation activity.

Annual Technical Summary Report:

None. This is a new start which did not begin until late in FY 1997 or early FY 1998.

2.	**Project Title**	Enzyme Catalysts for a Biotechnology-Based Chemical Industry
	Principal Investigator(s)	Frances H. Arnold
	Project Site	California Institute of Technology

Description:

Improving understanding of the molecular basis of protein stability and enzyme catalysis, combined with the ability to create large quantities of proteins of virtually any amino acid sequence, gives us the ability to redesign natural proteins to fit the requirements of industrial applications. As a result, biotechnologists no longer have to limit themselves to designing processes around natural biocatalysts--designing a biocatalyst to fit the process is gradually becoming an achievable goal. The ability to carry out biochemical syntheses in organic solvents, where solubilities are greatly enhanced and new chemistries are available, greatly expands the scope and potential applications of biocatalysis in the chemical industry. Unfortunately, most enzymes respond unfavorably to transfer to polar organic solvents; they are highly destabilized and their catalytic activities are often reduced by orders of magnitude. The overriding goal of this project is to develop design rules for engineering enzymes at the level of their amino acid sequences to improve stability and catalytic activity in polar nonaqueous solvents. The strategy employed is to use both random and site-directed mutagenesis techniques to alter the amino acid sequence of Subtilisin E, a serine protease with numerous potential applications in organic synthesis and preparation of novel polymers. This research will provide general tools and design rules for engineering stable and efficient biological catalysts. This project seeks to understand and enhance engineering enzyme stability and reactivity in nonpolar and nonaqueous solvents, using the prototype enzymes. Design tools will be provided to guide the selection and efficiency of macromolecular catalysts.

1996-1997 Accomplishments:

Directed evolution of thermostable pNB esterase and subtilisin

We have dramatically increased the thermostability of an industrial enzyme (pNB esterase from Eli Lilly) by sequential rounds of point mutagenesis and screening. Six generations of directed evolution have increased the enzyme T_m (melting temperature) from 49°C to 63°C. The amino acid substitutions responsible for this increase in stability have been identified. Similarly, we have significantly increased the stability of subtilisin E by both random mutation and recombination.

Analysis of recombination

This year we analyzed the role of recombination in directed evolution (*J. Mol. Biol., 1997*). Computer simulations of recombination and screening illustrate the trade-offs between recombining fewer parent sequences (in order to reduce screening requirements) and lowering the potential for further evolution. Search strategies that may substantially reduce screening requirements in certain situations were described.

Optimizing the DNA shuffling technique and its use to identify functional mutations

We described modifications to the method of DNA shuffling that significantly reduce the error rate (*Nucleic Acids Research*, 1997). Furthermore, we used the high-fidelity *in vitro* gene recombination coupled with sequence analysis of a small sampling of the shuffled library exhibiting the evolved behavior to identify those mutations responsible for the behavior in a background of neutral and deleterious mutations (*PNAS*, 1997).

Major technical advance

A new method for *in vitro* recombination

During this past year we have filed a utility patent application and prepared a publication on a novel method for *in vitro* recombination of mutated genes. This recombination method provides for the rapid accumulation of beneficial mutations and removal of neutral and deleterious mutations from a population of parent genes, enhancing the rate of evolution (*Nucleic Acids Research*, submitted).

1998 Planned Activities:

We plan to continue using random recombination and point mutation to improve the thermostability of pNB esterase and several subtilisins. In this project we will be investigating the linkage between key features, such as thermostability and high activity. It has been claimed that high thermostability is necessarily linked to lower activity at room temperature. We are exploring these fundamental linkage questions with thousands of mutants. We will also continue to investigate alternative methods for recombination of beneficial mutations and removal of deleterious mutations.

Annual Technical Summary Report

Enzymes have enormous potential for reducing energy requirements and environmental problems in the chemicals and pharmaceutical industries. Those who try to use enzymes, however, are stymied by the fact that these catalysts have evolved over billions of years to perform very specific biological functions and to do so within the context of a living organism. Some of the features required for function in a complex chemical network are undesirable when the catalyst is lifted out of context (e.g. product inhibition). Conversely many of the properties we wish an enzyme would have clash with the needs of the organism, or at least were never required: high stability, the ability to function in nonnatural environments or catalyze nonnatural reactions.

The explosion of tools that has come out of molecular biology during the last 20 years has made it possible for us to evolve enzymes for features never required in nature. We can speed up the rate and channel the direction of evolution by controlling mutagenesis and the accompanying "selection" pressures. Darwinian evolution carried out in the test tube offers a unique opportunity for biotechnology: the ability to tailor enzymes for optimal performance in a wide range of applications. Thus it is possible, for example, to evolve enzymes that carry out reactions on nonnatural substrates or even to carry out reactions for which there is no counterpart in nature. Due to the vast size of the potential sequence space, however, explorations by directed evolution must be guided by sound principles and workable strategies. During the past year this laboratory has continued to make significant progress in the evolution of industrial enzymes as well as in developing general methods for *in vitro* evolution.

Directed evolution of thermostable esterases

B. subtilis para-nitrobenzyl (pNB) esterase catalyzes the hydrolysis of the para-nitrobenzyl esters of various cephalosporin-type antibiotics, a necessary step in their large-scale synthesis. In previous years we dramatically increased the activity of this enzyme toward nonnatural substrates and in nonnatural environments. This past year we enhanced its thermostability and developed general methods by which the stability of virtually any enzyme can be improved. Using sequential rounds of mutagenesis and screening, we accumulated amino acid substitutions which increased the T_m of pNB esterase by 14°C. Similarly, we significantly increased the stability of subtilisin E by both random mutation and recombination.

A powerful method for *in vitro* recombination

A simple and efficient method for *in vitro* mutagenesis and recombination of polynucleotide sequences was developed this year. Recombination is a powerful tool for directed evolution. The method involves priming template polynucleotide(s) with random-sequence primers and extending to generate a pool of short DNA fragments which contain a controllable level of point mutations. These fragments are subjected to denaturation followed by annealing and further enzyme-catalyzed DNA polymerization to produce a library of full-length sequences. Screening or selecting the expressed gene products leads to new variants with improved functions.

A method for identifying functional mutations

This year we also developed a method for distinguishing functional from nonfunctional or deleterious mutations in homologous genes. High-fidelity *in vitro* gene recombination ('DNA shuffling') coupled with sequence analysis of a small sampling of the shuffled library exhibiting the evolved behavior allows identification of those mutations responsible for the behavior in a background of neutral and deleterious mutations. Functional mutations are expected to occur in 100% of the sequenced screened sample; neutral mutations are found in 50% on average, while deleterious mutations do not appear at all. When used to analyze ten mutations in a laboratory-evolved gene encoding a thermostable subtilisin E, this method rapidly identified the two responsible for the observed protease thermostability; the remaining eight are neutral with respect to thermostability, within the precision of the screening assay. A similar approach, coupled with selection for growth and survival of the host organism, could be used to distinguish adaptive from neutral mutations.

Analysis of recombination

This year we analyzed the role of recombination in directed evolution (4). Computer simulations of recombination and screening illustrate the trade-offs between recombining fewer parent sequences (in order to reduce screening requirements) and lowering the potential for further evolution. Search strategies that may substantially reduce screening requirements in certain situations were described.

References

Shao, A., H. Zhao, L. Giver and F. H. Arnold "Random Priming *in vitro* Recombination: an Effective Tool for Directed Evolution,", *Nucleic Acids Research*, submitted.

Zhao, H. and F. H. Arnold "Functional and Non-functional Mutations Distinguished by Random Recombination of Homologous Genes,", *Proc. Natl. Acad. Sci. USA*, **94**, 7997-8000 (1997).

Zhao, H. and F. H. Arnold "Optimization of DNA Shuffling for High Fidelity Recombination,", *Nucleic Acids Research.*, **25**, 1307-1308 (1997).

Moore, J.C., H.-M. Jin, O. Kuchner and F. H. Arnold "Strategies for the *in vitro* Evolution of Protein Function: Enzyme Evolution by Random Recombination of Improved Sequences," *J. Molecular Biology* **272**, 336-347 (1997).

3.	**Project Title**	*In Vitro* Evolution of Oxidation Catalysts: Toward Efficient Biochemical Production of Commodity Chemicals
	Principal Investigator(s)	Frances H. Arnold
	Project Site	California Institute of Technology

Description

This effort embarks on the *in vitro* evolution of key oxidative enzymes with the ultimate goal of creating efficient oxidation biocatalysts for the chemical industry. The target enzymes for this project are cytochrome P450, a soluble, cofactor dependent monooxygenase and chloroperoxidase (CPO) which exhibits properties and catalytic activities typical of peroxidases as well as P450-type monooxygenases. These enzymes catalyze a variety of oxidation reactions including epoxidation (CPO, P450), conversion of alkanes to alcohols (P450), and halogenation (CPO). The P450 utilizes molecular oxygen as the oxidant and requires an NAD(P)H cofactor as well as electron-tansfer protein(s), while CPO utilizes hydrogen peroxide as the oxidant and requires no cofactor or protein partners. In order to evolve these natural enzymes into useful industrial catalysts, their stability and activity, particularly in organic media, will have to be increased. Our specific goals for this project will be to 1) express these enzymes in bacterial (or yeast) hosts suitable for molecular evolution and 2) develop rapid, miniaturized and automatable screening protocols for these properties. With further research support from industrial partners, we will utilize these systems for the molecular evolution of specific industrial catalysts.

1996-1997 Accomplishments:

New effort in FY 1997/1998

1998 Planned Activities:

Development of expression systems suitable for in vitro *evolution*. This task focuses on the functional expression of the target oxidative enzymes, CcP, HRP, CPO, Cyt P450 and perhaps one more candidate in microbial hosts. Genes for CcP, HRP, CPO have been obtained and constructs for expression in *Escherichia coli* and yeast will continue to be prepared and tested. Ideally the proteins will be screened into the culture medium where they can be conveniently assayed.

Development of screens for enhanced biocatalyst stability. Screening assays sufficiently sensitive to detect small enhancements in stability will be put into place. Assay and incubation conditions will be chosen carefully to reflect desired properties in the ultimate catalysts.

Characterization of wild type enzymes: stability in organic solvents and against hydrogen peroxide. Wild type enzymes will be fully characterized with respect to their stabilities under various conditions. This defines the starting point for directed evolution of these properties. Wild type enzymes expressed in microbial hosts may differ fron naturally-occurring enzymes, due to differences in post-translational modifications. Thus, the recombinant enzymes must also be compared to the naturally occurring ones.

Application of directed evolution to enhance biocatalyst stability. Once tasks 1 and 2 have been achieved for a given target enzyme, random mutant libraries will be generated and screened for improved stability in organic solvents and toward hydrogen peroxide.

Annual Technical Summary Report:

None. This is a new start which did not begin until late in FY 1997 or early FY 1998.

4.	**Project Title**	Theory of Biocatalysis: Electron Transfer Reactions
	Principal Investigator	David Beratan
	Project Site	University of Pittsburgh Department of Chemistry

Description:

The principal project goal is the computer-aided design of enzyme biocatalysts with tailored catalytic rates. The focus of this project has been on electron transfer enzymes, a major class of biological enzymes. In this project, the theory of *protein electron tunneling pathway* has been developed, and implemented as computer models, which connect enzyme electronic structures and their reaction rates.

Specific project objectives include:

1. Develop algorithms to map the key residues in proteins between electron transfer sites that mediate electronic coupling and allow electron transfer reactions to proceed with great speed and specificity;
2. Identify "hot" and "cold" spots with respect to electron transfer in native and modified proteins;
3. Develop an understanding of primary, secondary, tertiary, and quaternary structural effects on electron transfer rates;
4. Use knowledge gained to stabilize energetic charge separated states and to enable the development of semisynthetic/modified protein energy conversion systems.

1996-1997 Accomplishments:

We achieved all four milestones planned for the year and have published our work extensively in the scientific literature.

<u>Milestone/Progress</u>

- Compute rates of two-electron transfer as a function of free energy in stepwise and concerted regimes.

  We have expanded our studies of two-electron transfer reactions to known reactions in solultion, and used the formalism to treat three-state electron transfer processes.

- Estimate electron transfer rates in experimental DNA systems constructed by the Meade, Barton, and Harriman groups.

  We have completed an extensive evaluation of DNA electron transfer experiments using our new theories.

- Develop a multiple-pathway representation of tunneling interactions in proteins using the concept of electron transfer contact maps.

  We completed our development of electron transfer contact maps.

- Conduct exploratory investigations of inverse design strategies for new materials.

  We completed our analysis of inverse design strategies.

1998 Planned Activities:

None. Project ended 9/97.

Annual Technical Summary Report:

We have made substantial progress in four areas: (1) multi-electron electron transfer processes, (2) coupled quantum/molecular dynamics studies of protein electron transfer, (3) quantum mechanical studies of DNA electron transfer processes, and (4) inverse design of new materials.

Multi-electron electron transfer processes

Chemical, biological, and electrode based electron transfer (ET) processes involve multi-electron events. However, an adequate framework in which to describe these complex reactions does not yet exist. A theory for two-electron transfer reactions in Debye solvents was developed. The theory is formulated by generalizing Zusman's model of ET reactions to those involving three parabolic wells: one for the doubly reduced donor, one for the singly reduced donor/singly reduced acceptor, and one for the doubly reduced acceptor. The ET processes are described in terms of diffusional motion along a one-dimensional reaction coordinate with tunneling transitions at the intersection points of the parabolas. Two competing mechanisms of two-electron transfer arise. One is a process with two sequential single electron steps $D^{=}A \rightarrow D^{-}A^{-} \rightarrow DA^{=}$. The other involves ET in one concerted two-electron step ($D^{=}A \rightarrow DA^{=}$). The general rate expressions for two-electron transfer are obtained. When the stepwise mechanism dominates, the free energy of activation is predicted to depend upon the driving forces of the two sequential steps but is independent of the overall driving force of the reaction. When concerted two-electron transfer dominates, the Marcus relation is obtained with a reorganization energy associated with the shift of two electrons. The dynamical solvent effect in two-electron ET processes is predicted to be unusual. Two distinct regimes exist, each with essentially linear τ_L dependence (with τ_L the solvent longitudinal relaxation time): one for slow solvents and one for fast solvents. A combination of solvent and free energy studies could be used to elucidate the mechanism of multi-electron processes in chemical and biological systems.

Coupled quantum/molecular dynamics studies of protein electron transfer

Coupled quantum mechanics and molecular dynamics simulations were performed to investigate the role of secondary structure and small geometrical fluctuations on the long range electronic coupling in peptides. Molecular dynamics simulations were performed to obtain typical relevant conformations of the oligopeptides and self-consistent Hartree-Fock calculations at the semiempirical quantum theory level were performed to extract the long range electronic coupling. α-helical oligopeptides in the initial structure show a dominant contribution due to hole mediation over a long range of tunneling energy, while the initial β-stranded oligopeptides have more equal contributions from both hole and electron mechanisms. Geometrical fluctuations lead to drastic changes in the couplings, including cancellations arising from destructive interference from distinct pathways or between the electron and hole contributions to the

coupling. The computations highlight the danger of drawing conclusions from electronic structure calculations of electronic coupling in proteins based upon computations on single geometries.

DNA electron transfer processes

The earliest studies of electron-transfer proteins raised the question of whether or not π-electron residues might facilitate electron transport. Three recent long-range electron-transfer experiments utilizing DNA bridges revisit this provocative, yet unresolved, question. The distance dependence of electron transfer in DNA is not a matter of purely academic concern; it controls the mechanism of DNA damage and repair in cells, and is being exploited in new molecular probes of DNA sequence. We have presented a theoretical analysis based upon very large scale self-consistent-field quantum calculation of all valence electrons (as many as ~ 3300) in these three systems. This computation is the first performed on such large macromolecules, and also the first to extract long-range electronic interactions at this level of theory. DNA electron-transfer is found to be mediated by through-space interactions between the π-electron containing base pairs, but the magnitude of the coupling facilitated by this channel drops rapidly with distance, as a consequence of the ~ 3.4 Å noncovalent gap between base pairs. These predictions are in agreement with most of the experimental data. The rapid decay of electron-transfer rates with distance computed here suggests that biologically controlled DNA electron-transfer events, of importance in DNA repair, must function over a relatively short range. Moreover, the predicted distance dependence of electron-transfer in DNA is strikingly close to that found in proteins.

Inverse design of new materials

An "inverse" molecular design strategy was developed to assist in the development of new molecules with optimized properties. This approach is based on a molecular orbital view and can be used to tailor ground state or excited state properties subject to particular constraints. In this scheme, wave functions are sought that optimize a chemical or electronic property, and then a Hamiltonian is constructed that generates these optimized wave functions. Analysis of the chemical properties in the optimized systems may suggest new synthetic targets. Examples were analyzed in order to optimize certain optical properties in simple structures.

References

S.S. Skourtis and D.N. Beratan, "High and Low Resolution Theories of Protein Electron Transfer," *J. Bioinorg. Chem.*, 2, 378-386 (1997).

D.N. Beratan, S. Priyadarshy, and S.M. Risser, "DNA: Wire or Insulator?" *Chemistry and Biology*, 4, 3-8 (1997).

J.N. Onuchic, S.M. Risser, S.S. Skourtis, D.N. Beratan, "The Design of Starburst Dendrimer Electron Transfer Systems," in *Molecular Electronics,* M.A. Ratner and J. Jortner, editors, Chemistry for the 21st Century Series, Blackwell Science, Oxford, 1997.

J. Wolfgang, S.M. Risser, S. Priyadarshy, D.N. Beratan, "Secondary Structure Conformations and Long-range Electronic Interactions in Oligopeptides," *J. Phys. Chem. B*, 101, 2987-2991 (1997).

L.D. Zusman and D.N. Beratan, "A Three-state Model for Two-electron Transfer Reactions," *J. Phys. Chem. A*, 101, 4136-4141 (1997).

S. Priyadarshy, S.M. Risser, and, D.N. Beratan, "Dna Is Not a Molecular Wire: Protein-like Electron Transfer in an Extended Pi-electron System," *J. Phys. Chem.*, 100, 17678-17682 (1996).

S. Priyadarshy, S.M. Risser, and, D.N. Beratan, "Dna Double Helix Mediated Long Range Electron Transfer," *Int. J. Quantum Chem. Quantum Biol. Symp.*, 60, 65 -71 (1996).

C. Kuhn and D.N. Beratan, "Inverse Strategies for Molecular Design," *J. Phys. Chem.*, 100, 10595-10599 (1996).

L.D. Zusman and D.N. Beratan, "Two-electron Transfer Reactions in Polar Solvents," *J. Chem Phys.*, 105, 165-176 (1996).

I.V. Kurnikov and D.N. Beratan, "*Ab initio* Based Effective Hamiltonians for Long Range Electron Transfer: Hartree-fock Analysis," *J. Chem. Phys.*, 105, 9561-9573 (1996).

S. Priyadarshy, S.S. Skourtis, S.M. Risser, and D.N. Beratan, "Bridge Mediated Electronic Interactions: Differences Between Hamiltonian and Green Function Partitioning in a Non-orthogonal Basis," *J. Chem. Phys.*, 104, 9473-9481 (1996).

D.N. Beratan and J.N. Onuchic, "Chapter 2: *The Protein Bridge Between Redox Centers*," in Protein Electron Transfer, D.S. Bendall, editor, Bios Scientific Publishers, Oxford, pp. 23-42, 1996.

J.N. Onuchic and D.N. Beratan, "Appendix A: *Electron Transfer Rate Calculations*," in Protein Electron Transfer, D.S. Bendall, editor, Bios Scientific Publishers, Oxford, pp. 273-283, 1996.

5.	**Project Title**	Advanced Bioreactor Concepts for Non-traditional Bioprocessing
	Principal Investigators	Brian Davison, John Barton, and Laurel O'Connor
	Project Site	Oak Ridge National Laboratory

Description:

This project will advance practical and fundamental knowledge of dynamics and immobilization for nontraditional biocatalytic systems. The goals of this project are as follows: development and testing of new nonaqueous bioreactor configurations; understanding and modeling of the kinetic properties of biocatalysts and the dynamics of bioreactor systems; and development of enhanced bioreactors for continuous production of industrial products (enzymes, chemicals, etc.).

1996-97 Accomplishments:

- The program centered upon the area of nonaqueous bioprocessing. The use of enzymes in organic media for chemical production offers several advantages over traditional heterogeneous catalysis, including greater substrate solubility, elimination of undesirable side reactions with water, thermodynamic equilibria more favorable than those in water, and easier and less expensive product recovery. The focus was on organic- and gas-phase biocatalysis in reactor systems in which scaleup and other design considerations were likely to be of concern. ORNL also contributed to the development of a document entitled "Biotechnology for National Industrial Needs, part II: Development of Biotechnology Capabilities in Extreme Environments for the Chemical and Related Products Industry." We also assisted David Boron and Gene Petersen in the development of a companion roadmap for this area.

- Gas-phase transesterification reactions were demonstrated in both batch and continuous reactor configurations for production of organic compounds such as butyl acetate and isoamyl acetate. Short-term trials have demonstrated the feasibility of this unique nonaqueous approach for chemical synthesis which is particularly amenable for situations in which the substrates are sparingly soluble in water. Effects of pH, humidity, and synthesis type on activity were measured.

- Immobilization and reactor design were investigated for the use of high-efficiency immobilized enzymes for dehalogenation in a CRADA with Dow Chemical. Dehalogenation addresses a direct industrial need to transform a current by-product waste of organic synthesis and allow it to be recycled in the process. The basic organism has been cultured for enzyme production at high yield and at 70-L scale. Dow Chemical has developed a greatly improved genetically modified organism which is being used at ORNL. ORNL has successfully immobilized the active enzyme into a novel porous glass matrix. The enzyme retained 10-20% of the original activity. We are continuing to try and increase the enzyme loading and activity.

- The 19th Symposium on Biotechnology for Fuels and Chemicals was partially supported.

- A cohesive nomenclature and methodology for classifying nontraditional biocatalytic processes was developed and accepted for publication in a peer-reviewed journal. Such categorization provides a means by which to classify new work within the nontraditional arena, as well as to broaden or refine current research. Furthermore, a technology

roadmap was developed which outlines nontraditional areas and their associated development issues which still require examination, in terms of both bridging and fundamental research, before these techniques will be adopted by the private sector.

- ORNL and NREL coordinated a targeted industrial reviewed project selection for key projects in nonaqueous biocatalysis. Subsequently, ORNL arranged subcontracts with the University of Georgia in "Synthetic Biocatalysis by Hyperthermophilic Enzymes" and by University of Iowa and University of California-Berkeley in "Activation and Stabilization of Enzymes for Use in Organic Solvents." ORNL also monitored an ongoing subcontract with the Massachusetts Institute of Technology in Rational Enhancement of Enzyme Performance in Organic Solvents" (see related reports).

1998 Plans:

No FY 1998 funds are budgeted for this effort. FY 1997 projects, tasks and subcontracts will be completed during the FY98 using carryover moneys. Two or more additional publications are expected based upon ORNL work. ORNL will emphasize the vapor-phase enzymatic reactor and organic phase biocatalysis. Several publications will be prepared from the subcontracted work.

Annual Technical Summary Report:

Dow CRADA on Dehalogenation

The dehalogenation of process by-products is being investigated in a confidential CRADA with Dow Chemical Co. The dehalogenase was immobilized in a novel porous glass matrix. This matrix is a rugged stable form which complexes around the enzyme from a sol of alkyl-orthosilicates. The enzyme was active and able to carry out the desired transformation. The retained activity was up to 20% of the original enzymatic activity. We are continuing to improve enzyme loading and activity despite the harsh conditions during the immobilization.

Nonaqueous Gas-phase Biocatalysis

Enzymatic reactions in nonaqueous phase offer a number of advantages over traditional aqueous-based processes, including greater substrate solubility, elimination of undesirable side reactions, more favorable thermodynamic equilibria, and simplified product recovery (Klibanov, 1986). This project investigated reactor systems in which scale-up and other design considerations were likely to be of concern. The proof-of-principle issues examined suggested that vapor-phase reactions can be utilized for production of a variety of esters. Furthermore, these reactions can proceed quickly in the presence of immobilized biocatalyst, with apparent reaction rates over uncatalyzed rates by up to two orders of magnitude.

Both bench- and intermediate-scale continuous gas reactors were constructed to examine nonaqueous enzymatic conversions. A diagram of one typical system is shown as Figure 1. Reactors containing immobilized porcine lipase were operated in both single pass and recycle modes to examine the potential for using gas phase substrates for esterification and transesterification reactions (production of higher molecular weight esters from smaller compounds). Enthalpy and free energy calculations revealed that the forward reactions for these reactions are energetically favorable, but not to such a degree that equilibrium is reached rapidly. Higher conversions (about threefold enhancement) were achieved when continuous modes (versus single pass) were employed due to longer effective residence times with the enzyme. Equilibrium was reached in as little as 3 hours, depending upon operating conditions. The reverse reactions were also found to proceed slowly in the presence of lipase, requiring several hours before reaching equilibrium.

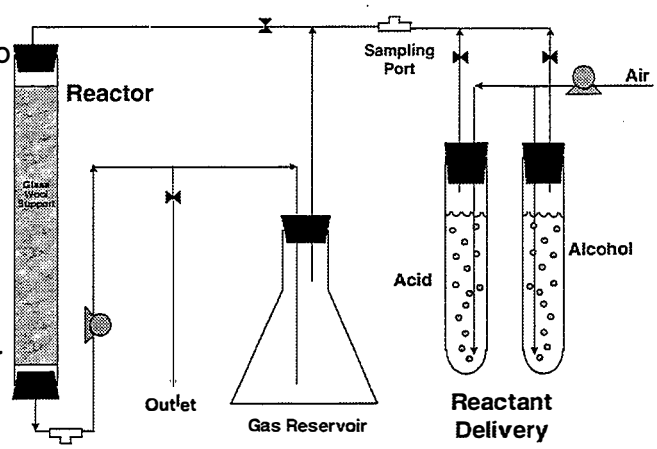


Figure 1. Schematic of gas-phase reactor system.

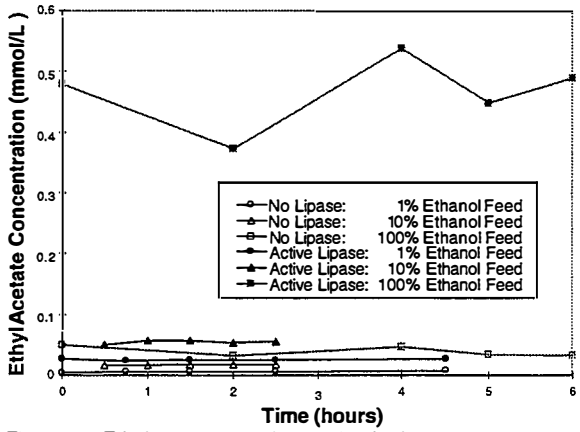


Figure 2. Ethyl acetate production in single-pass experiments. Reactors with active lipase present produce ester at rates approximately one order of magnitude higher than non-lipase controls. Also note that production is strongly affected by initial substrate

An example of ethyl acetate production in single-pass experiments is given as Figure 2. These results showed an order-of-magnitude increase in esterification rate for ethyl acetate with active lipase present as compared to controls with no lipase present. In addition, the direct dependence of the single-pass conversion rate is noted by using varying levels of substrate (ethanol). These results demonstrate the proof-of-principle process for continuous bioconversion using immobilized enzymes in a vapor phase. Another example of continuous batch esterification is provided in Figure 3 for production of ethyl propionate from ethanol and propionic acid. This figure also displays two control results. It is interesting to note a boiled lipase control still displayed measurable activity even after being subjected to high temperature denaturation.

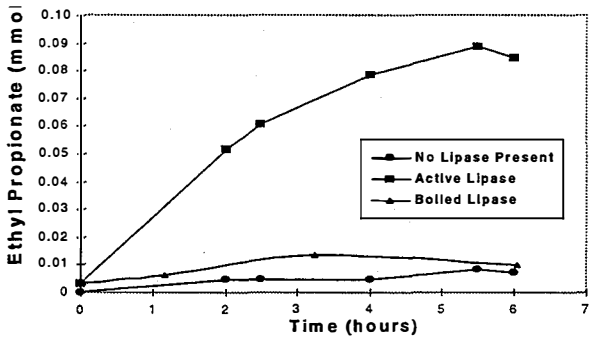


Figure 3. Ethyl propionate production as a function of time. It is interesting to note that boiled lipase still displays some activity above the lipase-free control.

Transesterfication reactions were also examined during FY97. These reactions proceed via combination of esters with alcohols to produce higher molecular weight esters and lower molecular weight alcohols/aldehydes. These reactions are desirable over standard esterification reactions because they do not involve an acidic substrate (such as propionic acid) which can denature the immobilized enzyme. These reaction schemes were used to examine the effect of desiccation and humidity on lipase activity.

One of the reactions examined was

Ethyl Acetate + Isoamyl Alcohol $\longrightarrow$ Acetaldehyde + Isoamyl Acetate

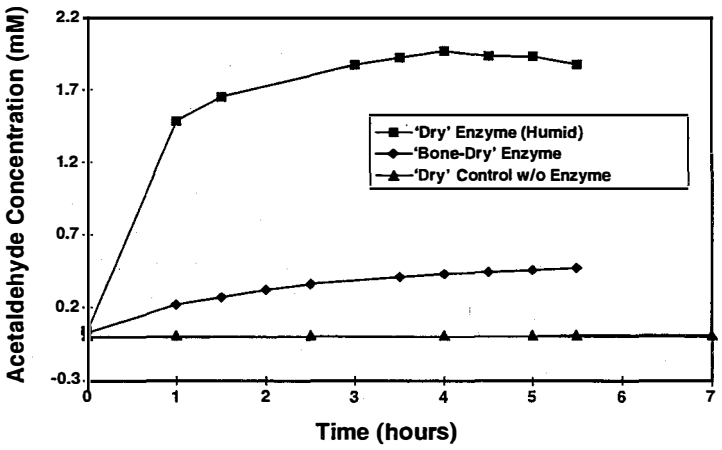


Figure 4. Production of acetaldehyde from a transesterification reaction under conditions of 0% and 60% relative humidity. In the case in which both the gas stream and enzyme exist under more humid conditions, higher production rates are attained. 'Bone' dry lipase is less active, but still produces considerably higher levels than dry controls with no enzyme present.

Some of the data collected for this synthesis are presented in Figure 4. In this experiment, a transesterification reaction was used to produce acetaldehyde and isoamyl acetate (banana oil) under conditions of 0 and 60% relative humidity. Humid conditions resulted in considerably higher production. Interestingly enough, under completely desiccated conditions, immobilized lipase enzyme still displayed substantial activity. These experiments were conducted at ambient temperature and pressure. Immobilized enzyme was dried by placing it in a dessicator for 4 days prior to use.

The packing material used for experiments was clean glass wool. This material was found to provide high surface area, good structural support, and low pressure drop in both bench- and intermediate-scale reactors. The material is inappropriate for aqueous or organic liquid systems due to its tendency to compress under those conditions. BET isotherm measurements of the wool packing material were conducted to accurately determine the interfacial area available for sorption/binding. Clean glass wool packing was found to have 0.75 +/- 0.01 m^2/g of available surface area. Glass wool covered with lipase (in the amounts used for these experiments) was found to have 1.57 +/- 0.01 m^2/g available surface area.

This project was also selected by ORNL Review as worthy for exposure via a State of The Lab Highlight entitled, "Advanced Liquid-Free Bioprocessing Technology."

6.	**Project Title**	Activation and Stabilization of Enzymes for Use in Organic Solvents
	Principal Investigator(s)	Jonathan Dordick and Douglas Clark
	Project Site	University of Iowa and University of California, Berkeley

Description:

This work focuses on developing general techniques to activate and stabilize enzymes for use in organic solvents. This is to be accomplished by careful and rational control of the enzymic microenvironment and encompasses both enzymes suspended and dissolved in organic solvents. Our previous efforts have been directed toward developing a mechanistic and molecular-level understanding of enzyme activity and activation in organic solvents by studying enzyme structure, function, and dynamics in such media. This work builds on preliminary findings and attempts to develop general strategies for preparing highly active and stable enzymes in both heterogeneous and homogeneous nonaqueous environments. The expected results of this proposed study are enhanced functional utility of a wide range of enzyme catalysts in both polar and nonpolar organic solvents as well as improved understanding of the molecular events that govern enzyme function and stability in the extreme, yet commercially critical environment of organic solvents. The specific objectives of this effort are as follows:

1. Optimize the activity and stability of enzymes in organic solvents by lyophilization with salts and other excipients. Such "biocatalyst engineering" has been shown to improve enzyme acticity in organic solvents by several orders of magnitude. The rationale for this work is that further improvement should be possible through the optimal selection of excipient and control of enzyme hydration. These parameters will be optimized with the goal of producing biocatalysts with the highest possible activity and stability in organic solvents.

2. Maximize the activity of enzymes solubilized in organic solvents via surfactant ion-pairing, and expand the range of enzymes and solvents amenable to this methodology.

Several model enzymes will be examined including the subtilisns, Carlsberg and BPN', bovine α-chymotrypsin, thermolysin, and soybean hull peroxidase. Some will be provided through our industrial partnerships (Genencor and Diversa, Inc.). Cross-linked versions of some of these enzymes from Altus Biologics will also be evaluated.

1996-1997 Accomplishments:

New start in FY 1997.

1998 Planned Activities:

This project will primarily involve several model enzymes: subtilisins Carlsberg and BPN', bovine α-chymotrypsin, thermolysin, and soybean hull peroxidase (obtained from Enzymol International). We will also have numerous protein engineered variants of BPN' made available to us in our collaboration with Genencor, which will be used to probe the mechanisms of enzyme activation in dehydrated media. We will also have access to CLECs (subtilisin, thermolysin as well as others) through our collaboration with Altus Biologics.

Specific tasks include:

1. Optimization of enzyme activation by salts
2. Effect of organic solvents on salt activation
3. Salt activation of lyophilized CLECs
4. Stability of activated enzyme preparations

which will be performed during efforts described in the following topical research areas.

Engineering the catalyst microenvironment

a) *Optimization of enzyme activation by salts.* We will examine the catalytic activity and catalytic rate enhancement of the aforementioned enzymes lyophilized in the presence of a variety of salts. The following properties will be investigated in this work:

1) Effect of salts with different chaotropic/kosmotropic properties;
2) Effect of the pH of the lyophilizing solution;
3) Effect of the water activity, a_w, of the enzyme-salt powder;
4) Effect of active-site polarity on salt-induced activation via use of polar and nonpolar site-directed mutants of subtilisin BPN' provided by Genencor.

b) *Effect of organic solvents on salt activation.* We will examine the magnitude of the activation effect by different salts as a function of the polarity and hydrophobicity of the organic solvent. Because salt activation is believed to stem from a combination of enzyme hydration and active-site polarity, we anticipate that the magnitude of activation will be highest in nonpolar solvents. Nevertheless, substantial activation can be achieved in polar organic solvents upon use of the appropriate salt as protein-water interactions may become sufficiently strong to enable the protein to retain critical hydration in the presence of highly water-solubilizing polar solvents.

c) *Salt activation of lyophilized CLECs* - We will suspend CLECs in aqueous solutions containing different salts and lyophilize the preparations. The catalytic activity and stability of these salt-added lyophilized CLECs will be compared to controls prepared without added salts as well as to native CLECs dried by washing with a water-stripping solvent. We anticipate that the same mechanisms that give rise to enzyme activation due to the presence of the salt may result in activation of CLECs in nearly anhydrous organic media as compared to the controls in the absence of salt.

Stability of activated enzyme preparations

Catalytic activity will be determined in organic solvents as a function of temperature, and catalyst half-lives will be measured at multiple temperatures. Salt-activated enzyme preparations, CLECs and salt-added lyophilized CLECs will be employed in this phase of the work.

Annual Technical Summary Report:

None. This is a new start which did not begin until late in FY 1997 or early FY 1998.

7. Project Title Biocatalysis by Design
Principal Investigator(s) William A. Goddard, III and Siddarth DasGupta
Project Site California Institute of Technology Materials Simulation Center

Description:

This project is directed to the development of tools for atomic scale modeling and simulation of biological systems, and the prediction of critical parameters for experimental validation of the models. The project will also assess the critical impediments to the widespread application of biotechnology in industry, and will examine possible approaches to overcoming the barriers, using the tools developed at the Materials and Molecular Simulations Center (MMSC). The MMSC was established as a place where academic, government, and industrial researchers can work together on theoretical studies related to catalyst design and structure using molecular dynamics, quantum mechanics, and computer graphics capabilities. The MMSC is a means of technology transfer between the researcher and the end user. The needs of industry therefore direct the project goals and orientation.

Methods and techniques to permit the prediction of the structure, and hence function, of proteins are under investigation. Project goals include the design of proteins that recognize specific DNA sequences. These proteins would be used to develop regulatory proteins for controlling biocatalysts. In particular, the pseudo-spectral generalized valence bond (GVB) techniques will be used. The current chemical mechanical calculations, the first step in molecular mechanics simulations, are extremely CPU intensive. Consequently, these calculations are restricted to molecular systems that are small (10-20 heavy atoms). However, most industrial applications of interest are much larger in number of atoms of interest and are thus much more computer intensive (in terms of method of calculation) than is practically feasible. Recent developments in the pseudo-spectral GVB techniques, which use ideas originally developed to solve fluid dynamics simulations, have made it more feasible to calculate systems of 100 atoms, which would help substantially *ab initio* calculations on crystals and systems with periodic boundary conditions. Massively parallel computing capabilities will be necessary to accomplish this goal and to expedite these calculations into reasonable cost/time operations.

1996-1997 Accomplishments:

Highlights of the topical areas that were evaluated and investigated throughout the past year are either briefly summarized or references and outcomes are noted where specific output resulted in publishable information or information useful to industry.

Process Simulations – Mario Blanco. We have continued our industrial collaborations in the areas of process simulation. These include:

Scale Control : *Chevron Petroleum Technology Company.* Work was designed to develop a new scale dissolver formulation.
Polymer Compatibilizers and Silane Coupling Agents: *Owens Coming*. Work effort involved developing new software for compatibility predictions that save laboratory efforts.
Wear Inhibitors: *Chevron Research Technology Center.* Work involves developing a self-assembling monolayer model for metal wear inhibition by dithiophosphates.

Gasoline Reformer Catalysts: *Exxon Research.* Work included examining new catalysts for gasoline reformulation.
Water-Resistant Adhesives: *Avery Dennison.* Identified a mechanism for outstanding water resistance in new acrylic modified adhesives

Mechanism of Uncoating of the Rhinovirus – Nagarajan Vaidehi. We have published this work in the *Proceedings of the National Academy of Sciences* USA, **94**, pp 2466-2471 (1997).

Domain Motions in Phospho Glycerate Kinase (PGK) – Nagarajan Vaidehi The movie of the dynamics can be found at http://www.wag.caltech.edu

A Second Generation Force Field for Nucleic Acid Simulations – Ken A. Brameld & Siddharth Dasgupta. We have published the work. Brameld, K.A., Dasgupta, S., Goddard, W.A. III. "Distance Dependent Hydrogen Bond Potentials for Nucleic Acid Base Pairs from *ab initio* Quantum Mechanical Calculations (LMP2/cc-pVTZ)." *J. Phys. Chem B.*, **101**, 4851-4859 (1997).

Investigation of the Mechanism of Family 18 and Family 19 Chitinases – Ken. A. Brameld, William A. Goddard, III. Three papers were submitted for publication: a) Brameld, K. A.; Shrader, W. D.; Imperiali, B.; Goddard, W. A., III. "Substrate Assistance in the Mechanism of Family 18 Chitinases: Theoretical Studies of Potential Intermediates and Inhibitors." *J. Mol. Biol.*; b) Brameld, K. A.; Goddard, W. A., III. "The D-Boat Substrate Distortion Mechanism for Family 18 Chitinases." *J. Am. Chem. Soc.*; and c) Brameld, K. A.; Goddard, W. A., III. "The Role of Enzyme Distortion in the Single-Displacement Mechanism of Family 19 Chitinases." *Proc. Natl. Acad. Sci. USA.*

Resolving the Levinthal Paradox – Practical Protein Folding – Derek A. Debe. A paper was submitted for publication in *Nature*, D. A. Debe, M. J. Carlson, and W. A. Goddard III, Resolving the Levinthal Paradox: Practical Protein Folding.

The Early Stages of the Protein Folding Process – Ruth Ann Bertsch, Nagarajan Vaidehi, Yao-Chun Peng, Sunney I. Chan, & William A. Goddard, III. We have predicted that the folding time scales in our simulations would be in the nanoseconds scale for 100% folding yield. This paper is being communicated to *Protein Science*.

Sidechain Rotamer Predictions in Folded Proteins – Matt J. Carlson, Derek Debe, William A. Goddard III. Our rotamer prediction method compares favorably with other methods in the literature. The strength of the method lies in its good performance with a smaller rotamer set, about half the number of rotamers of other current methods. This is achieved with our custom force field used in the energy evaluation. Thus, our method has the capability to operate on a large number of candidate protein folds to better fit within our hierarchical framework.

1998 Planned Activities:

This project is no longer funded and will conclude in FY 1998

Annual Technical Summary Report:

The following individual reports represent work supported in part by the Department of Energy and leveraged with industry and other federal agencies. It is listed by specific project with the names of those individuals specifically responsible for the work.

Mechanism of Uncoating of the Rhinovirus – Nagarajan Vaidehi

We have used our new Massively Parallel molecular dynamics simulation code (MPSim, which can treat very large systems – 100,000 to 1,000,000 atoms) to study the rhinovirus (HRV) responsible for the common cold. HRV has a protein shell (about 300A diameter) with about 500,000 atoms surrounding the RNA in its core. The pathway for viruses to multiply involves binding to a receptor on the cell surface and then exporting of this virus, but there are over 100 serotypes, making it difficult to design drugs or a vaccine for them. Experimental antigen studies indicate that the RNA comes out through the channel at the center of a pentamer. It has been speculated that drugs effective against HRV-1A constrict the pentamer channel so that the RNA cannot be exported, while drugs against HRV-14 recognize the cell receptor binding site.

To test this hypothesis we examined the energy cost to open the pentamer channel of the virus sufficiently wide for RNA to leave. We find that for HRV-1A with normal cofactor or sucrose the pentamer channel can open to 20 Å without significant energy costs, whereas the three best drugs from Winthrop all restrict the opening to 12 Å, too small for the RNA to be exported. Indeed WIN-52035, a drug which requires much larger concentrations to be effective can open to 15 Å. Based on these results we designed three new drugs and tested them using this model. We find that these new drugs are as effective as the best known one for HRV-1A. On the other hand we find that the WIN drugs for HRV-14 do not lead to pentamer channel stiffening and in the case of HRV-14 the drug action is through preventing binding to the cell and not through uncoating. This is accordance with the experimental speculation for HRV-14. We have published this work in the *Proceedings of the National Academy of Sciences USA*, **94**, pp 2466-2471 (1997).

Domain Motions in Phospho Glycerate Kinase (PGK) - Nagarajan Vaidehi

The enzyme PGK catalyzes a key direct phosphoryl transfer step in the glycolytic pathway. Under physiological conditions, PGK facilitates the transfer of phosphoryl group from 1,3-biphospho D-glycerate to MgADP, thus forming MgATP and 3-phospho-D-glycerate (3-PG). PGK consists of two major domains, namely the N-domain and the C-domain. Each globular domain has about 1400 atoms in an α/β, twisted open sheet structure composed of a single sheet of six parallel β-strands sandwiched between four helices. Crystal structures of PGK from various species show that the substrates bind to the opposite domains on PGK. If so, how does PGK bring the domains together to facilitate the reaction?

Using hierarchical Newton Euler Inverse Mass Operatior (NEIMO) algorithm, we performed MD simulations with the two domains as rigid bodies and the torsional degrees of freedom in the interdomain region. Counterions Na+ and Cl- were used to account for solvation of the polar side chain aminoacids. These simulations were performed on the Origin 2000 supercomputer using the MPSim code. We find large domain motions occurring and the open PGK structure goes from open to close structure and swings back and forth around the closed form.

Further analysis of the domain motions in terms of Euler angles show that the axis of rotation during the dynamics coincided with the axis of rotation in going from the experimental open form of PGK to the closed form of PGK. It should be emphasized that this domain motion was not observed with the conventional Cartesian dynamics. The dynamics structures are close to the closed crystal structure although not exactly the same in the interdomain region. This is because the crystal structures for the open and closed form have sequence insertions in the interdomain region. However, the conserved regions agree within 1.5 Å with the closed crystal structure. These results are exciting.

We also performed NEIMO big domain dynamics at 300K on the closed form of the crystal structure. It is extremely interesting to note that the dynamics opens up the domains initially within 25ps but goes back to the closed form (exactly) after 25ps and stays there. This shows that when substrates are bound the more stable conformation is the closed form than the open form. We also see an energy hump when the structure opens up and goes back to equilibrium subsequently. The results of these simulations are under preparation for publication. The movie of the dynamics can be found at http://www.wag.caltech.edu

A Second Generation Force Field for Nucleic Acid Simulations - Ken A. Brameld and Siddharth Dasgupta

We have been developing a new force field (FF) for molecular mechanics and dynamics simulations of nucleic acids, parameterized from quantum mechanical calculations on small model systems. The substructures critical for the conformation of nucleic acids may be broken down into three subunits; the purine and pyrimidine base pairs, a ribose or deoxyribose sugar and a phosphodiester backbone. We have now completed parameterization for all subunits and will begin carrying out validation studies.

Base Pair Hydrogen Bonding

Hydrogen bonding is one of the key nonbonded interactions found in many naturally occurring systems to maintain structure and specificity. The essential recognition and specificity of nucleic acid base pairing, which is responsible for the transfer of genetic information, stems from the specific formation of hydrogen bonds between Watson-Crick base pairs. We have carried out quantum mechanical calculations (Hartree-Fock) to optimize the Watson-Crick A-T and G-C base pairs and the Hoogsteen A-T pair.

Initial calculations with a 6-31G** basis set yielded hydrogen bond lengths that were 0.15 Å too long as compared to X-ray crystal structures. Using the larger, more diffuse cc-pVTZ basis set, the average error was reduced to 0.05 Å. To determine the complexation energy, the energy difference between the free bases and the base paired complex must be corrected for basis set superposition errors (BSSE).

Hartree-Fock (HF) calculations do not include electron correlation, therefore will not accurately describe the dispersion or "Van der Waals" energies. Hydrogen bonding is primarily an electrostatic interaction. However, we have found that correlation is critical for determining accurate complexation energies. The HF/cc-pVTZ calculations greatly underestimate the A-T Watson-Crick (-9.2 kcal/mol) and A-T Hoogsteen base pair energies (-9.4 kcal/mol) when compared to the experimentally determined energy of -13.0 kcal/mol. Using the HF/cc-pVTZ optimized geometry, we calculated the LMP2/cc-pVTZ(-f) energies and found correlation affects to increase the A-T Hoogsteen hydrogen bond energy by 2.3 kcal/mol. The Watson-Crick G-C pair only increased by 0.9 kcal/mol.

After correcting the quantum mechanical results for thermal and zero-point energies, we find good agreement between the theoretical energy of -21.4 kcal/mol and the experimental energy of -21.0 kcal/mol for the GC base pair. In contrast our calculated energy of -10.8 kcal/mol still falls short of the experimental energy, -13.0 kcal/mol.

In addition to examining the minimum geometry and energy of the base pairs, we have also studied the potential as a function of distance. These calculations were carried out including dispersion (LMP2/cc-pVTZ(-f)). We have found that a simple repulsive Morse potential, in

conjunction with a full electrostatic description, may be used to accurately describe the entire potential surface. The attractive portion of the potential at large distances is dominated by electrostatics, which eventually give way to the repulsive Morse potential.

Sugar Conformations

Conformational energies for ribose, deoxyribose and C3'-fluroribose have been examined using *ab initio* quantum mechanics (HF/6-31G**//LMP2/cc-pVTZ(-f)). The energy differences between the low energy conformational pucker states of these furanose sugars is important in the simulation of nucleic acids. Sugar pucker conformations are critical in defining global helical properties of double-stranded DNA and impart local structural variations in the phosophodiester backbone.

Within the 5-membered furanose ring, the lowest energy pucker conformations will have one or two atoms which lie out of the plane defined by the remaining three or four atoms. The completely planar conformation is oftentimes 3-5 kcal/mol higher in energy. Interconversion between these various pucker states occurs along a pseudorotation pathway, during which the planar conformation does not occur. For the ribose and deoxyribose sugars, the two local minima defined as 2'-endo and 3'-endo are of primary interest, as is the barrier height between these conformations.

Due to size limitations in the number of atoms for which quantum mechanical calculations are practical, a complete base + sugar + phosphate cannot be investigated. Instead, smaller cluster models which have the intrinsic properties of the complete system must be used. In the case of the ribose and deoxyribose calculations, we examined many substituted amines at the C1' position as replacements of the full purine or pyrimidine base. We found a C1' pyrrole substitution to be a good compromise between accuracy and feasibility of the calculations. Along with a methyl substitution at C4', this constituted the cluster model used for all subsequent energy calculations.

For deoxyribose, we find the 2'-endo conformation is the global minimum with the 3'-endo conformation 0.55 kcal/mol higher in energy. The barrier height between interconversion of these two conformations is 1.70 kcal/mol. There are few experimental data with which to compare these results. Based on NMR data, the C3'-endo conformation is believed to be 0.5 kcal/mol higher in energy than C2'-endo and the barrier height to be between 1.0-3.0 kcal/mol. Thus our calculations are in agreement with the limited experimental data available.

The pseudorotation potential for ribose is complicated by the presence of two hydroxyls (C2' and C3'). It is known from X-ray structures of RNA that the ribose sugar prefers a C3'-endo pucker. Our calculations indicated a global minimum at C2'-endo, similar to that seen for the deoxyribose sugar. Examination of this optimum structure revealed the formation of a hydrogen bond between the C3' and C2' hydroxyls. This hydrogen bond cannot form in the RNA polymer as the O3' hydrogen of our cluster model is a phosphate in the full system. Substitution of a hydroxyl with a fluoride (which maintains the electronic features of a O-P moiety) alleviated this problem. For the C3'-fluoro ribose, we found the C3'-endo conformation to be the global minimum. The C2'-conformation is 1.7 kcal/mol higher in energy and the barrier is between the two conformations is 2.4 kcal/mol.

A force field has been parameterized to reproduce the pseudorotation potential energy curves for the ribose and deoxyribose sugars. In conjunction with parameterization of the phosophodiester backbone using a dimethylphosphate cluster model, the MSCFF (MSC Force

Field) is now complete for nucleic acids. Validation studies of known systems will be carried out to ensure the accuracy of this new FF. Among the systems to be studied include small high resolution crystals such as GpC and ApU and larger DNA structures such as a B-DNA decamer.

Investigation of the Mechanism of Family 18 and Family 19 Chitinases - Ken. A. Brameld and William A. Goddard, III.

Summary

Chitinases are critical for the normal life cycles of insects and fungi by controlling molting and growth, yet are also found in plants as a defense against fungal pathogens. The mechanism of the enzymatic hydrolysis of chitin, a β-(1,4)-linked N-acetyl-glucosamine (NAG) polysaccharide, is not well understood; however analogies with lysozymes suggest a role for an oxocarbenium ion intermediate. We carried out *ab initio* quantum mechanical calculations at the HF/6-31G** level to optimize the geometry and energy for three putative transition states during the enzymatic hydrolysis of chitin. These calculations indicate *that the stable oxocarbenium intermediate is a dioxocarbenium species* in which the acetamido carbonyl forms a covalent bond to C1' of N-acetyl-glycosamine. The resulting intermediate has a delocalized positive charge on the acetamido and the six-membered glucose ring retains a predominately chair conformation.

Using molecular mechanics simulations, several conformations of a hexaNAG substrate bound to the active site of chitinase A have been examined. However, only one conformation is both stable during dynamics simulations and is in an orientation suitable for the first step of acid catalyzed hydrolysis: protonation of the linking anomeric oxygen between GlcNAc residues D and E. For this structure, GlcNAc residue at binding subsite D is distorted to a boat conformation. QM calculations reveal that protonation of a boat conformation leads directly to glycosidic bond cleavage and an oxazoline ion intermediate with no barrier.

While the family 18 chitinases use substrate distortion and anchimeric assistance to achieve the enzymatic hydrolysis of chitin, the family 19 chitinases operate under a completely different hydrolysis mechanism. Using molecular dynamics simulations, we have examined the binding of a hexaNAG substrate and two potential hydrolysis intermediates (an oxazoline ion and an oxocarbenium ion) to a family 19 barley chitinase. We find the hexaNAG substrate binds with all sugars in a chair conformation, unlike the family 18 chitinase which causes substrate distortion. Glu 67 is in a position to protonate the anomeric oxygen linking sugar residues D and E while Asn 199 serves to hydrogen bond with the C2' *N*-acetyl group of sugar D, thus preventing the formation of an oxazoline ion intermediate. In addition, Glu 89 is part of a flexible loop region allowing a *conformational change to occur within the active site* to bring the oxocarbenium ion intermediate and Glu 89 closer by 4-5 Å. A hydrolysis product with inversion of the anomeric configuration occurs due to nucleophilic attack by a water molecule which is coordinated by Glu 89 and Ser 120.

Family 18 Hydrolysis Intermediates

We have explored the possibility of an oxocarbenium ion being internally stabilized by the acetamido carbonyl of a NAG residue, by carrying out *ab initio* calculations on a C1'-deoxy-N-acetyl-glucosamine ion. We examined three conformations with various orientations of this internal carbonyl by constraining the C--N--C2'--C3' torsion and optimizing the rest of the structure. The first case allows any conformation and yields the global minimum which is a dioxocarbenium species. The carbonyl rotates towards C1' forming a covalent bond of length

1.49 Å. The two additional structures have the carbonyl constrained in either an extended conformation or allowed to coordinate to the positive O1', but without forming a covalent bond. In both these cases the glucose adopts a half-chair geometry due to sp^2 C1' center. This contrasts with the dioxocarbenium ion which can maintain a mildly distorted full chair conformation.

The calculations discussed above did not include the solvent present in the normal aqueous environment. In order to estimate the effects of solvent polarization we calculated self-consistently the wave function while including a Poisson-Boltzmann continuum description of the solvent. We find a similar relative stability of the three structures with the dioxocarbenium ion lowest in energy. Without solvation, the coordinated acetamido structure has an energy of 16.1 kcal/mol, and the extended acetamido conformation has an energy of 29.8 kcal/mol, relative to the oxocarbenium structure. The inclusion solvation preferentially stabilizes the extended conformation (due to the greater solvent accessible surface area about the carbonyl); however the energetic ordering of these conformations remains unchanged with the following relative energies: 0.0, 16.9 and 21.7 kcal/mol.

Allosamidin and related derivatives have been shown to be chitinase inhibitors and are believed to act as transition state analogs. We also calculated the optimum geometry of allosamizoline, the aglycon portion of allosamidin, so as to compare this inhibitor to the dioxocarbenium species. There is a striking match between the two structures, particularly in the charged region of the acetamido. The RMS coordinate difference between equivalent non-hydrogen atoms is 0.21 Å. We also calculated the electrostatic potential surfaces and find a very good correlation between the inhibitor and proposed transition state. This strongly supports our theoretical findings (but does not rule out alternative transition state intermediates).

Family 18 Substrate Distortion

Using molecular dynamics simulations, we examined the plausible conformations for a hexaNAG substrate bound to the active site of chitinase A. We found that:

1. The hydrolysis mechanism of chitinase A (a family 18 chitinase from *Serratia marcescens*) involves substrate distortion. We find that only one conformation is stable during dynamics simulations.
2. The first step of acid catalyzed hydrolysis (protonation of the linking anomeric oxygen between GlcNAc residues D and E) requires a *boat* conformation for the GlcNAc residue at binding subsite D.
3. *Ab initio* quantum mechanical (QM) calculations (HF/6-31G**) predict that protonation of a GlcNAc in a boat conformation leads to spontaneous anomeric bond cleavage to yield an oxazoline ion intermediate.

We also studied several conformations of two possible hydrolysis intermediates -- the oxocarbenium ion and the oxazoline ion. Only the oxazoline ion orients in the enzyme active site so as to allow stereo-selective attack by water. This leads to retention of configuration in the anomeric product as observed experimentally.

It is possible that all family 18 chitinases share a common mechanism. Hence, we suspect that distortion of the D-subsite of the substrate into a boat form is required for any glycosyl hydrolase which has only one acidic residue in the active site.

Family 19 Conformational Change

In order to investigate the detailed mechanism of family 19 chitinases, we have carried out molecular dynamics (MD) simulations on the barley seed endochitinase complexed with a hexaNAG substrate and two possible intermediates (triNAG-oxocarbenium ion and triNAG-oxazoline). We conclude:

1. The hexaNAG substrate binds with all sugars in a chair conformation such as to favor protonation by Glu 67 of the anomeric oxygen linking sugar residues D and E.
2. Asn 199 serves to hydrogen bond with the C2' *N*-acetyl group of sugar D, preventing the formation of an oxazoline ion intermediate.
3. Glu 89 is situated in a flexible loop region allowing a *conformational change to occur within the active site* to bring the oxocarbenium ion intermediate 4-5 Å closer to Glu 89.
4. Glu 89 and Ser 120 coordinate with a water molecule which may be activated for nucleophilic attack; this would yield a product with an inverted anomeric configuration (as observed).
5. Inhibitors designed to be transition state analogs of the more planar oxocarbenium ion are predicted to be selective against family 19 chitinases over family 18 chitinases.

Papers submitted for publication

Brameld, K. A.; Shrader, W. D.; Imperiali, B.; Goddard, W. A., III. "Substrate Assistance in the Mechanism of Family 18 Chitinases: Theoretical Studies of Potential Intermediates and Inhibitors." *J. Mol. Biol.* Submitted for publication.

Brameld, K. A.; Goddard, W. A., III. "The D-Boat Substrate Distortion Mechanism for Family 18 Chitinases." *J. Am. Chem. Soc.* Submitted for publication.

Brameld, K. A.; Goddard, W. A., III. "The Role of Enzyme Distortion in the Single-Displacement Mechanism of Family 19 Chitinases." *Proc. Natl. Acad. Sci. USA.* Submitted for publication.

Resolving the Levinthal Paradox – Practical Protein Folding - Derek A. Debe

The protein folding problem, predicting the folded conformation of a protein from amino acid sequence, is perhaps the most critical and certainly most intriguing unsolved problem in structural biology. At the Materials and Process Simulation Center here at Caltech, we have developed a computationally feasible, hierarchical approach to solving this previously intractable multiple-minima problem. The basis of this approach is an algorithm which exhaustively samples the complete set of conformational topologies for an n-residue protein. This coarse grained sampling allows a vast number of conformation states to be eliminated from consideration as the global free energy minimum.

The major difficulty in protein structure prediction is sampling the unique native fold despite a seemingly intractable number of conformational alternatives (e.g. $3^{100} \approx 10^{47}$ conformations for 100 amino acids). In association with Prof. Goddard at the Materials and Process Simulation Center at Caltech, I have developed the Generic Protein (GP) Direct Monte Carlo method for generating large ensembles of non-overlapping polypeptide chains. Using a massively parallel implementation of this method at the National Center for Supercomputing Applications at Urbana-Champaign, Illinois, we have demonstrated that the GP direct Monte-Carlo approach is capable of exhaustively sampling all possible fold topologies for polypeptides up to 100 amino acids in length. We show that the size of this set is $\sim 3\times10^{7}$, substantially smaller than 10^{47}. This result has significant implications with regard to how a protein samples its unique native

fold *in vivo* in time scales on the order of milliseconds (resolving the Levinthal Paradox). Furthermore, the algorithm allows for the crucial reduction in conformation space required for successful *ab initio* prediction of 3-dimensional structure from primary sequence.

The GP method is the foundation of a hierarchical approach (Figure 1) to the protein structure prediction problem. The GP method successfully samples all possible fold topologies for a peptide chain of length n, and consequently a GP contains a near-native 3-dimensional structure for all possible primary amino acid sequences. After generation of the GP ensemble, a recognition filter designed to select promising 3-dimensional folds for a specific amino acid sequence may be used in order to reduce the ensemble size by several orders of magnitude. Thus the recognition step of the prediction hierarchy yields a manageable ensemble of structures which contains a near-native structure for the specific amino acid sequence being considered. When coupled with successful recognition approaches, the GP method reduces the *ab initio* folding problem to the equivalent of the homology modeling problem, the amenable task of refining a structure with the correct global fold to comply with sequence specific energetic interactions.

In the final hierarchical step, full atom models of the remaining candidate structures are created, allowing for further refinement and recognition of the final structure using full atom/full solvation molecular dynamics procedures and other accepted techniques for homology model refinement.

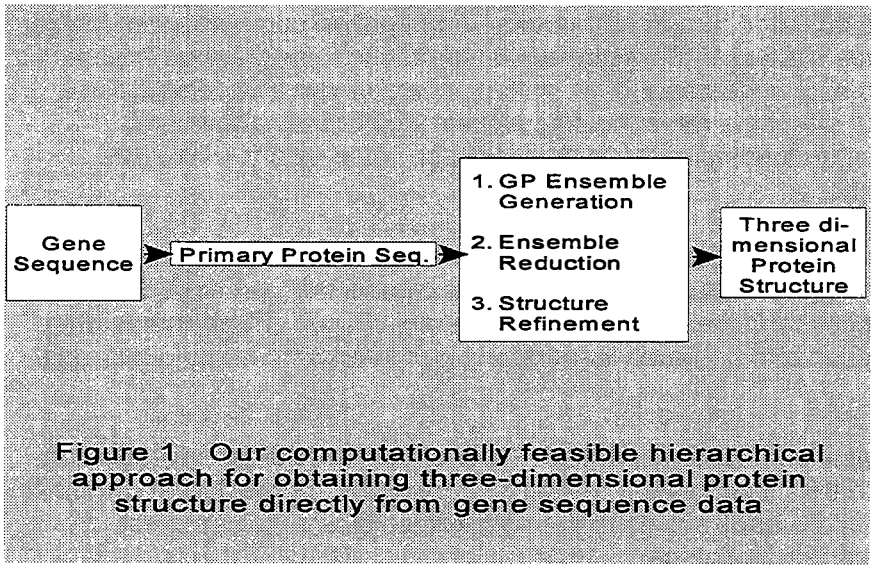


Figure 1 Our computationally feasible hierarchical approach for obtaining three-dimensional protein structure directly from gene sequence data

We have demonstrated the success of the GP algorithm and submitted our results for publication to *Nature*. Current research concerns implementing variations of the algorithm which use a small number of known distance constraints (such as the location of disulfide bonds or long range NOE's) to drastically decrease the size of the original ensemble which must be generated in order to generate near-native structures. Other research focuses on developing and refining the recognition techniques used to cleave the size of the original GP ensemble.

Paper submitted for publication:

D. A. Debe, M. J. Carlson, and W. A. Goddard III, Resolving the Levinthal Paradox: Practical Protein Folding, *Nature.*

The Early Stages of the Protein Folding Process - Ruth Ann Bertsch,* Nagarajan Vaidehi,† Yao-Chun Peng,† Sunney I. Chan,* & William A. Goddard, III†

Using the NEIMO algorithm, molecular dynamics can fold a 20-mer of polyalanine, starting from the extended, non-helical state, into an α-helix. Under specific simulation conditions (e.g. 450°K, 1e(r) dielectric constant, and a 15 Å cutoff radius for nonbonded interactions), the polyalanine has a 81.8% probability (+/- 6.2% with 95% confidence), depending on the initial momenta, of forming a helix in 500 ps. Nonhelical hydrogen bonds may be most responsible for preventing the failures from forming a helix in 500 ps. The calculated helix propensities (in the present study) are Ala > Leu > Gln > Gly. This order is the same as that found experimentally. This material is being communicated to *Protein Science*.

Understanding the process of protein folding is important today

Understanding the kinetic process of protein folding is a significant industrial and biological problem today. Theoretically a folding proteinaceous polymer will eventually adopt the conformation of its thermodynamic minimum. However, the free energy of stabilization of the native state of a protein is very small, and many local, thermodynamic minima exist. We might expect that the protein could become trapped in one of these minima, preventing it from folding on a reasonable time scale.

When a protein is trapped in a local minimum, it becomes susceptible to irreversible aggregation, lowering the yield of folded, functional protein. Aggregation competes with folding early in the folding process; thus, studies of the early events of protein folding are particularly important. Aggregation is also sequence-dependent. In fact, some amino acid residues in a protein sequence may primarily affect the rate of folding without significantly affecting its function or stability. A better understanding of the folding process could lead to enhanced control over protein folding yields.

Unfortunately, the early events of protein folding are difficult to study experimentally because they are so fast. Secondary structure frequently forms within microseconds, yet experiments to study proteins folding at submillisecond time scales are technically difficult. Furthermore, they can only supply limited information and are rarely general enough to apply to many other water-soluble proteins.

Thus, computer simulations could contribute much to the field of protein folding. Investigating the behavior of simple folding motifs, such as the α-helix, is a tractable approach that computer simulations can use to elucidate the early events of protein folding. One of the simplest polypeptides which forms a helix is polyalanine.

Methods

Twenty-mers of polyalanine were built using the Peptide Builder of BIOGRAF. The terminal residues were neutral, except where otherwise noted.

Molecular dynamics simulations were conducted using the NEIMO-Hoover method. The equations of motion solved in the NEIMO-Hoover method are given by

$$M(\theta) + C(\theta, \theta) - \xi(\theta) = T(\theta)$$

The solutions have been shown to be linear in the number of degrees of freedom. NEIMO reduces the degrees of freedom of the polymer, enabling faster conformational searches. NEIMO freezes the bond lengths and angles until only three torsional angles, f, y, and w, are free to move.

All simulations used the united atom AMBER force field and a timestep of 10 fs. Simulations were conducted for at least 500 ps. Except where noted, the dielectric constant was 1e(r) and distance dependent. Most simulations used cutoff radii of 15 Å for nonbonded interactions. In other words, most simulations ignored the nonbonded interactions outside a sphere of a specific radius.

Results:

The kinetics of α-helix formation in polyalanine and polyglycine eicosamers (20-mers) have been examined using torsional coordinate molecular dynamics (MD). Of 155 simulation experiments on extended $(Ala)_{20}$ carried out for 0.5ns each, 129 (83%) formed a persistent α-helix. In contrast, the extended state of $(Gly)_{20}$ only formed a right-handed α-helix in two of the 20 MD experiments (10%), and these helices were not as long or as persistent as those of polyalanine.

These simulations show helix formation to be a competition between the rates of forming local hydrogen (i.e. hydrogen bonds between any residue *i, i+2 , i+3, i+4 or i+5th* neighbor and forming nonlocal hydrogen bonds (HBs) between residues widely separated in sequence.

Local HBs grow rapidly into an α-helix; but, nonlocal HBs usually retard helix formation by trapping the polymer in irregular, "balled-up" structures. Most trajectories formed some nonlocal HBs, sometimes as many as eight. But, for $(Ala)_{20}$, most of these eventually rearranged to form local HBs that lead to α-helices. A simple kinetic model describes the rate of converting nonlocal HBs into α-helices.

Torsional coordinate MD speeds up folding by eliminating bond and angle degrees of freedom and reducing dynamical friction. Thus observed 210 ps half-life for helix formation is likely to be a lower bound on the real rate. However, we believe the sequential steps observed here mirror those of real systems.

Mechanisms of Helix Formation:

Example simulation of a helix folding. Figure 1 depicts a simulation of a helix folding at 400°K with a 15A cutoff for nonbonded interactions. Figure 1 exemplifies many simulations.

Local minima. Within the first 5 ps, the polyalanine in all of the simulations relaxed to a configuration where dihedral angles lay in a well around $\phi >> -80°$ and $\iota » +70°$. Figure 2 illustrates this -80°, +70° well and the α-helical well at $\phi >> -60°$ and $\iota » -40°$. Each graph in Figure 2 is the trajectory of the ϕ/ι angles one residue took during a simulation at 450°K, 15A cutoff radius. The -80°, +70° orientation is not highly favored by proteins folded at equilibrium. (See Figure 3.) However, the orientation allows strong hydrogen bonding to form between residues separated by one intervening residue (*i, i+2* hydrogen bonds). These minima appeared in all simulations tested, from temperatures ranging from 300°K to 500°K, whether nonbond interactions were calculated explicitly or only out to a cutoff radius.

Figure 1. The trajectory of folding polyalanine during the NEIMO-Hoover simulation at 450° K with a spline cutoff at 15 Å and a distance-dependent dielectric of 1ε r. In this simulation, the helix first forms at residues 8 to 11. Then the amino terminal half of the polymer forms a helix. Finally, residues 14 to 18 folded starting from the center and propagating up progressively.

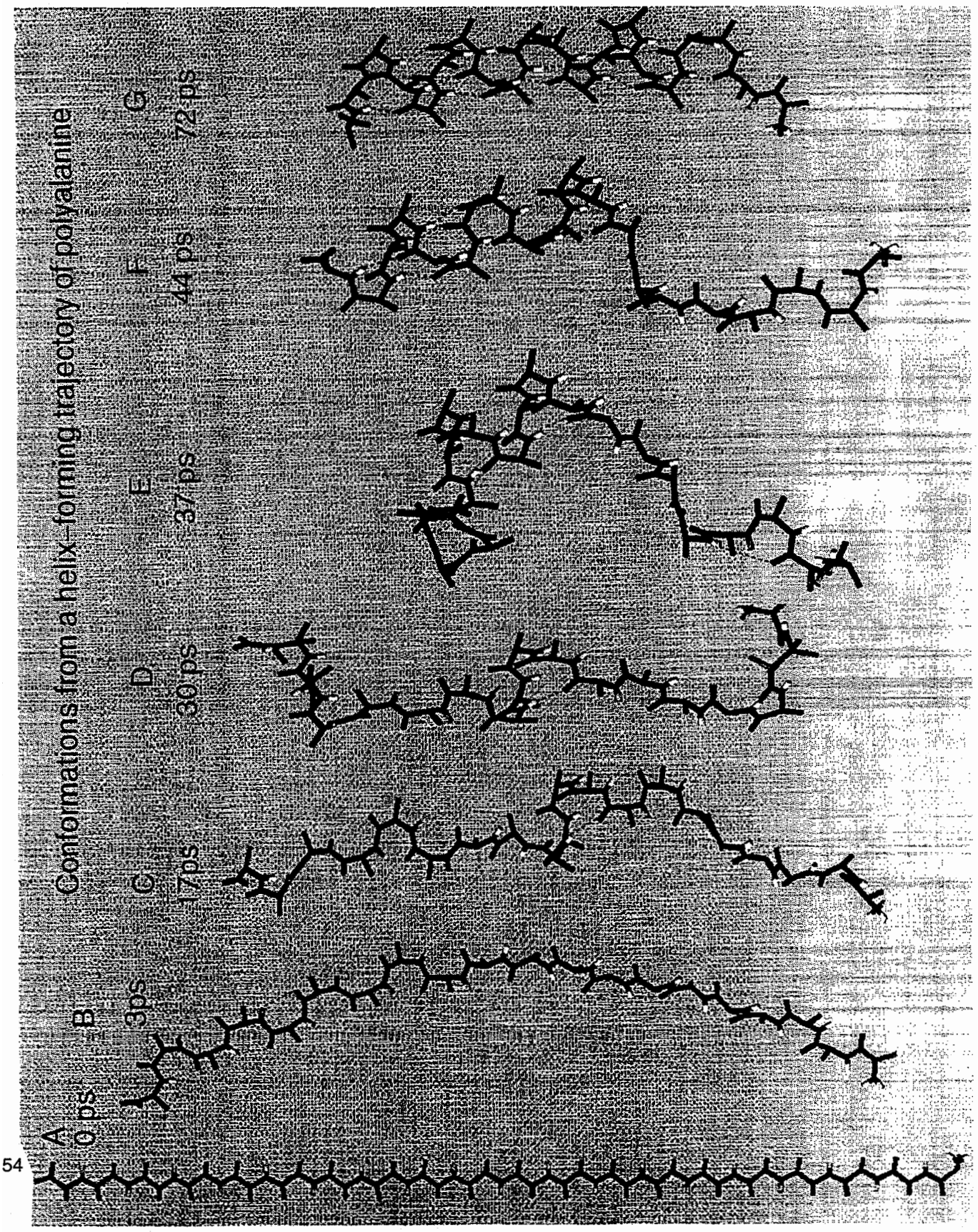


Figure 2. The trajectory that the dihedral angles ϕ and ı of residues 9-11 and 17 took during the NEIMO-Hoover dynamics simulation at 450°K using a nonbond cutoff of 15 Å. These residues clearly illustrate the existence of two energetic minima separated by a large barrier. The well at ~-60°, -40° characterizes α-helices. The well at ~-80°, +70° characterizes structures with strong, *i, i+2* hydrogen bonds. The polyalanine in all of the simulations immediately relaxed into this latter well from the fully extended state.

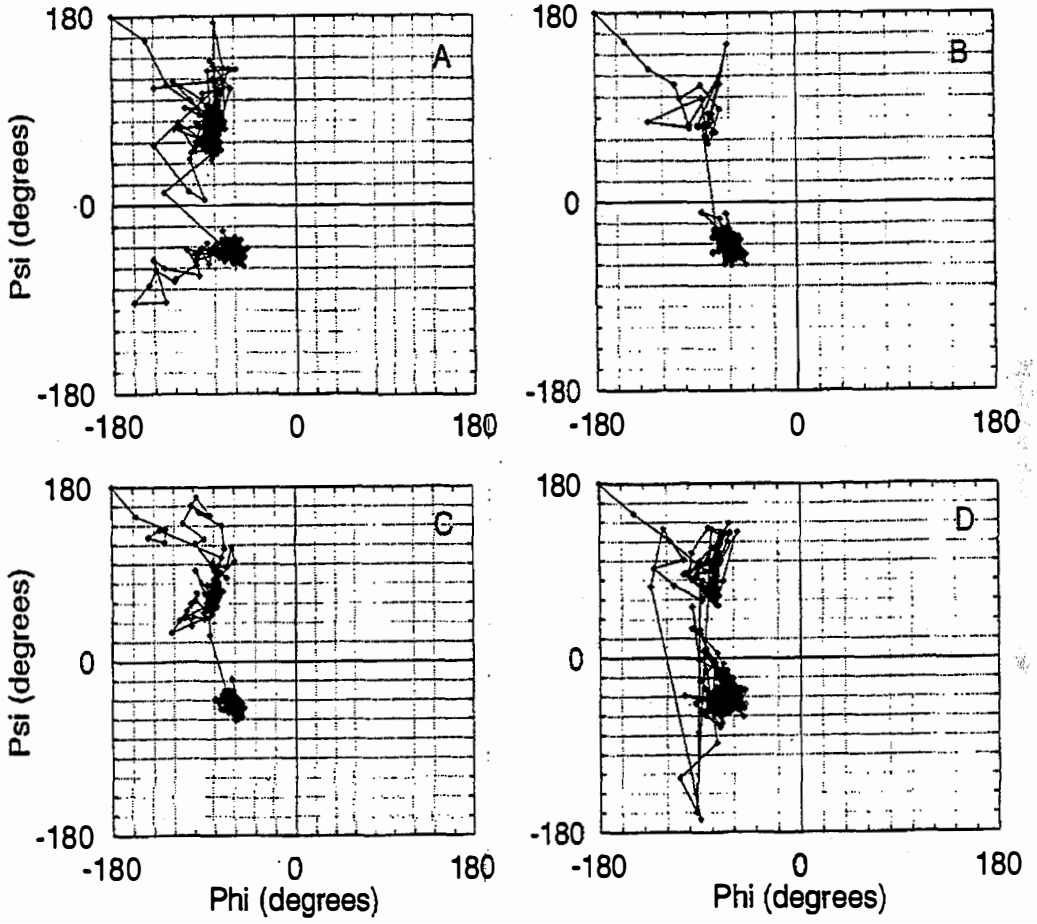
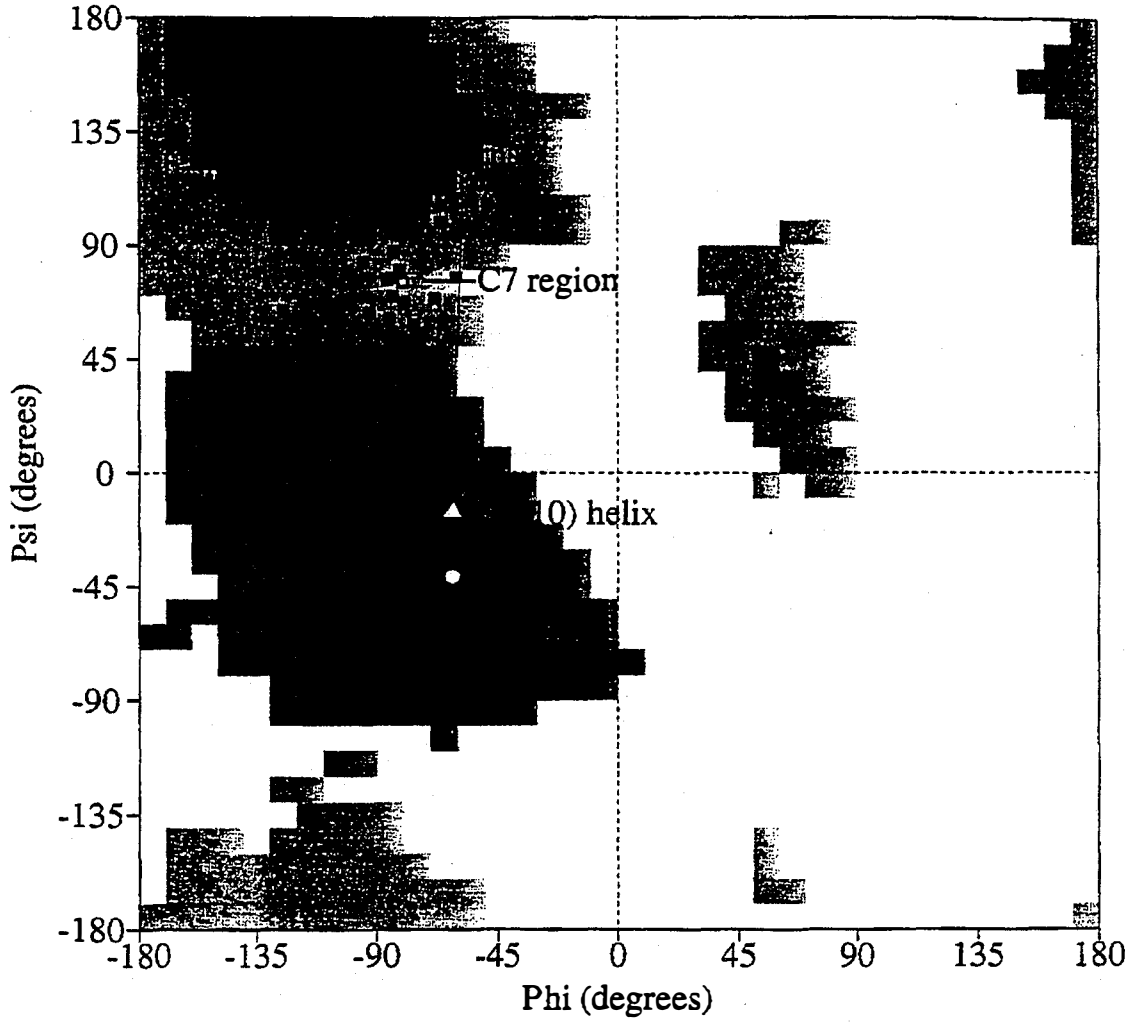


Figure 3. Ramachandran plot of the ϕ/ιangles of each residue in the polyalanine at 50 ps (squares) and 250 ps (triangles). The cluster of squares in the upper left-hand quadrant is the energetic minimum chiefly occupied by the structures during the first 100 ps of the simulation when the residues were strongly hydrogen bonded to their neighbors two residues away (*i, i+4* hydrogen bonds). The cluster of triangles in the lower right-hand quadrant represents the ϕ,ι angles of the α-helix formed at 250 ps. The darkest regions A and B are the regions where good quality X-ray crystallographic structures expect to place 90% of their residues. (Plot courtesy of PROCHECK.[6])



[6] Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; & Thornton, J.M. (1993) PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Cryst. 26*, 283-291.

Nucleation of helix formation. Helix-coil transition theory predicts that the rate-limiting step of folding of a 20-mer of polyalanine at equilibrium will be the formation of the first helical, *i, i+4* hydrogen bond.[7] (An α-helical hydrogen bond joins the backbone carbonyl of residue *i* and the amide proton of residue *i+4*.) In contrast, our non-equilibrium simulations of the folding process indicate that nucleation is not rate-limiting but that propagation of the α-helical hydrogen bonds is. Although some simulations only had one nucleation event, some at 450°K and 15 Å cutoff had two or three nucleating events. Sometimes two residues separated by one intervening residue would nucleate a patch of helix together.

Propagation. Propagation of the helix along the polymer did not appear to follow a specific order. Some simulations formed a helix in the middle of the polymer first, then the amino terminus, then the carboxy terminus, and then all three helices fused. Some simulations showed a helix forming at the amino terminus first, then propagating to the middle, then joining with a C-terminal helix. At least one simulation in the group of 126 helix-forming simulations at 450°K, 15 Å cutoff exhibited one of the combinations of fusing and/or propagating the N-terminal third, the middle third, and the C-terminal third until an entire helix was formed. There was no pattern to the method or order that the polymer thirds formed a helix.

Nonhelical hydrogen bonds. Probably the best predictor of whether or not a simulation will form a helix within 500 ps is whether or not the polymer becomes entrapped in a nonhelical conformation, glued together by nonhelical hydrogen bonds. In some cases, the polymer can only break out of the nonhelical conformation and form a helix after it first breaks the tethering nonhelical hydrogen bonds. Using Arrhenius-like behavior for the decay of nonlocal H-bonds we calculated the free energy for breaking a single nonlocal H-bond is approximately 0.25 Kcals/mole. The helix formation consists of two time scales. One is the delay time for the helix formation which is the time taken for the structure to attain the maximum number of nonlocal H-bonds that it had and 2) the decay time which is the time taken for all the nonlocal H-bonds to decay to helical H-bonds. Using these two times we have predicted that the folding time scales in out simulations would be in the nanoseconds scale for 100% folding yield. This paper is being communicated to *Protein Science*.

Other Sequences.

Polyglycine is known to be a helix breaker. We found that for polyglycine only 10% of the simulations formed a right-hand helix and these helices were short lived and not stable. Polyglycine sampled all four quadrants of the Ramachandran plot .

Further tests were done to calculate the helix propensities for various aminoacid substitution in the polyalanine chain. The percentage helicity for alanine (86.21%) > Leucine (79.3%) > Glutamine (63.3%). This is in accordance with the helix propensity data from experiments for these residues.

In addition to DOE, this work was supported by the National Science Foundation and the National Institutes of Health.

---

[7] Cantor & Schimmel (1980) *Biophysical Chemistry: The Behavior of Biological Macromolecules, III.* (San Francisco, W.H. Freeman & Comp., pp. 1046-1072.

Sidechain Rotamer Predictions in Folded Proteins - Matt J. Carlson, Derek Debe, and William A. Goddard, III

Within our research group's proposed hierarchical generic protein (GP) folding method, after several candidate protein folds have been selected, it is desirable to perform a full atom simulation on the candidate folds. We have developed a simple and efficient method to place sidechain atoms onto the main chain of a candidate protein fold.

All possible sidechain conformations are represented by a set of rotamers, or rotational states, where sidechain bonds and angles are fixed and a single rotamer state is determined by the dihedral angles formed in the sidechain. Some sidechains, such as alanine, have only one possible sidechain placement. Other longer sidechains, such as lysine, have several torsion angles and require more rotamer states to properly represent its possible conformations. In total, 108 rotamers were constructed from the protein database to represent the 20 amino acids, with the longest chains having up to 16 rotamer states.

This is a smaller set of rotamers than used in most of the current literature, but the GP hierarchy requires the placement of sidechains on many candidate folds, so a tradeoff on total accuracy was made for the sake of speed. This is allowable since the rotamer addition step occurs prior to further atom refinement steps to be performed on candidate folds.

With 108 rotamers in our data set, the number of possible total states of the protein of length N is about 6^N. This is computationally prohibitive to calculate, so our algorithm progressively eliminates unlikely rotamer states. A special force field, described later, is used to evaluate the energy of prospective rotamer placement.

The first step in the prediction is to examine all rotamer states for a specific residue and calculate their interaction with the backbone of the protein. The interaction energy between the sidechain and the backbone of the protein is calculated and rotamers with extremely poor backbone interactions are removed from the rest of the calculation forming a rotamer set R_1. The rotamer with the best backbone interaction is retained for each residue and used to construct the initial all-atom protein template. This template protein is used as the "ground state" protein which is then perturbed by systematically changing the rotamer state of each residue and calculating its interaction with the rest of the "ground state" protein. Only rotamers which survived the backbone interaction criteria are used. The two rotamers that have the best improvement from the protein "ground state" are kept and this new protein is considered to be our new template. In addition, a subset of rotamers, R_2, is constructed out of all rotamers that are lower in energy than the initial "ground state". The new template protein is used as a "ground state" protein for a similar rotamer analysis with rotamer set R_2. The best two rotamer improvements are preserved and a new template is created. This process of rotamer investigation using rotamer set R_2 is continued until no rotamer in R_2 lowers the energy of the current template. Finally, as a precaution, all rotamers from set R_1 are used in a search for new lower energy templates. All R_1 rotamers are checked against the template and the best two improvements are kept until no rotamers in the R_1 rotamer set are lower in energy than the template protein. This final template is used as the prediction of sidechain position for this protein fold candidate.

The force field used in all energy calculations was a modification of the standard Dreiding generic force field. Due to the steep inner wall of most 12-6 van der Waals (vdW) interactions, it was necessary to use a Morse term for all VdW terms. In addition, it was necessary to reduce the standard gamma value for many of the force field VdW interactions. Gamma values varying

between 7 and 12 allowed the forcefield to be more forgiving of close interactions that often arise when attempting to pack a sidechain rotamer into a constricted environment. This softening of the potential was able to offset the roughness of our search due to the small number of rotamers in our rotamer set.

A final approximation was made in the representation of LYS and ARG. Both amino acid types have long sidechains. Torsion data in the database are sparse for torsions beyond the third torsion of each sidechain. In addition, these sidechains are often difficult to pack within a protein fold due to their length. To accommodate these two longer sidechains, only atoms required to preserve their first three torsion angles were used when calculating interaction energies. The tips of these rotamers are then added back onto the residues once the final all-atom prediction is made.

Our rotamer prediction method compares favorably with other methods in the literature. The strength of the method lies in its good performance with a smaller rotamer set, about half the number of rotamers of other current methods. This is achieved with our custom force field used in the energy evaluation. Thus, our method has the capability to operate on a large number of candidate protein folds to better fit within our hierarchical framework.

Process Simulations – Mario Blanco

We have continued our industrial collaborations in the areas of process simulation. These include:

Scale Control : Chevron Petroleum Technology Company: Develop a new scale dissolver formulation
Polymer Compatibilizers and Silane Coupling Agents: Owens Corning: Develop new software for compatibility predictions that save laboratory efforts
Wear Inhibitors: Chevron Research Technology Center: Develop a self-assembled monolayer model for metal wear inhibition by dithiophosphates
Gasoline Reformer Catalysts: Exxon Research: Work on new catalysts for gasoline reformulation
Water resistant adhesives: Avery Dennison: Identified a mechanism for outstanding water resistance of new acrylic modified adhesives

Perhaps the most important finding of this quarter is the development of a new formulation to be used for scale (Barite) dissolution in oil and gas production. Large amounts of DTPA are used to control the precipitation of insoluble salts such as barite. The co-precipitation of barium and other divalent cations such as radium caused a great many problems in oil productions, from shutdowns for maintenance and repair, to large amounts of low level radioactive solids that need to be disposed of.

The new formulation replaces 50% of an expensive chemical (diethylpentaaminoacetate or DTPA) with a common aminoacetate ligand in catalytic concentrations. Thus, the new barite dissolver formulations require less than half the usual amount of DTPA dissolver (0.09M as compared with 0.18M) in combination with a second aminoacetate ligand in concentrations << 0.09M. This work was done in collaboration with USC and Chevron Petroleum Technology Company.

We are now proceeding with a joint industrial partnership to develop new, scale inhibitors to prevent the formation of naturally occurring radioactive materials (NORM) in oil production.

These are co-precipitates of insoluble radium salts that occur in the normal operations of secondary oil recovery. We have identified four companies willing to fund a 3 year effort to develop new water soluble materials capable of maintaining radium in solution at selectivities (over barium) greater than 10^9.

The basic idea is:

The co-precipitation of radium, from the produced water into the host scale matrix (most often barium sulfate), creates a radioactive deposit (so-called NORM).

If one could find a chemical (ligand) that would bind strongly with dissolved radium, then radium would not enter the solid scale matrix, and hence any scale that does form would not be radioactive.

Radium concentrations in the produced water are typically very low (less than 1 part per billion); in theory then, it would require only a low dose of radium specific ligand (RSL) to keep the radium in a dissolved form in the produced water. If successful, this approach would be an inexpensive NORM prevention method.

To be practical, the RSL must be a ligand that is extremely selective for dissolved radium because typical oil field produced water has substantial concentrations of other cations that will compete with the radium to bind with the ligand.

Chevron sponsored a small molecular modeling study to consider the feasibility of identifying a RSL at the California Institute of Technology (Caltech) under the direction of Prof. William A. Goddard III. This study indicated a rigid macrocycle preorganized with the right cavity size and lined with nitrogen donor atoms might provide sufficient selectivity for radium over competing divalent cations. Discussions with Dr. J.M.J. Frechet of the University of California, Berkeley indicate they can synthesize macrocyclic core polymeric surfactants expected to meet these requirements.

We recommend PERF member companies sponsor a small follow-up theoretical study at Caltech (one year post-doc) to better define the chemical structure(s) that are viable RSL candidates. Later studies would incorporate an experimental program to test the performance of the most promising RSL candidates.

Relevant Publications:

Jiang-S.Y., Frazier-R., Yamaguchi-E.S., Blanco-M., Dasgupta-S., Zhou-Y.H., Cagin-T., Tang-Y.C. and Goddard-W.A., the SAM Model for Wear Inhibitor Performance of Dithio-phosphates on Iron-oxide, *Journal of Physical Chemistry B,* V101, 39, pp. 7702-7709,1997.

M. Blanco, Y. Tang, P. Shuler, and W.A. Goddard III, Activated Complex Theory of Barite Scale Control Process, *Molecular Engineering*, in press

Ramachandran-S., Tsai-B.L., Blanco-M., Chen-H., Tang-Y.C., Goddard-W.A., Atomistic Simulations of Oleic Imidazolines Bound to Ferric Clusters, *Journal of Physical Chemistry A*, 101,1,83-89,1997.

Ramachandran-S., Tsai-B.L., Blanco-M., Chen-H., Tang-Y.C., Goddard-w.a., Self-assembled Monolayer Mechanism for Corrosion Inhibition of Iron by Imidazolines, *LANGMUIR*, V12, No 26, 6419-6428, 1996.

Jiang-S.Y., Dasgupta-S., Blanco-M., Frazier-R., Yamaguchi-E.S., Tang-Y.C., Goddard-W.A., Structures, Vibrations, and Force-fields of Dithiophosphate Wear Inhibitors from Ab-initio Quantum-chemistry, *Journal of Physical Chemistry*, 100, 39, 15760-15769, 1996 .

Blanco, M., Tang, Y., Goddard, W.A., Applications of Simulations to Upstream Petroleum Technologies, Technology Exchange Meeting, Dhahran, Saudi Arabia, 1995

Abner J. Mintz, Master Thesis, June 1997, Materials and Process Simulation Center, Caltech

8. Project Title Separations by Reversible Chemical Association
Principal Investigator(s) C. Judson King
Project Site Department of Chemical Engineering
University of California at Berkeley
Lawrence Berkeley Laboratory

Description:

Investigate the application of reversible chemical complexation as a means of separation of polar organic compounds from dilute aqueous solutions. Solutes of interest include oxygenated compounds produced by fermentation, such as carboxylic acids, alcohols, and amino acids. These approaches can substantially reduce energy requirements, and minimize the wastes, of conventional chemical separations (distillation, evaporation, conventional solvent extractions) in industry.

General objectives of the project include the measurement of phase equilibria and the rates at which equilibria are established; establish guidelines for the selection of complexation agents; develop suitable methods of process integration (into other unit operations), and to define the most promising applications of the method.

Specific objectives include the identification of regeneration methods for organic acids extracted from solutions; identification of process approaches for the extraction of organic acids at pH above the pKa of the acid; and identification of suitable extraction agents for glycol and other polyfunctional alcohols.

1996-1997 Accomplishments:

Recovery of Carboxylic Acids -- Regeneration of Adsorbents.

Regeneration of Adsorbents

We have developed an effective way of regenerating basic solid sorbents laden with carboxylic acids. Instead of leaching with trimethylamine (TMA) in aqueous solution, as we had done heretofore, we have used TMA in an anhydrous organic solvent, e. g., methyl ethyl ketone (MEK). TMA in MEK leaches lactic acid from Dowex NMA- 1, a tertiary amine sorbent, essentially completely, and TMA (and MEK) are removed from lactic acid completely upon heating, through a cracking reaction. It had not been possible to remove TMA completely in this way from the aqueous leachant, because ionization of the TMA carboxylate suppresses the volatility of TMA. Furthermore, leaching with the organic solution of TMA in a fixed bed gives a substantial focusing effect, whereby the recovered TMA carboxylate is captured in the band ahead of the displacing organic TMA solution. This gives a much higher concentration of the carboxylic acid in the feed to the thermal decomposition reaction, an effect that is economically beneficial.

Sustained Adsorption Capacity

We have confirmed a prediction that the presence of carbon dioxide during adsorption of a carboxylic acid onto a basic sorbent such as Dowex MWA- I serves to buffer the pH and thereby enable a high degree of recovery of the carboxylic acid. Removal of carbon dioxide then enables regeneration of the sorbent. Similar effects have been found by others in solvent extraction processes.

Fractionation among Solutes

Stillage from fermentation ethanol plants contains both lactic and succinic acids and glycerol. We have found that Dowex MWA-I fractionates effectively between the carboxylic acids and glycerol. One potential means of separating between lactic and succinic acids is to utilize the much lower solubility of succinic acid for selective precipitation. We found that the solubility of succinic acid in ketone solvents is only slightly influenced by the simultaneous presence of lactic acid.

Adsorption of Poly -OH Compounds onto Activated Carbons

We have found that regeneration of glycols, glycerol, sugars and sugar alcohols from activated carbons is incomplete in the presence of water, but is complete following prior removal of co-adsorbed water. Furthermore, there is a substantial focusing effect for fixed-bed leaching of glycols with acetone, and of glycerol with methanol. This effect affords what should be an economical regeneration process.

1998 Plans:

We will work toward conclusion of the research and the two Ph. D. projects based thereupon. We will explore further the use of carbon dioxide to sustain capacity during adsorption of carboxylic acids, measuring and interpreting the phenomena involved and identifying optimal processing conditions. Three reports and at least two papers on the carboxylic acids research will be prepared for publication.

We will carry out research to identify the causes for the unexpected deleterious effect of co-adsorbed water on the regenerability by solvent leaching of carbons that have adsorbed glycols, glycerol, sugars and/or sugar alcohols from aqueous solution. We shall then prepare a report and a paper.

Annual Technical Summary Report:

Work during 1996-97 concentrated upon four areas C (1) regeneration of basic sorbents useful for recovery of carboxyhc acids from aqueous solutions at pH above the pKa of the carboxylic acid, (2) use of carbon dioxide to sustain capacity of basic sorbents at pH above the pKa of the adsorbing carboxyhc acid, (3) fractionation among solutes in carboxylic-acid-beafing solutions, and (4) use of carbon adsorbents for recovery of glycols, glycerol, sugars, sugar alcohols and other poly -OH substances from aqueous solution. In addition, we published a paper on prior research dealing with adsorption of phenol from aqueous solutions by means of activated carbons.

<u>Regeneration of Adsorbents</u>

In previous research we used regeneration by an aqueous solution of trimethylamine (TMA) to recover carboxylic acids from solid sorbents and liquid extractants in the form of the TMA carboxylate. The carboxylate can then be decomposed thermally, to release TMA for recycle and to provide product acid. This approach gives high recovery for sparingly soluble carboxylic acids, such as succinic and adipic acids. However, removal of TMA is incomplete for highly soluble carboxylic acids, such as lactic acid.

We have now developed an effective way of regenerating basic solid sorbents laden with the more soluble carboxylic acids. Instead of leaching with trimethylamine in aqueous solution, we have used TMA in an anhydrous organic solvent, e. g., methyl ethyl ketone (MEK) or acetone. TMA in MEK leaches lactic acid from Dowex MWA- 1, a tertiary amine sorbent, essentially completely at reasonable stoichiometric ratios, and TMA (and MEK) are then removed from the organic solution of TMA lactate essentially completely upon heating, through a cracking reaction. In the organic solution ionization of the TMA carboxylate does not suppress the volatility of TMA.

Furthermore, we have found that leaching with the organic solution of TMA in a fixed bed gives a substantial focusing effect, whereby the recovered TMA carboxylate is captured in a small band ahead of the displacing organic TMA solution. This gives a much higher concentration of the carboxylic acid in the feed to the thermal decomposition reaction, an effect that is economically beneficial.

We also examined the effects of temperature and heat-up time on the rate of the cracking reaction. The reaction is complete in well under an hour if heat-up is sufficiently rapid. We also found that TMA forms a relatively insoluble adduct with succinic acid. This adduct can be recovered as a solid, and can be completely decomposed by heat to form solid succinic acid.

Sustained Adsorption Capacity

We have confirmed a prediction that the presence of carbon dioxide during adsorption of a carboxylic acid onto a basic sorbent such as Dowex NMA- I serves to buffer the pH and thereby enable a high degree of recovery of the carboxylic acid. Removal of carbon dioxide then enables regeneration of the sorbent. Similar effects have been found by others in solvent extraction processes.

Fractionation among Solutes

Stillage from fermentation ethanol plants contains lactic and succinic acids and glycerol. We have found that Dowex MWA- I fractionates effectively between the carboxylic acids and glycerol. One potential means of separating between lactic and succinic acids is to utilize the much lower solubility of succinic acid for selective precipitation. We found that the solubility of succinic acid in ketone solvents is only slightly influenced by the simultaneous presence of lactic acid.

Adsorption of Poly -OH Compounds onto Activated Carbons

We have found that regeneration of glycols and glycerol from activated carbons is incomplete in the presence of water, reaching an asymptote of about 75% removal of adsorbed solute. Vacuum stripping of co-adsorbed water serves to make the solute completely recoverable with the anhydrous solvent. Acetone and methanol have been the solvent used for recovery of adsorbed glycols and glycerol, respectively. The reasons for this striking effect are not obvious, and we are carrying out further research to elucidate the cause.

We have carried out experiments wherein anhydrous solvent is used to displace adsorbed glycols or glycerol from carbons. The solvent serves to displace the co-adsorbed water, giving essential complete removal of the solute. Furthermore, there is a substantial focusing effect in which the displaced solute is concentrated into a small band ahead of the displacing solute. This effect affords what should be an economical regeneration process.

We have found that aqueous solutions of acetone serve to regenerate sugars and sugar alcohols adsorbed onto carbon. The water is needed to provide sufficient solubility for the recovered solute.

Papers published

J. E. Kilduff & C. J. King, "Effect of Carbon Adsorbent Surface Properties on the Uptake and Solvent Regeneration of Phenol", *Ind. Eng. Chem. Research,* 36, 1603-1613 (1997).

9.	**Project Title**	Rational Enhancement of Enzyme Performance In Organic Solvents
	Principal Investigator	Alexander Klibanov
	Project Site	Massachusetts Institute of Technology

Description

This project seeks to elucidate the mechanisms of enzyme stability in organic solvents and to apply the knowledge to processes of industrial importance. The project will also investigate gas-phase biocatalysis. The phenomenon of "molecular memory" of enzymes will be explored. Rigid protein configurations in enzymes retain the conformation (imprint) induced by the original ligand/stability factor. The result of this molecular memory is an enzyme with altered catalytic characteristics or a protein with new binding site, which might be able to perform new chemical reactions.

1996-1997 Accomplishments:

The overall goal of this project is to rationally enhance the catalytic performance of enzymes in organic solvents, which should broaden the utility of enzymes as industrial catalysts. During the past year, we have focused on (1) maximizing the catalytic activity of cross-linked enzyme crystals in organic solvents by controlling the acid-base balance in the system; (2) assessing the secondary structure of enzymes in neat organic solvents as compared to those in water and in aqueous-organic mixtures; (3) explaining and predicting the substrate specificity in organic solvents of lyophilized enzymes (for which, in contrast to their crystalline counterparts, the structure is not known); (4) examining mechanistically the catalytic activity and secondary structure of subtilisin dissolved (as opposed to suspended) in organic solvents. The following are our major advances:

- We have discovered that the catalytic activity of cross-linked enzyme crystals suspended in organic solvents can be dramatically increased by adding appropriate “organic buffers” (i.e., mixtures of an acid and its conjugate base). This phenomenon has been rationalized mechanistically.
- We have directly demonstrated by means of FTIR spectroscopy that while many neat organic solvents do not denature enzymes, mixtures of the same solvents with water do. This explains why enzymes are catalytically active in neat organic solvents (and in water) but not in aqueous-organic mixtures.
- We have developed a simple thermodynamic methodology that can explain the enzymatic substrate specificity in organic solvents even if the enzyme structure is not known.
- We have explained, using FTIR structural analysis, why subtilisin is catalytically active when dissolved in some organic solvents but not in others.

1998 Planned Activities:

We shall continue to address the key problems pertaining to the rational enhancement of enzyme catalysis in organic solvents. Specifically, the following directions of research will be pursued.

(1) We will investigate the phenomenon of “enzyme memory” in organic media. Specifically, the dependence of the catalytic activity of enzymes in a given solvent will be examined as a function of the enzyme sample history, i.e., what treatments/solutions the enzyme

had been exposed to prior to placement in the solvent. This study will point to the optimal enzyme preparations in organic solvents and should be mechanistically insightful as well.

(2) We will endeavor to develop a new molecular-modeling methodology to explain quantitatively the enzymatic enantioselectivity both in water and in organic solvents. This methodology will combine the currently accepted vacuum molecular mechanics with the continuum solvation method, thereby allowing us to take into account the interactions of the solvent with the enzyme and the substrate.

(3) To shed more light on enzyme mechanisms in organic solvents, we will solve the X-ray crystal structures of a complex of the model enzyme subtilisin with a transesterification substrate (S)-*sec*-phenethyl alcohol, as well as that of the acylenzyme cinnamoyl-subtilisin in acetonitrile. Comparison of these structures with those in water and with that of the free enzyme in acetonitrile will a) reveal any changes induced by the substrate; b) show whether the substrate and/or the acyl moiety displace any bound water and/or solvent from the active center, thus contributing to the energetics of the enzyme-ligand binding; and c) verify the formation of the non-covalent enzyme-substrate complex in an organic solvent.

(4) We will continue our investigation of processes catalyzed by enzymes dissolved (rather than suspended) in organic solvents. To this end, we will study the protease-catalyzed cleavage of proteins (where both the enzyme and the proteins are dissolved) in glycerol. Particular attention will be paid to the selectivity (positions of cleavage in glycerol compared to that in water. In addition, amylase-catalyzed cleavage) of starch will be examined; the resultant glycerol glucosides may have interesting biomaterial properties.

Annual Technical Summary Report:

The overall goal of this project is to rationally enhance the catalytic performance of enzymes in organic solvents, which should broaden the utility of enzymes as industrial catalysts. During the past year, we have focused on (1) maximizing the catalytic activity of cross-linked enzyme crystals (CLCs) in organic solvents by controlling the acid-base balance in the system; (2) assessing the secondary structure of enzymes in neat organic solvents as compared to those in water and in aqueous-organic mixtures; (3) explaining and predicting the substrate specificity in organic solvents of lyophilized enzymes (for which, in contrast to their crystalline counterparts, the structure is not known); and (4) examining mechanistically the catalytic activity and secondary structure of subtilisin dissolved (as opposed to suspended) in organic solvents. Below the technical highlights of the past year's work are outlined and summarized.

- Lyophilized subtilisin Carlsberg exhibits "pH memory" in organic solvents, i.e., its catalytic activity in such media is profoundly affected by the pH of the aqueous solution from which the enzyme was lyophilized. In contrast, cross-linked crystalline subtilisin displays no appreciable pH memory. This disparate behavior has been rationalized by us in terms of the different protonation states of the two enzyme preparations in organic solvents. Since CLCs of enzymes in organic solvents offer performance and mechanistic advantages over their lyophilized counterparts, the opportunities for pH maximizing their catalytic activity in such media have been explored. It has been found that subtilisin CLCs in pentanone can be activated more than 100-fold by adding a "buffer" consisting of a mixture of a suitable acid and its sodium salt (conjugate base). The maximal activation has been achieved (1) when the pK_a value of the buffer acid in water lies in the appropriate range, (2) at an optimal acid/base composition of the buffer mixture, and (3) when the buffer mixture is present in a sufficient concentration.

- FTIR spectroscopy was used by us to quantitatively assess the secondary structure of proteins in aqueous–organic mixtures ranging from pure water to a pure solvent. For every such solution/suspension, the α-helix content of the protein was independently calculated from the amide I and amide III spectral regions (which gave essentially identical results). In all cases studied (two unrelated enzymic proteins – lysozyme and subtilisin; three dissimilar water-miscible solvents – acetonitrile, tetrahydrofuran, and 1-propanol), the protein secondary structure was much more native-like in pure organic solvents than in most water–solvent mixtures, e.g., 60% (v/v) organic solvents. In fact, placing lyophilized (as well as crystalline) proteins in the anhydrous solvents tested had no appreciable effect on the α-helix content, whereas the latter declined markedly in the 60% (v/v) solvents. This behavior was found to be kinetically controlled, i.e., to be due to inherent restrictions on protein conformational mobility in anhydrous, in contrast to aqueous–organic, media.

- A simple methodology has been successfully employed by us to explain the solvent dependence of the substrate specificity of enzymes in organic media. This methodology, which does not require the knowledge of the enzyme structure and is thus applicable to lyophilized and other noncrystalline enzyme preparations predicts that the k_{cat}/K_M ratio for two substrates should be proportional to their Raoult's law activity coefficients. This approach has been validated for two enzymes, subtilisin Carlsberg and α-chymotrypsin, catalyzing the propanolysis of unnatural (in addition to natural) ester substrates in a variety of anhydrous solvents.

- FTIR spectroscopy has been used by us to quantify the α-helix and β-sheet contents of subtilisin Carlsberg dissolved in several nonaqueous, as well as aqueous, solvents. Independently, the catalytic activity of the enzyme has been measured in the same solvents. While our previous FTIR studies revealed no connection between the secondary structure and enzymatic activity for subtilisin *suspended* in various organic solvents, a very different situation is observed herein for the *dissolved* enzyme. Specifically, if either the α-helix or β-sheet content in a given solvent is higher or lower than in water, no appreciable enzymatic catalysis is observed. Conversely, when the secondary structure of subtilisin dissolved in a given nonaqueous solvent is similar to that in water, so is the enzymatic activity.

10. Project Title Strategies for Biocatalyst Design in Extreme Conditions for Industrial Applications

Principal Investigator(s) Robert Meglen and Gene Petersen

Project Site National Renewable Energy Laboratory

Description

Both random and site-directed mutagenesis techniques have been successfully used to alter the primary sequence of enzymes to improve stability and catalytic activity in nonpolar solvents. Since enzyme activity is in part due to the three-dimensional structure of the enzyme, one may infer that substitutions in primary sequence induce secondary and tertiary structural changes that affect enzyme activity. Successive application of random mutagenesis provides a valuable database for exploring the relationship between structural modification and enzyme activity. However, if a "rational design" technology is to emerge from enzyme modification experiments, it will be necessary to develop data analysis tools for converting these databases into information that relates specific structural changes to enzyme activity and permits predictive capacity. *The goal of this project is to use existing data from random mutagenesis experiments to develop quantitative structure activity models that can be used to assist in the selection of potential enzyme mutation strategies.*

Previous reports described the development of five modeling strategies and the expansion of our model-generated crystal structures database to 130 T4 lysozyme mutants. Molecular modeling was used to expand the database of known structures from 130 computer generated 3D structures to 212 variants which have been produced in the laboratory but not subjected to x-ray crystallographic determination. A position independent quantitative structure activity relationships (QSAR) predictive model was extended to 1197 *de novo* T4 lysozyme mutants. Predictions for potential experimental mutations were made and experimental verification of these predictions are planned. A position independent QSAR predictive model was extended to two additional proteins, 1748 *de novo* staphylococcus nuclease mutants and 247 subtilisin E mutants.

1996 - 1997 ACCOMPLISHMENTS:

The goals for fiscal year 1997 were to further refine and develop the structural database for quantitative structure activity modeling, to select additional model systems, and to develop and test several predictive models for enzyme activity.

- QSAR modeling has been extended to green fluorescence protein (GFP). X-ray crystal structures for modeling 13 published mutants were combined with published fluorescence spectra.
- Molecular modeling (Insight/Discover/MSI) has also been used to generate 413 synthetic crystal structures on *de novo* GFP mutants.
- Initial exploratory studies have been initiated to combine empirical data from infrared spectroscopy to the prediction of secondary structure such as -helix and β- pleated sheets.

1998 Planned Activities:

Future work will be extended to experimental testing of selected predictions from among the 1197 *de novo* T4 lysozyme mutants. Namely, the model predicts that any of four aromatic amino acid substitutions (tryptophane, phenylalanine, histidine, tyrosine) for serine at residue 38

should lead to activities in excess of 500% of the wild type. Replacement of glutamic acid at residue 128 with one of these aromatic amino acids is predicted to enhance activity by 300-400%. Modest increases of 200-300% are predicted for substitution of asparagine by any of several amino acids (Ile, Leu, Val, Ala, Pro, Met, Gly, Cys).

Additional published and unpublished data will be acquired from currently active GFP researchers in order to test the new model predictions that have been pursued during the past 6 months. Specifically, the feasibility of performing some experimental verifications on selected test mutants to verify our predictions will be performed. (These experiments have a lower priority than the T4 lysozyme mutations but could provide additional verification for the QSAR modeling work.) Experiments to combine empirical data from infrared spectroscopy into the prediction of secondary structure will be continued.

Annual Technical Summary Report

Structure-Activity Relationships

The high cost of synthesis and/or testing put a premium on predictive ability in the synthetic drug industry. As a result a growing repertoire of computational tools has been brought to bear on the development of QSAR. Ignorance of the process relating molecular structure to biological activity, or the complexity of those processes, precludes the direct computation of those relationships and one must rely upon an empirical model to develop predictive capability. Application of QSAR methods to macromolecules is in its infancy, but the underlying assumptions and problems of prediction for enzymes suggest that the application of QSAR tools from the drug industry might be advantageous.

Evaluation of Molecular Modeling for Representing Three-Dimensional Protein Structures

Successful QSAR modeling requires two data matrices; one containing the experimentally measured properties to be predicted (activities, free energies of unfolding, etc.), and a matrix that contains structural information and salient physicochemical forces that is used as the basis for constructing a predictive model of the properties. *T4 lysozyme* was chosen as the initial validation system for modeling because data matrices on more than 170 variants exist in the literature. Attempts at extending QSAR modeling to systems beyond T4 lysozyme have been performed on two additional systems, *Staphylococcus nuclease* and *Subtilisin E*. However, these enzymes have limited data for testing and validating the 3-dimensional structural component of the modeling. A search for additional systems was performed.

An additional protein system has been identified for testing the QSAR modeling approach and its generality. The GFP isolated from the Pacific jellyfish, *Aequorea victoria*, was selected. This protein has become a valuable marker for molecular and cell biologists [1]. By linking the GFP gene to those encoding for other proteins one can visually track the protein products in living cells [2-5]. The key features of green fluorescent protein that suggest its utility for QSAR modeling derives from the unique protein structure (β - barrel). The cylindrical shape [6-8] of GFP consists of 11 strands of β-sheet which form the walls of the cylinder (see Figure 1). Six short α-helix segments cap the two ends of the cylinder. The chromophore responsible for the protein's fluorescence, located near the geometric center of the cylinder, results from the oxidation and cyclization of the three residue sequence Ser^{65}-Tyr^{66}-Gly^{67} [9-10]. Thirteen GFP mutants have been experimentally produced in various laboratories by substituting various amino acids within the chromophore and at significant distances from the chromophore [11-14]. While modifications within the chromophore have been observed to affect the excitation

maximum and the fluorescence intensity, amino acid substitutions significantly outside of this region also affect the protein's spectral characteristics [8]. These observations suggest that GFP could provide a useful system for exploring the QSAR prediction modeling of proteins.

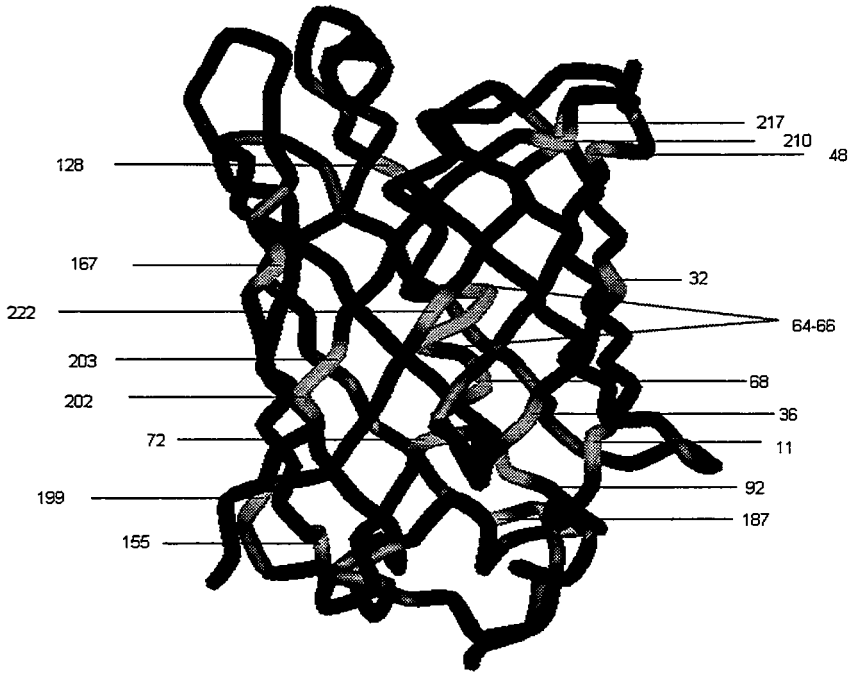


Figure 1. Green fluorescent protein residue locations at which mutations have been made. Additional locations not shown on figure: 5, 8, 16, 22, 25, 41, 57, 69, 71, 76, 83, 100, 105, 115, 118, 156, 160, 176.

Previous reports described work that showed significant structural changes can be induced by substituting residues at key "leverage" points in *T4 lysozyme*, *Staphylococcus nuclease* and *Subtilisin E*. Therefore, modeling studies were focused on residue locations that were anticipated to induce significant structural changes. Figure 1 shows a rendered version of the barrel protein backbone. The light shaded regions with pointers show the locations where mutants (19 possible amino acid substitutions to the wild type at each designated location) have been synthesized and dynamics minimized. Additional locations for mutations are not shown in this figure to preserve clarity. These locations were selected because previous QSAR models for lysozyme suggested that residues at the beginning or end of secondary structural features (α-helix orβ-sheets) might lead to significant tertiary structural modifications. Based on the initial results it appears that the calculated stabilization energies are significantly different for several of these locations.

Crystal structures for all 13 published mutants that have been obtained and used as starting points for three dimensional QSAR modeling. Excitation maxima, fluorescence intensities and other spectral features were obtained and used as the dependent variables matrix (y-matrix). In

addition, the amino acid principal property files (described in previous reports) that comprise the independent variable matrix (x-matrix) in QSAR modeling were also developed. It is clear that much of the experimental work on generating mutants has been focused on the residues within or immediately adjacent to the chromophore sequence Ser^{65}-Tyr^{66}-Gly^{67}.

List of GFP Mutants Generated by Molecular Modeling

(and which have been produced experimentally)

Wild type	S202F, T203I
203I	T203H,S65T
E222G	T203Y, F64L, S65G, S72A
Y66W	T203Y, S65G, V68L, S72A
Y66H	T203W, S65G, S72A
I167V	
I167T	
S65T	

Lysozyme studies reported last year were limited to the pseudo-crystalline state for the enzyme using the X-ray crystal structures as starting points and solvent effects on structure were minimal. While these results have been promising, studies of the GPF β -barrel indicate that solvent-protein interactions are very important. Therefore, molecular dynamics studies of the effect of water on the structural stability of the protein have been initiated. Figure 2 illustrates the significance of solvent stabilization. The upper left frame shows the secondary rendered version of the wild type protein without solvent (X-ray structure basis). The rendered version of the final structure after 10 ps of dynamics calculation (at 300° K) is shown in the upper right frame. The β barrel has clearly disassembled, leaving only four small sections of hydrogen-bonded β sheets and mostly random coil. One explanation for this is that the strength of the hydrogen-bonded β-sheets is not sufficient to maintain the tertiary structure. The lower left frame shows the same X-ray based structure with a 5 Å solvation shell (approximately 1200 water molecules) at time zero. The frame on the lower right shows the same assembly after 10 ps. While the water shroud has begun to disperse (evaporate), the tertiary structure of the protein is relatively unchanged. This indicates that the water-protein interactions provide significant stabilization for the structure. The solvation energy stabilization for hydrophilic residues is well documented in the literature and this result is expected. However, the significance of this observation is that our computational methods are reproducing the known characteristics of the proteins. Furthermore, since the present QSAR study of the green fluorescence protein is sensitive to the stabilization energy, continued exploration of the role that solvation energy plays in computed stabilities. However, the number of mutants that can be treated with full solvation dynamics is limited by the speed of the computations. A 10 ps dynamics run without solvation (approximately 1000 vibrations of each atom) takes about 60 hours. Addition of 1200 water molecules doubles the computational time (about 124 hours). Goddard and co-workers at Cal Tech have developed alternate computational strategies for simulating the effect of water that should shorten the calculation time.

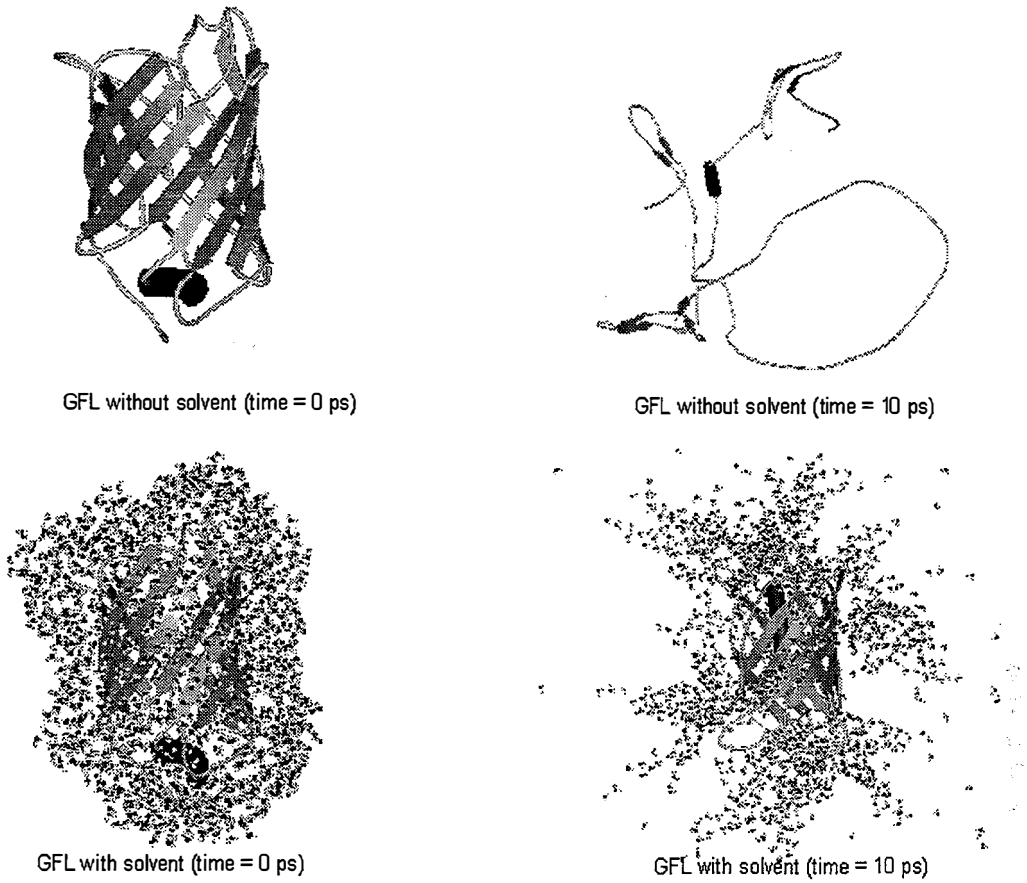


Figure 2 Comparison of the structural stabilization effect of water on the dynamics calculation of Green Fluorescent Protein.

Empirical Measurements related to Secondary Structure

While the main goal for this work has been to use existing data from random mutagenesis experiments to develop quantitative structure activity models that can be used to assist in the selection of potential enzyme mutation strategies, the ultimate selection tool would be coupled with an experimental measurement tool. That is, a rapid experimental measure of structural changes which could be used in conjunction with modeling would improve the rate at which one could screen mutants and assist in the directed evolution of desired enzyme modification. While X-ray crystallography and NMR measurements provide structural change information, they are not yet sufficiently routine and rapid to serve as screening tools. Other spectroscopic methods such as; fluorescence, circular dichroism, FTIR, Raman, are commonly used to analyze the polypeptide folding of proteins. Recently, several publications have described the use of FTIR spectra for secondary structure analysis [15-21]. An exploratory study has been initiated to examine the potential of IR spectroscopy to provide empirical inputs to the QSAR modeling already performed. (A rapid experimental structural probe would be a valuable adjunct to QSAR modeling.)

The differences in amide bond geometric orientations of α-helix, β-sheet, turn, and random coil structures allow for differences in vibrational frequencies and distinct IR signals in the amide I, II, and III bands (See Figure 4). The amide I vibrational region (1600-1700 cm^{-1}) corresponds to the C=O stretch weakly coupled with C-N stretch and N-H bending. The amide II region (1500-1600 cm^{-1}) represents C-N stretch strongly coupled with N-H bending. The amide III region (1200-1300 cm^{-1}) is N-H in-plane bending coupled with C-N stretching and also includes C-H and N-H deformation vibrations.

The five proteins used in this study were chosen because they represent varying content of secondary structural compositions. Alpha-chymotrypsin and immunoglobulin G are both primarily β-sheet proteins with little β-helix structure. Lysozyme was used because of its high α-helix content and this protein was the initial model upon which QSAR predictions were first undertaken. Similarly bovine serum albumin was chosen because it has relatively high α-helix content. Subtilisin E was used in this study because it represents a protein with moderate content of both α -helix and β-sheet. Table 1 shows the secondary structural compositions (determined by X-ray structural analysis) of the proteins used in this study.

Table 1
Secondary Structural Compositions of Proteins Used
(determined by x-ray structural analysis [Fu, et. al., Reference 5])

	Alpha Helix	Beta Sheet	Other
alpha chymotrypsin	8%	50%	42%
bovine serum albumin	55%	27%	18%
immunoglobulin G	3%	67%	30%
lysozyme	45%	19%	36%
subtilisin E	28%	23%	49%

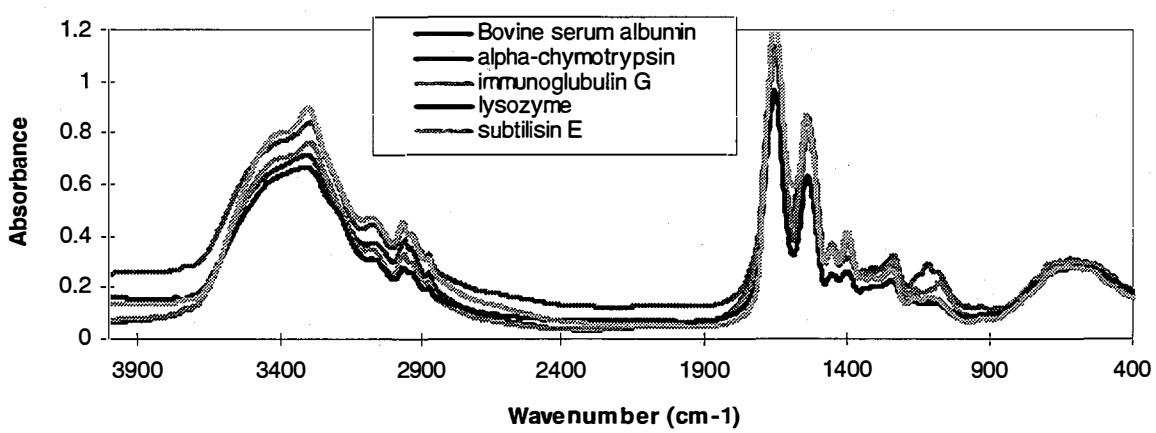
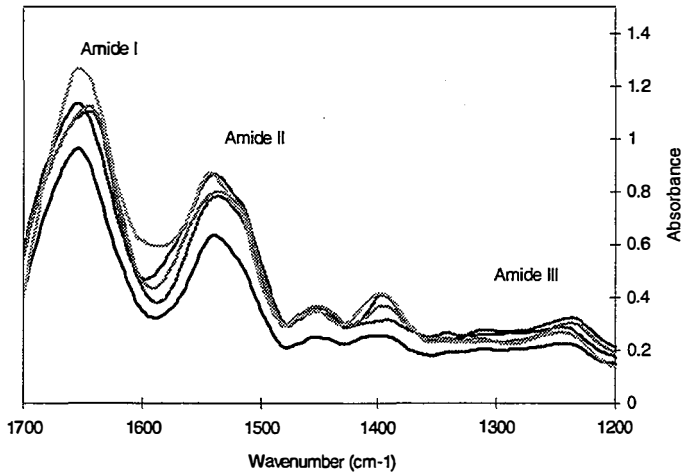


Figure 3. Infrared Spectra of Proteins with Varying Amounts of Alpha Helix and Beta Pleated Sheets

Figure 4. Expanded Region Showing Amide Bands



Transmission FTIR spectra were obtained on commercially available proteins. Approximately 10 mg of each protein were pelletized with KBr and spectra were obtained. Replicate spectra were obtained on 2 to 10 pellets for each protein. Figures 3 and 4 show representative FTIR spectra of five proteins used in this study.

The spectra were used in a projection to latent structures (PLS) calibration model. This multivariate data analytic method relatesthe systematic information in a matrix **X** to the information in a matrix **Y** with the purpose of predicting **Y** from **X**. In this case, the first step consisted of obtaining several protein spectra with known α-helix andβ sheet compositions. The X matrix consists of the absorption spectra (intensity versus wavelength in the range 400-4000 cm^{-1}). The Y matrix consists of the known (independently determined) alpha and beta content. The statistical technique, PLS, simultaneously calculates multivariate projections of the predictor variables (spectral intensities) and the dependent variables (α-helix andβ-sheet) so that the projection scores of the two data blocks are maximally correlated. In this way, a quantitative expression of the relationship (multivariate calibration curve) between the two matrices is obtained. The model derived in this way can then be used to extract the predicted structural composition of unknown samples for which the spectra have been obtained.

Figure 5 shows a plot of a fully cross-validated[8] model accuracies for α-helix based on FTIR for the study proteins. The root-mean-square (RMS) prediction accuracies are all approximately 5%, close to the precisions of other methods. Figure 6 shows a plot of a fully cross-validated[†] model accuracies for beta-sheet based on FTIR for the study proteins. The RMS prediction accuracies are all approximately 5%, also close to the precisions of other methods.

---

[8] **STATISTICAL NOTE:** It is common practice to construct a calibration model from one data set, and to subsequently assess the model accuracy by making "predictions" on an independent set not included in the construction of the model. Typically one half to two thirds of the samples are used to construct the model, while the remainder are "held back" for testing prediction accuracy. In the accuracy assessments shown here we have chosen the more rigorous assessment criterion of **full cross-validation.** As with other methods, accuracy assessment figures show the model predicted results plotted on the y-axis; and the "true," "known," or independently determined results plotted on the x-axis. However, in the full cross validation procedure, EACH predicted point was held back (not included) in the model construction. Namely, all of the other points are used to construct the model. Thus, an accuracy plot with 100 predictions plotted against the true or measured values is based on 100 models consisting of the 99 samples other than the one under scrutiny. This procedure allows for a more parsimonious use of the data collection and permits a robust accuracy assessment.

Figure 5. Prediction Accuracy Alpha Helix Percent in Several Proteins Based on PLS Modeling from Infrared Spectra

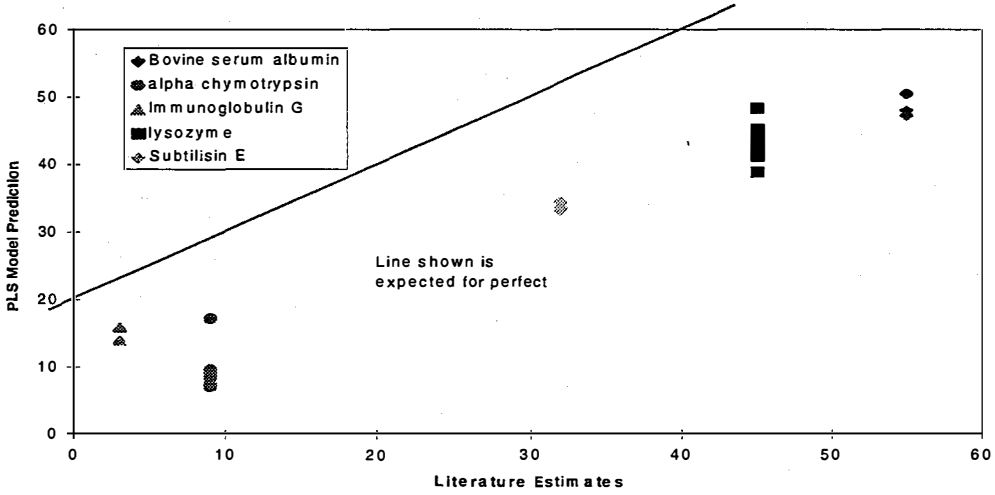
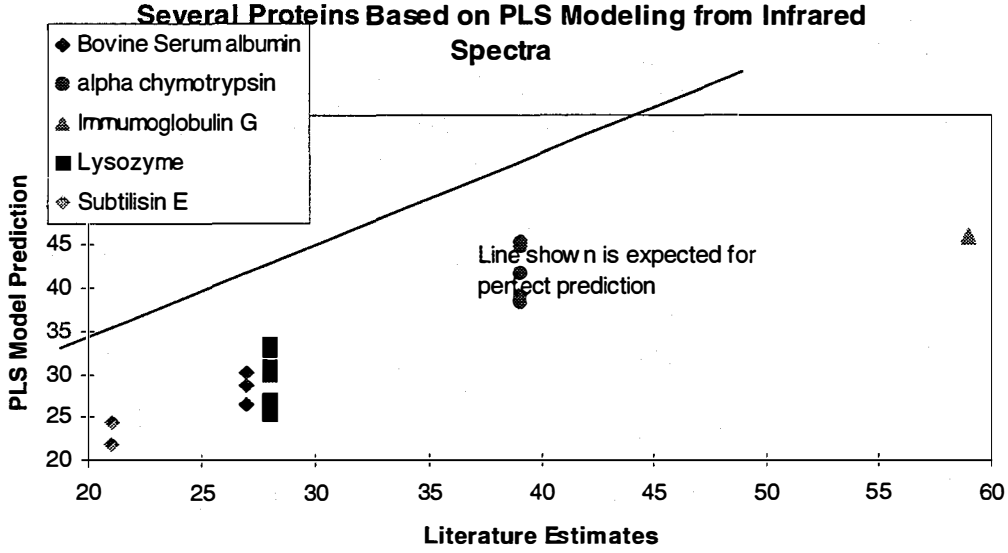


Figure 6. Prediction Accuracy Beta Pleated Sheet Percent in Several Proteins Based on PLS Modeling from Infrared Spectra



The success of the FTIR technique suggests that this method could provide a useful experimental adjunct to the computational method used in QSAR modeling. Additional work is planned.

Literature Cited

1. Special Issue, (**1996**), *Gene*, **173**, 5-11, "Fluorescent Proteins and Applications."
2. Hassler, S., (**1995**), *Biotechnology*, **13**, 103, "Green Fluorescent Protein: The Next Generation."
3. Stearns, T., (**1995**), *Current Biology*, **5**, 262-264, "The Green Revolution."
4. Cubitt, A., Heim, R. Adams, S, Boyd, A., Gross, L, and Tsein, R., (**1995**), *TIBS*, **20**, 448-455, "Understanding, Improving and Using Green Fluorescent Proteins"

5. Cheng, L., Fu, J., Tsukamoto, A. And Hawley, R., (**1996**), *Nature Biotech.*, **14**, 606-609, "Use of Green Fluorescent Protein Variants to Monitor Gene Transfer and Expression in Mammalian Cells."
6. Gura, Trisha, (**1996**), *Science*, **273**, 6-Sept-96, 1336, "Structure of Gene-Tag Protein Solved."
7. Ormo, M, Cubitt, A., Kallio, K., Gross, L., Tsien, R., Remington, S., (**1996**), *Science*, **273**, 6-Sept-96, 1392-1395, "Crystal Structure of the *Aequorea victoria* Green Fluorescent Protein."
8. Yang, F., Moss, L. G., and Phillips, G. N., (**1996**), *Nature Biotech.*, **14**, 1246-1251, "The Molecular Structure of Green Fluorescent Protein."
9. Niwa, H., Inouye, S., Hirano, T., Matsuno, T., Kohma, S., Kubota, M., Ohashi, M., and Tsuji, F., (**1996**), *Biochemistry*, **93**, 13617-13622, "Chemical Nature of the Emitter of the Aequorea Green Fluorescent Protein."
10. Heim, R. Prasher, D., Tsein, R., (**1994**), *Proc. Natl. Acad. Sci. USA*, **91**, 12501-12504, "Wavelength Mutations and Posttranslational Autoxidation of Green Fluorescent Protein."
11. Ehrig, T., O'Kane, D., and Prendergast, F., (**1995**), *FEBS Letters*, **367**, 163-166, "Green Fluorescent Protein Mutants with Altered Fluorescence Excitation Spectra."
12. Delgrave, S., Hawtin, R., Silva, C., Yang, M. And Youvan, D., (**1995**), *Biotechnology*, **13**, 151-154, "Red-Shifted Excitation Mutants of the Green Fluorescent Protein."
13. Hastings, J. Woodland, (**1996**), *Gene*, **173**, 5-11, "Chemistries and Colors of Bioluminescent Reactions: a Review."
14. Cormack, B., Valdivia, R., and Falkow, S., (**1996**), *Gene*, **173**, 33-38, "FACS-Optimized Mutants of the Green Fluorescent Protein (GFP)."
15. K. Kaiden, T. Matsui, and S. Tanaka, (**1987**), *Appl. Spectrosc.*, 41, 180-184, "A Study of the Amide III Band by FT-IR Spectrometry of the Secondary Structure of Albumin, Myoglobin, and g-Globulin."
16. K. Kato, T. Matsui, and S. Tanaka, (**1987**), *Appl. Spectrosc.*, 41, 861-865, " Quantitative Estimation of a-Helix Coil Content in Bovine Serum Albumin by Fourier Transform-Infrared Spectroscopy."
17 T. P. Abbott, W.W. Wolf, Y. V. Ku, R.O. Butterfield, and R. Kleiman, (**1991**), *Appl. Spectrosc.*, 45, 1865-1673, "FT-IR Analysis of Jojoba Protein Conformations in D2O."
18. B.R. Singh and Fuller, M.P. (**1991**), *Appl. Spectrosc.*, 45, 1017-1021, "FT-IR in Combination with the Attenuated Total Reflectance Technique: A Very Sensitive Method for the Structural Analysis of Polypeptides".
19. Fu, F.-N., D.B. Oliveira, W.R. Trumble, H.K. Sakar, and B.R. Singh, (**1994**), *Appl. Spectrosc.*, 48, 1432-1441, "Secondary Structure Estimation of Proteins Using the Amide III Region of Fourier Transform Infrared Spectroscopy: Application to Analyze Calcium-Binding-Induced Structural Changes in Calsequestrin.".
20. G. Careri, E, Gratton, P.-H. Yang, and J.A. Rupley, (**1980**), *Nature*, 284, 572-573, "Correlation of IR Spectroscopic, Heat Capacity, Diamagnetic Susceptibility and Enzymatic Measurements on Lysozyme Powder."
21. Griebenow, K. and Klibanov, A. M., *J. Biol. Chem.* (**submitted 1995**), " Lyophylization-Induced Changes in the Secondary Structure of Proteins."

11.	**Project Title**	Biological Separation of Phosphate from Ore
	Principal Investigator	R. D. Rogers
	Project Site	Industrial Biotechnology and Process Engineering *at the* Idaho National Engineering and Environmental Laboratory

Description:

The project is aimed at bioseparation of insoluble phosphate complexes from ore matrix. To achieve this, selection and culturing of microorganisms capable of dissolving and concentrating insoluble phosphate forms have been made. The mechanism by which microorganisms dissolve the phosphate complexes are being investigated. Promising bioreactor configurations will be tested at the bench scale.

1996-1997 Accomplishments

- A new cost analysis on the production of a pelleted micro-encapsulated, biologically mediated, slow release phosphate fertilizer (dry pellet) was prepared. The study validated previous assumptions obtained from production of a water-based biobead that the concept is economically viable with attractive environmental benefits.
- Component parts of the pellet were selected and tested for comparability.
- Growth studies conducted to evaluated manufactured biopellets showed that plants response to the pellets was comparable to that of phosphate fertilizer.
- Preliminary agreements are being reviewed with phosphate fertilizer producers for the development and licensing of the biotechnology.

1998 Planned Activities:

Due to termination of funding, there is no activity planned for FY 1998.

Annual Technical Summary Report:

The importance of microbial solubilization of phosphate was demonstrated 45 years ago. Numerous studies have since been conducted on the effectiveness of microbes in solubilizing insoluble forms of phosphorous. The literature reports that 1-33% of the phosphate contained in rock phosphate (RP) can be released by microbial activity in batch tests. Using a naturally occurring microorganism selected from screening over 800 isolates, researchers at the Idaho National Engineering and Environmental Laboratory showed that a selected microorganism was capable of solubilizing over 85% of available RP. Enhancement was also maintained with increasing RP concentrations when the process was carried out under continuous conditions with a bench scale bioprocess unit.

Because of the positive potential of the proposed method a large-scale continuously stirred tank (CTR) bioreactor system was developed for the processing of RP. During the development of the CTR technology (which has had a patent issued) it became apparent that there were other potential applications of the bioprocess which could accomplish the same end goal of producing a "green" phosphate produce. Results of the testing suggest that the bioprocess could be applicable to the development of a slow release, pelletted fertilizer bead. The bead technology has already progressed to the development of a model process for production.

Based on this model, the envisioned application of the technology will be in devising a method for encapsulation of the component parts of the biosolubilization process (microorganism, carbon source, RP) into 2-3 mm beads for the production of slow release phosphate fertilizer products (patent pending). The incremental decrease in the cost of pollution management from use of this methodology alone would increase the profitability of the industry. It is expected that the U.S. industry would also be able to compete more effectively with an increasing supply of non U.S. product currently being sold in the country.

A first generation (FG) prototype bead consisting of a source of carbon, rock phosphate ore (RP), and the active bacterium (E37) has been manufactured and tested both in laboratory studies and with plants in a greenhouse setting. Encouraging results from this study have been provided in a preceding report. The FG product was manufactured, stored, and used under fully hydrated conditions. Experience with this prototype indicated that a dehydrated product would be more practicable for storage, handling, and application. Therefore, efforts has been directed towards the development of a second generation (SG) dehydrated bead.

Various studies conducted during the development of the SG product indicated that, under the right conditions, the dried components of the bead could be combined under pressure to form a pelleted product. Before manufacturing could begin it was necessary to develop methods for the production, storage, and blending of viable, dried E37 with other components. Various methods were also evaluated for pressure production of a desirable pellet.

Laboratory studies with the pelleted material showed that the pellet had superior storage and handling characteristics when compared to the FG bead. Solubilization of rock phosphate could be initiated after extended periods of storage. Data indicated that there was at least a five-fold extension in product shelf life. A plant study in an environmentally controlled growth chamber was conducted to determine the ability of the pelleted product to provide sufficient soluble phosphate to sustain the growth of barley in phosphate deficient soil. Results of this study showed that the pellet was capable of promoting plant growth comparable to plants provided soluble ortho phosphate. Refinements of production and testing were ongoing until the end of FY 1997.

12.	**Project Title**	Efficient Biocatalytic Synthesis of Polyesters from Bulk Monomers
	Principal Investigator	Alan J. Russell
	Project Site	University of Pittsburgh

Description:

Technology Vision 2020 for the U.S. chemical industry provides a vision and long-term strategy for the development of the Nation's chemical process industry. The document calls for the development of new catalysts and reaction systems for economical and environmentally safe processes. We propose to use expertise developed in our laboratories to further develop and understand an approach to biocatalytic poyesterification which shows remarkable promise. The large scale synthesis of a plastic by "green" technology could have an impact on energy utilization and the environment. Bulk polymerization is the standard method for polyester synthesis. The research thrust concerns enzyme catalyzed polymer synthesis in which there is no solvent. This bulk polymerization will mirror conventional synthesis but will eliminate the need for extremes of temperature and corrosive acid catalysts. This research is the first detailed investigation of rapid an efficient synthesis of polyesters from bulk monomers under ambient conditions with very low biocatalyst concentrations. This project builds on a collaborative project with Bayer Corporation in which we have developed, patented and licensed a method for the biocatalytic synthesis of short chain polyester polyols in supercritical fluids. That combination of reaction and product separation, from which the proposed research has grown, is an example of research that meets the needs expressed in Vision 2020 and the BCTR program.

1996-1997 Accomplishments:

New start in FY1997

1998 Planned Activities:

All efforts are designed for 15 months work.

Task 1: Benchmark enzyme/reactor configurations for bulk polyesterifications. This will involve use of enzymes to synthesize unique polymers from bulk monomers and maximizing polyester molecular weight during biocatalytic polyesterification of bulk monomers. This work will involve using thermostable lipases from Diversa, Inc. and design of a small scale bioreactor to address mass transfer problems.

Task 2: Investigate the vaious monomer end groups designed to shift equilibrium of polyesterification to synthesis. Use of monomers other than divinyl adipate (DVA) in irreversible transesterification will be evaluated to establish which end-groups are most effective to enhance rates, to establish a desired degree of polymerization, and to evaluate the effect of reaction by-products on the enzyme catalyst system.

Task 3: Maximization of achievable molecular weights. Evaluation of small scale bioreactor in task 1 related to reducing limitations to polymerization by diffusion (mass transfer) and to role of exothermicity (temperature) on molecular weight distribution of products. This effort will only be initiated and not completed within the 15 months of work.

Annual Technical Summary Report:

None. This is a new start which did not begin until late in FY 1997 or early FY 1998.

13. Project Title Over Production of Lignocellulosic Enzymes of *Coriolus versicolor* by Genetic Engineering Methodology

Principal Investigator(s) Arthur Williams

Project Site Howard University

Description:

The project seeks to understand the biological and chemical processes involved in the secretion of the enzyme polyphenol oxidase (PPO) by the hyphae, the basic unit of the filamentous fungus *Coriolus versicolor.* These studies are made to determine rational strategies for enhanced secretion of PPO, both with the use of recombinant DNA techniques (Howard University), and without (Clark-Atlanta). This effort is done in concert with work at Clark-Atlanta University but focuses on recombinant DNA techniques to enhance enzyme production.

The major thrust of this project was twofold: (1) to mass produce *C. versicolor* tyrosinase (polyphenol oxidase) by genetic engineering as well as cultural manipulations; (2) to use polyphenol oxidase (PPO) as a biocatalyst in the processing of lignocellulose as a renewable energy resource. Wood-decay within forests, a significant renewable photosynthetic energy resource, is caused primarily by *Basidiomycetous* fungi, i.e., white-rot fungi. In the case of the white-rot fungus *C. versicolor,* its degradative capacity results from the ability to elaborate extracellular cellulolytic and ligninolytic enzymes. With regard to the latter, at least one of the enzymes, PPO, appears within the external milieu in a highly time-dependent fashion when *C. versicolor* is cultured in a defined growth medium.

In this study, the assessment of genomic and CDNA recombinant clones with regard to the overproduction of PPO continued. During antibody (anti-tyrosinase) screening of genetic clones, enhanced production of polyphenol oxidase in several variants was observed. Further, immunoprecipitation and ELISA protocols were employed to detect and quantify tyrosinase (PPO) during cultural manipulations. Catechol and tyrosine appeared to stimulate the production of PPO in bulk cultures. On the other hand, diethylthiocarbarmte (DETC) exhibited an inhibitory effect on PPO Activity. In the attempt to over produce PPO, it was noticed that the activity of catechol oxidase (CTO) was higher than PPO. Summarily, owing to the fact that the tyrosinase complex possesses multiple activities, e.g., monophenol oxidase, PPO, and CTO, the employment of it as a biocatalyst to process lignocellulose appears feasible.

Further, immunocytochemical techniques were employed to assess the mechanism(s) involved in the secretion of PPO by the hyphae. That is, immunogold was used to label (i.e., tag) native PPO during certain growth conditions. Data from the immunocytochemical analysis suggested a putative route of secretion of PPO in *C. versicolor.* Also, factors influencing PPO secretion were examined.

<u>Background</u>

Wood Decay in the Forest – Renewable Energy Resource:

Wood-decay within Forests, a renewable photosynthetic energy source, is caused by *Basidiomycetous* fungi, e.g., white-rot fungi which possess the ability to degrade lignin, cellulose and hemicellulose, the main polymers of wood. In the case of at least one of these fungi, *C. versicolor,* this ability results from the fungus' capacity to elaborate both extra-cellular cellulolytic and ligninolytic enzymes, synthesis and secretion of these enzymes can occur *in situ* and *in vitro.* With regard to the latter, at least one of these enzymes, PPO, makes its appearance

within a culture medium in a highly time-dependent fashion. Presumably, extracellular PPO originates via secretion of intracellular PPO. However, this presumption requires verification. Because PPO, an enzyme capable of converting diphenols to diquinones and oligomerizing the lignin derivative, syringic acid, appears to be inducible, it is conceivable that the *C. versicolor* culture system could be a model for achieving over production of enzymes, a stated mission of the DOE project. In addition, the system, which can be scaled-up to industrial levels of enzyme production, seems to be the one in which the regulation of synthesis and/or secretion of both cellulolytic and ligninolytic enzymes can be investigated separately; i.e., experiments can be designed whose results may lead to the independent control of synthesis and/or secretion.

Lignocellulose – A Forest By-Product:

Besides its presence within wood, lignocellulose often constitutes an unwanted component in the paper-pulp industry and renders certain agricultural commodities less digestible to ruminants. Thus, an available supply of lignocellulolytic enzymes could be marked industrial value. To maintain an adequate supply, substrate induction coupled to hyphal "culture" and/or genetic engineering techniques could be employed to overproduce these enzymes.

The cellulases and ligninases that are secreted by *C. versicolor* can degrade cellulose and lignin, the main organic polymers of wood. *C. versicolor* can be grown in 'batch culture', which makes the over-production and the enhanced secretion of interest, commercially and biologically.

Objectives

1. Detect and quantify the expression of PPO in recombinant clones of *E. coli.*
2. Detect and quantify the overproduction of PPO during cultural manipulations.
3. Mutagenize cloned DNA inserts harboring the PPO gene.
4. Assess the expression of recombinant clones carrying mutagenized DNA inserts for their abilities to overproduce PPO.
5. Perfect the PPO system as a model for producing large scale quantities of industrial enzymes.
6. Characterize the catalytic properties and activities of the PPO complex.
7. Utilize PPO as a biocatalyst in wood bioprocessing, e.g., lignocellulose utilization.
8. Initiate the development of a course curriculum impacting upon the environment with heavy emphasis pertaining to biotechnology, manufacturing, and bioprocessing.

Methodology

Part 1. Genetic Engineering of the PPO Gene

1. Construction of PPO recombinant clones
2. Screening of recombinant clones of PPO
3. Trafficking of PPO in *E. coli* cells

Part 2. Biological Processing of the PPO Gene

1. Enzymatic catalysis of PPO.
2. Enzymatic assay of PPO.
3. Bioprocessing of PPO.

1995-1997 Accomplishments:

- Generated several PPO recombinant clones of *C. versicolor* by means of cDNA approach.
- Assessed the expression of the PPO gene in bacterial cells as well as a unique yeast variant.
- Quantified PPO production via antibody probing.
- Enhanced PPO production via cultural manipulations.
- Detected other enzymatic activities associated with the PPO complex.
- Initiated the utilization of PPO as a biocatalyst in the processing of lignocellulose.

Conducted Work in the Following Areas in 1996:

- Marketed the mass production of PPO via genetic engineering and cultural manipulations.
- Resolved other enzymatic activities associated with the PPO complex.
- Utilized multiple activities of PPO complex in the bioprocessing of lignocellulose.
- Demonstrated other industrial applications of PPO.
- Generated a patent of PPO proteins.

1998 Planned Activities:

Completed in FY96

Annual Technical Summary Report:

Results

During the period (8/17/95 - 8/16/96), the assessment of genomic and cDNA recombinant clones with regard to overproduction of PPO has continued (Figure 1). During antibody, anti-tyrosinase,

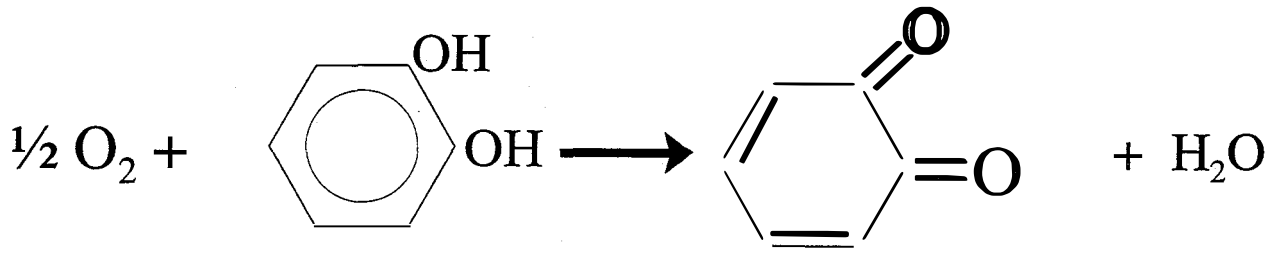


Figure 1 - Conversion of orthophenol to orthoquinone catalyzed by enzymes

screening of genetic clones, enhanced production of PPO in several variants was observed. That is, immunoprecipitation and ELISA protocols were employed to detect and quantify tyrosinase, i.e., PPO, during cultural manipulations. Cultural filtrates and tissue extracts were concentrated whereby standardized aliquots were tested in the ELISA protocol. On the other hand, corresponding aliquots were subjected to SDS-polyacrylanude gel electrophoresis subsequent to immunoprecipitation. Based upon these results, one major protein band was detected from the tissue extracts while no visible band was displayed from the filtrates during the 15-day growth period.

Other recombinant clones were generated by means of the cDNA approach (Figure 2). Specifically, two methods were employed to separate poly (A^+) mRNA from bulk RNA and poly A Tract System proved more efficient (Figure 3). An aliquot of the poly (A^+) sample was subjected to formaldehyde-agarose gel electrophoresis to confirm its presence. In order to determine whether the poly (A) mRNA was functional, samples were subjected to the BRL *in vitro* translation system and two protein bands, approximately 40-60 kDA were visible. These protein bands were transferred to nitrocellulose paper, using the Sartobolt apparatus and probed with fluorescein-PPO antibodies.

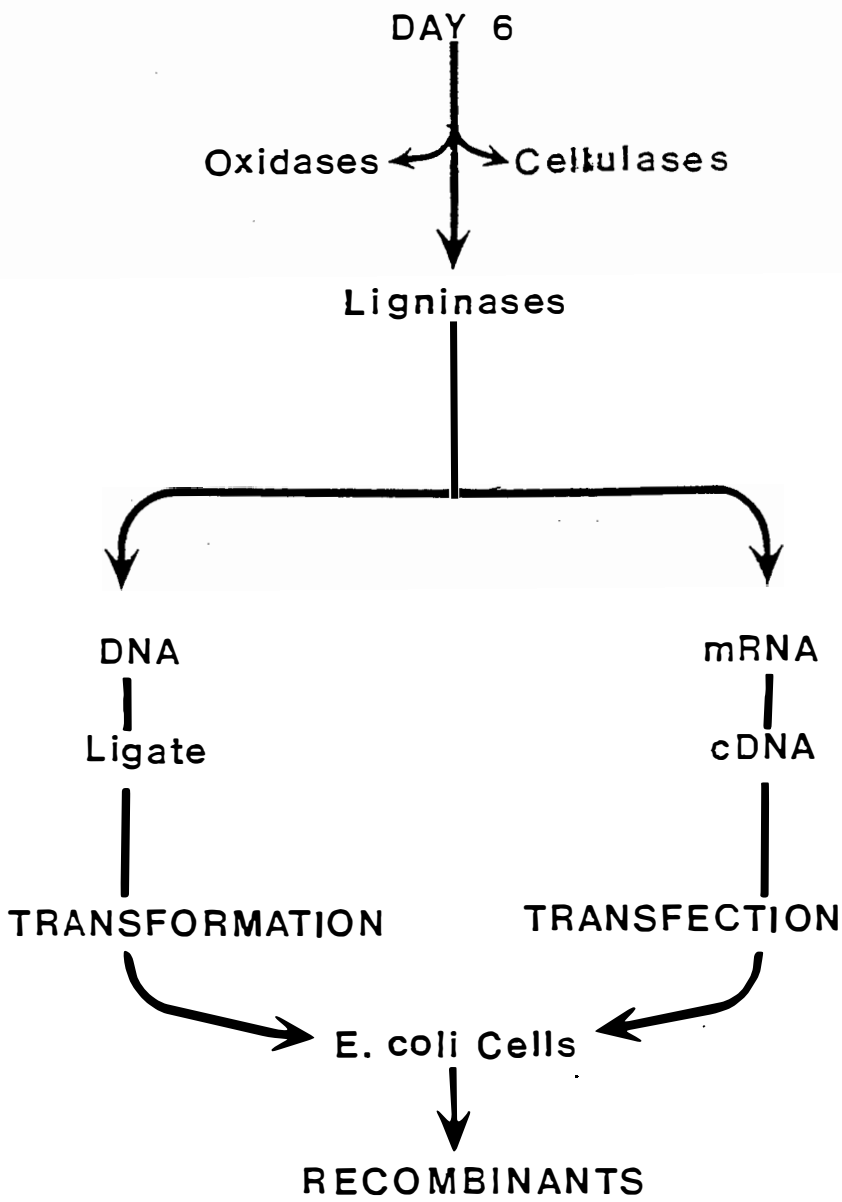


Figure 2. Plan of Study: A Molecular Approach.

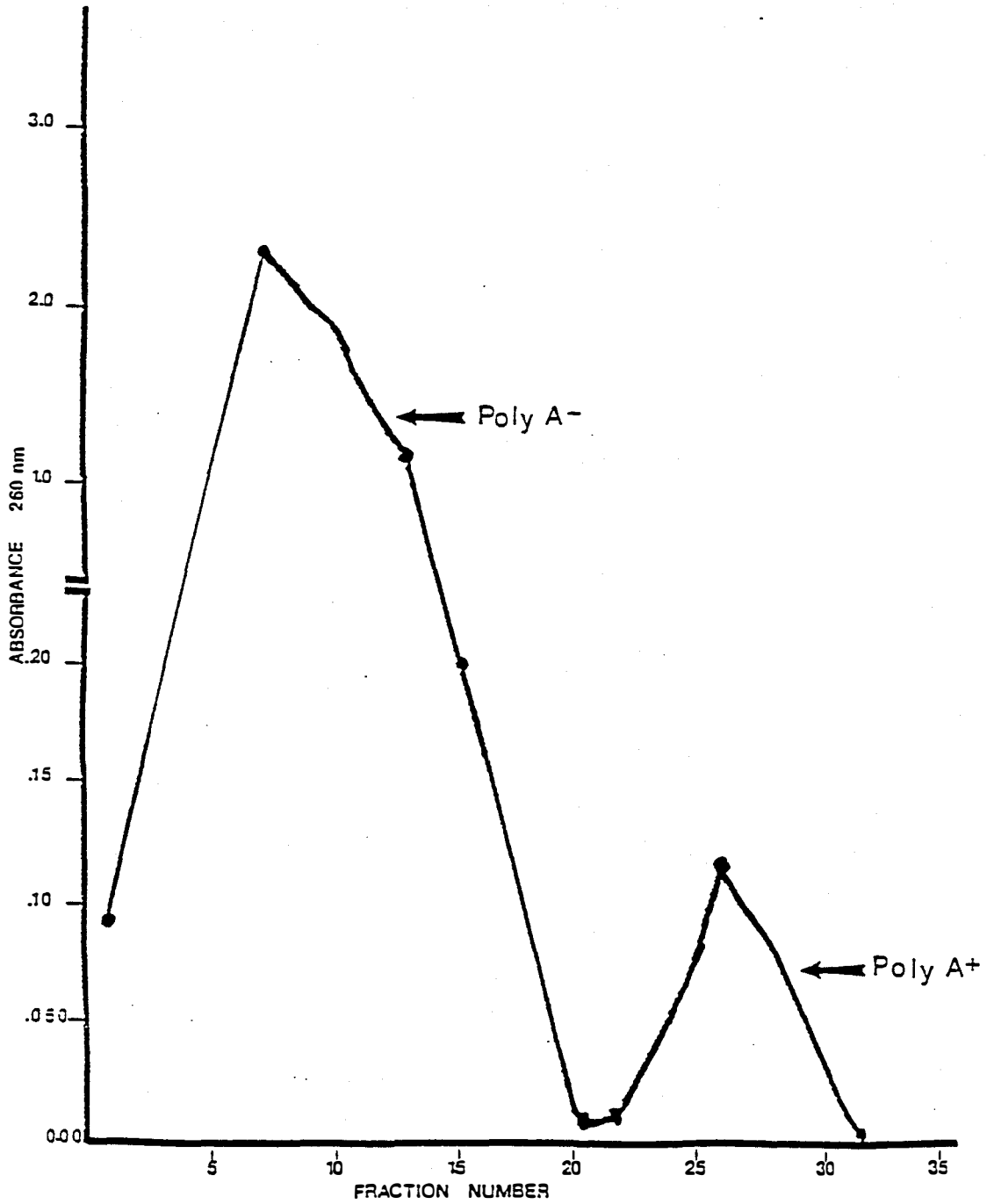


Figure 3. Fractionation of Purified RNA via Oligo (dT)-Cellulose Chromatography.

Additionally, the enzymatic profile of activities associated with the tyrosinase complex was determined. Catechol and tyrosine appeared to stimulate the production of PPO. The activity of CTO was even higher than PPO. Summarily owing to the fact that tyrosinase complex possesses multiple activities, e.g. monophenol, PPO, and CTO, its employment as a biocatalyst to process lignocellulose appears feasible from this study. This project has afforded now substantial understanding the PPO enzymatic complex.

Electron Microscopy: Immunocytochemical Analysis

PPO activity and production were studied by means of immunocytochemical methods. For electron microscopic immunocytochemistry, the mycelial mass was fixed in 3% glutaraldehyde and embedded in LR White medium. Subcellular localization of PPO with immunogold revealed that the 6th day mycelia showed maximum metabolic activity which was evidenced by the presence of (1) nuclei, (2) euchromatic material (3) rough endoplasmic reticulum; and (4) mitochondria. Vacuoles in the 6th day mycelial cells of *C. versicolor* were small and contained the enzyme polyphenol oxidase which was evidenced by the presence of immunogold particle in the vacuoles. There was an abundance of immunogold particles on the cell-wall and in the cytosolic area of the mycelial hyphae. The 9th day mycelia tissue displayed the maximum level of secretion of the enzyme PPO. This immunocytochemical study of PPO showed the intracellular storage of PPO and the distribution of PPO which was observed as the localization of immunogold particles conjugated to the secondary antibody (goat anti-rabbit IgG) which is specific to the primary antibody rabbit-antityrosinase in the vacuoles, secretory vesicles cell walls and the extracellular area. The overall putative route of secretion of PPO appears from (1) *E.R.* ⇒ (2) vacuoles ⇒ (3) secretory vesicles (adjacent to the interior of the cell wall) ⇒ (4) cell wall ⇒ (5) outside of cell wall. (see Figures 4, 5, and 6).

Significance of Biological Processing of PPO

The white-rot basidomycete, *C. versicolor,* secretes PPO as well as other lignocelluylolytic enzymes *in situ* and *in vitro.* The enzymes degrade cellulose and lignin, the principal polymers of wood. Wood is a component of trees which constitutes a renewable forest energy resource. Also, these enzymes may play a role in the recycling of waste from lignocellulose degradation.

To date, environmental biotechnology has a direct and fundamental role in bioremediation, biorestoration, waste treatment, environmental diagnostics and analysis, the preservation of biodiversity, and the development of environmental risk assessment technologies. Applications of biotechnology to energy production are likely to be an important part of our economic future by promoting production of fuels from renewable resources.

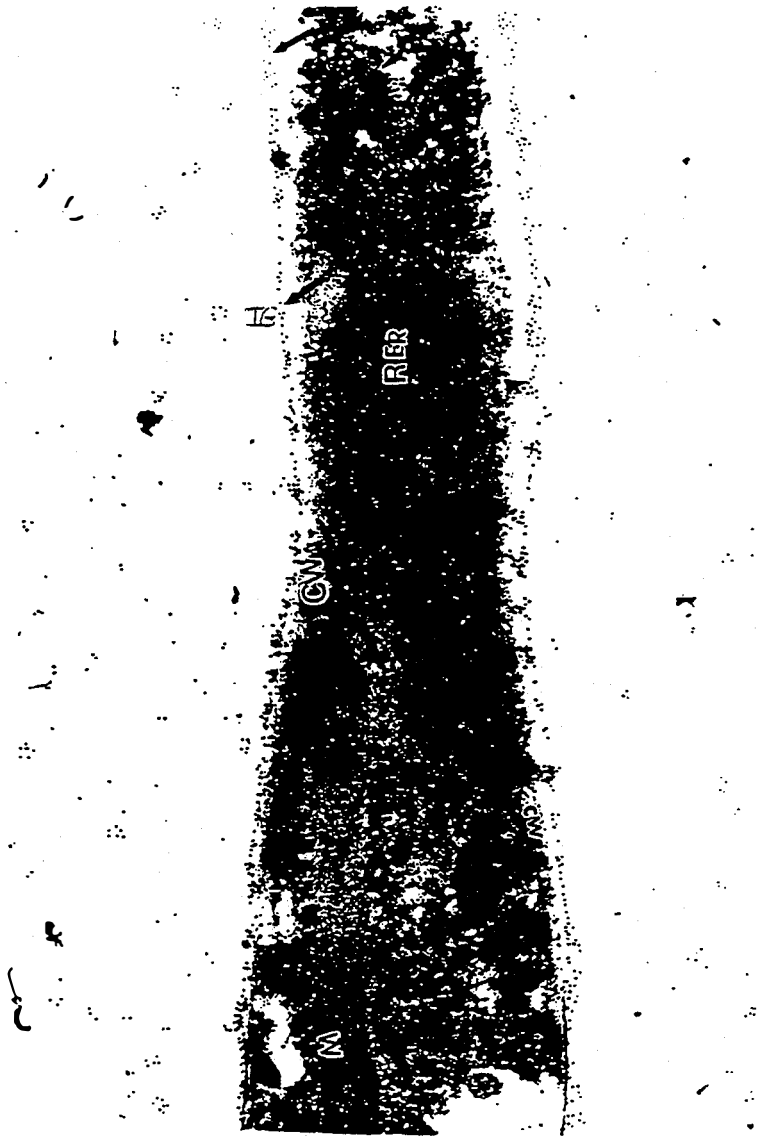


FIGURE 4

Sixth- day cell showing the presence of a large amount of immunogold particles all along the outside border of the cell-wall (CW) of a young hyphal cell of *Coriolus versicolor* from the sixth day. A large mitochodrion (M) is well-defined. A dense mass of rough endoplasmic reticulum (RER) traverses the entire cytoplasm. Presence of scattered immunogold particles (arrows) is well noted in the extracellular area. (×45,000)

Figure 5 - Sixth day hyphal cell (experimental) showing the abundance of immunogold particles of the cell wall (CW) area. A large amount of rough endoplasmic reticulum (RER) is scattered throughout the cytoplasm of the cell. A large amount of immunogold particles (arrow) are present in vacuoles and outside the cell. Presence of a nucleus is well defined (x33,000).

Figure 6 - Hyphal cell of *C. versicolor*, 9th day mycelial cell (experimental set) showing the presence of a dense mass of rough endoplasmic reticulum (RER). Note the deposition of immunogold particles (arrow heads) along the cell wall (CW), the vacuoles (V), and in the secretory vesicle (SV) (x27,000).

6.0 Program Output

6.1 Publications in Professional Journals, Books and Symposium Chapters, and DOE Technical Publications (alphabetical by institution)

California Institute of Technology

1. Zhao, H. and F. H. Arnold "Functional and Non-functional Mutations Distinguished by Random Recombination of Homologous Genes," Proc. Natl. Acad. Sci. *USA*, **94**, 7997-8000 (1997).

2. Zhao, H. and F. H. Arnold "Optimization of DNA Shuffling for High Fidelity Recombination," *Nucleic Acids Research.*, **25**, 1307-1308 (1997).

3. Moore, J.C., H.-M. Jin, O. Kuchner and F. H. Arnold "Strategies for the *in vitro* Evolution of Protein Function: Enzyme Evolution by Random Recombination of Improved Sequences," *J. Molecular Biology* **272**, 336-347 (1997).

4. Arnold, F.H., "Directed Evolution: Creating Biocatalysts for the Future," *Chemical Engineering Science* **51**, 5091-5102 (1996).

5. Zhao, H., L. You and F. H. Arnold, "Strategy for the Directed Evolution of a Peptide Ligase," Proceedings of the 13th Enzyme Engineering Conference, *Annals of the N.Y. Acad. Sci.* **799**, 1-6 (1996).

6. Shao, Z. and F. H. Arnold, "Engineering New Functions and Altering Existing Functions," Current Opinion in Structural Biology **6**, 513-518 (1996).

7. Arnold, F.H. and Moore, J.C. "Optimizing Industrial Enzymes by Directed Evolution," *Advances in Biochemical Engineering* **58**, 1-14 (1997).

8. Zhao, H. and F. H. Arnold,"Optimization of DNA Shuffling for High Fidelity Recombination," *Nucleic Acids Research.* **25**, 1307-1308 (1997).

9. Zhao, H. and F. H. Arnold, "Functional and Non-functional Mutations Distinguished by Random Recombination of Homologous Genes," Proc. Natl. Acad. Sci. *USA* **94**, 7997-8000 (1997).

10. Zhao, H. and F. H. Arnold, "Combinatorial Protein Design: Strategies for Screening Protein Libraries," *Current Opinion in Structural Biology* **7**, 480-485 (1997).

11. Kuchner, O. and F. H. Arnold, "Directed Evolution of Enzyme Catalysts," invited review for *Trends in Biotechnology*, in press.

12. Jiang-SY, Frazier-R, Yamaguchi-ES, Blanco-M, Dasgupta-S, Zhou-YH, Cagin-T, Tang-YC and Goddard-WA, "The SAM model for wear inhibitor performance of dithiophosphates on iron-oxide," *Journal of Physical Chemistry B*, **101**, 39, 7702-7709 (1997).

13. Brameld, K.A., Dasgupta, S., Goddard, W.A. III. "Distance Dependent Hydrogen Bond Potentials for Nucleic Acid Base Pairs from *ab initio* Quantum Mechanical Calculations (LMP2/cc-pVTZ)." *J. Phys. Chem B.*, **101**, 4851-4859 (1997).

14. Jiang-S.Y., Dasgupta-S., Blanco-M., Frazier-R., Yamaguchi-E.S., Tang-Y.C., Goddard-W.A., "Structures, Vibrations, and Force-fields of Dithiophosphate Wear Inhibitors from Ab-initio Quantum-chemistry," *Journal of Physical Chemistry*, **100**, 39, 15760-15769 (1996).

15. Blanco, M., Y. Tang, P. Shuler, and W.A. Goddard III, "Activated Complex Theory of Barite Scale Control Process" *Molecular Engineering*, in press

16. Ramachandran-S, Tsai-BL, Blanco-M, Chen-H, Tang-YC, Goddard-WA, "Atomistic simulations of oleic imidazolines bound to ferric clusters," *J. of Physical Chemistry* A, **101**,1,83-89 (1997).

17. Ramachandran-S, Tsai-BL, Blanco-M, Chen-H, Tang-YC, Goddard-WA, "Self-assembled monolayer mechanism for corrosion inhibition of iron by imidazolines," *LANGMUIR*, **12**, No. 26, 6419-6428 (1996).

Lawrence Berkeley National Laboratory

18. Kilduff, J. E and King, C.J. "Effect of Carbon Adsorbent Surface Properties on the Uptake and Solvent Regeneration of Phenol", *Ind. Eng. Chem. Research,* 36, 1603-1613 (1997).

19. Broekhuis, R.R., S. Lynn and C. J. King, "Recovery of Propylene Glycol from Dilute Aqueous Solutions by Complexation with Organoboronates in Ion-Pair Extraction," *Ind. Eng. Chem. Research*, **35**, 1206-1214 (1996).

20. Dai, Y. and C. J. King, "Selectivity between Lactic Acid and Glucose during Recovery of Lactic Acid with Basic Extractants and Polymeric Sorbents," *Ind. Eng. Chem. Research*, **35**, 1215-1224 (1996).

Massachusetts Institute of Technology

21. Xu, K. and A.M. Klibanov, "pH Control of the catalytic activity of cross-linked enzyme crystals in organic solvents". *J. Am. Chem. Soc.* **118**, 9815-9819 (1996).

22. Klibanov, A.M., "Why are enzymes less active in organic solvents than in water?". *Trends Biotechnol.* **15**, 97-101 (1997).

23. Griebenow, K. and A.M. Klibanov, "On protein denaturation in aqueous-organic mixtures but not in pure organic solvents". *J. Am. Chem. Soc.* **118**, 11695-11700 (1996).

24. Wescott, C.R. and A.M. Klibanov, "Thermodynamic analysis of the solvent effect on substrate specificity of lyophilized enzymes suspended in organic media". *Biotechnol. Bioeng.* **56**, 340-344 (1997).

25. Xu, K., K. Griebenow, and A.M. Klibanov, "Correlation between catalytic activity and secondary structure of subtilisin dissolved in organic solvents". *Biotechnol. Bioeng.* **56**, in press (1997).

University of Pittsburgh

26. Skourtis, S.S. and D.N. Beratan, "High and low resolution theories of protein electron transfer," *J. Bioinorg. Chem.*, **2**, 378-386 (1997).

27. Beratan, D.N., S. Priyadarshy, and S.M. Risser, "DNA: wire or insulator?" *Chemistry and Biology*, **4**, 3-8 (1997).

28. Onuchic, J.N., S.M. Risser, S.S. Skourtis, D.N. Beratan, "The design of starburst dendrimer electron transfer systems," in Molecular Electronics, M.A. Ratner and J.Jortner, editors, *Chemistry for the 21st Century Series*, Blackwell Science, Oxford, 1997.

29. Wolfgang, J., S.M. Risser, S. Priyadarshy, D.N. Beratan, "Secondary structure conformations and long-range electronic interactions in oligopeptides, *J. Phys. Chem.* B, **101**, 2987-2991 (1997).

30. Zusman, L.D. and D.N. Beratan, "A three-state model for two-electron transfer reactions," *J.Phys. Chem.* A, **101**, 4136-4141 (1997).

31. Priyadarshy, S., S.M. Risser, and, D.N. Beratan, "DNA is not a molecular wire: protein-like electron transfer in an extended pi-electron system, *J. Phys. Chem.*, **100**, 17678-17682 (1996).

32. Priyadarshy, S., S.M. Risser, and, D.N. Beratan, "DNA double helix mediated long range electron transfer," *Int. J. Quantum Chem. Quantum Biol. Symp.*, **60**, 65 -71 (1996).

33. Kuhn, C. and D.N. Beratan, "Inverse strategies for molecular design," *J. Phys. Chem.*, **100**, 10595-10599 (1996).

34. Zusman, L.D. and D.N. Beratan, "Two-electron transfer reactions in polar solvents," *J. Chem. Phys.*, **105**, 165-176 (1996).

35. Kurnikov, I.V. and D.N. Beratan, "Ab initio based effective hamiltonians for long range electron transfer: Hartree-Fock analysis," *J. Chem. Phys.*, **105**, 9561-9573 (1996).

36. Priyadarshy, S., S.S. Skourtis, S.M. Risser, and D.N. Beratan, "Bridge mediated electronic interactions: differences between Hamiltonian and Green function partitioning in a non-orthogonal basis," *J. Chem. Phys.*, **104,** 9473-9481 (1996).

37. Beratan, D.N. and J.N. Onuchic, "Chapter 2: The protein bridge between redox centers," in Protein Electron Transfer, D.S. Bendall, editor, Bios Scientific Publishers, Oxford, pp 23-42 (1996).

38. Onuchic, J.N. and D.N. Beratan, "Appendix A: Electron transfer rate calculations," in Protein Electron Transfer, D.S. Bendall, editor, Bios Scientific Publishers, Oxford, pp 273-283 (1996).

Recent book chapters acknowledging support of this project:

D.N. Beratan and J.N. Onuchic, "Chapter 2: The protein bridge between redox centers," in *Protein Electron Transfer*, D. Bendall, editor, Bios Scientific Publishers, Oxford, pp 23-42.

J.N. Onuchic and D.N. Beratan, "Appendix A: Electron transfer rate calculations," in *Protein Electron Transfer*, D. Bendall, editor, Bios Scientific Publishers, Oxford, pp 273-283.

Oak Ridge National Laboratory

39. Barton, J. W., E. K. Reed, B. H. Davison. Gas-phase Enzyme Catalysis Using Immobilized Lipase for Ester Production. *Biotechnology Techniques* **11**:747-750 (1997).

40. Davison, B. H. J. W. Barton, G. R. Petersen. Nomenclature and Methodology For Classification of Nontraditional Biocatalysis. *Biotechnology Progress*. 13:512-518 (1997).

41. Boron, D., G. R. Petersen and B. H. Davison. Biotechnology for National Industrial Needs, Part II: Development of Biotechnology Capabilities in Extreme Environments for the Chemical and Related Products Industry.

42. Barton, J.W., K. T. Klasson, L. J. Koran, Jr., and B. H. Davison. "Microbial Removal of Alkanes from Dilute Gaseous Waste Streams: Kinetics and Mass Transfer Considerations," *Biotechnol. Prog.* In press (1997).

43. Barton, J.W., S. M. Hartz, K. T. Klasson, and B. H. Davison. "Microbial Removal of Alkanes from Dilute Gaseous Waste Streams: Mathematical Modeling of Advanced Bioreactors Systems", *J. Chem. Technol. Biotechnol.* in press (1997).

6.2 Patent Awards/Filings, and Licenses

Filed

A random-priming technique for *in vitro* mutation and recombination of polynucleotide sequences) Ser No. 60/045,211

A versatile 'defined primer' approach for in *vitro* recombination of polynucleotide sequences, Ser. No. 60/046,256.

Recombination of polynucleotide sequences using random or defined primers, Utility application filed 8/97.

S.M. Husson & C. J. King, "Regeneration of Carboxylic-Acid-Laden Basic Sorbents by Leaching with a Volatile Base in an Organic Solvent".

Licenses

Caltech completed joint research projects with Recombinant BioCatalysis, Inc. and Thermogen, Inc. with agreements for licensing resulting technologies.

Significant Achievements:

Award:

C.J. King, Award in Separations Science and Technology, American Chemical Society, April 1997.

Special Presentations:

Frances H. Arnold

"Directed Evolution of a para-Nitrobenzyl Esterase," at "Design of Protein Structure and Function" Conference, Royal Swedish Academy of Sciences, Stockholm, Sweden, May 22-25, 1996.

Invited speaker for 50th Anniversary Celebration of the Office of Naval Research, at the National Academy of Sciences, May, 1996.