Broadly defined, biotechnology includes any technique that uses living organisms to make products, improve plants or animals, or to develop microorganisms for specific purposes. Biotechnology emerged as a modern science with the discovery of deoxyribonucleic acid (DNA) in the 1950s. In the 1970s scientists discovered one of the basic tools of biotechnology—recombinant DNA technology. During the 1990s technical innovation in this field occurred at an exponential rate. New biological techniques based on recombinant DNA technology are now being applied to R&D in pharmaceuticals, agriculture, and the chemical industry.

New biotechnology was first proven to be a source of innovation in pharmaceutical R&D, contributing as both a production technology and a research tool. Currently, both dedicated biotechnology companies and established multinational pharmaceutical companies are using the tools and techniques of biotechnology in their drug discovery and development efforts.

Now the platform technologies seen as the elements of a paradigmatic change in pharmaceutical R&D are being applied to biotechnology product and process development in other industrial sectors. Four important platform technologies that are finding application outside the pharmaceutical industry are combinatorial chemistry, high-throughput screening, genomics, and bioinformatics.

In this report, PEP reviews these four important platform technologies and some of their applications in agriculture and the chemical industry. We also cite over 150 related patents and describe the more than 100 companies primarily active in these technologies. The market status throughout the world is summarized for three industrial sectors—pharmaceuticals, agriculture, and chemicals—where biotechnology has been commercialized. An overview of development costs in the pharmaceutical industry as a whole is provided in the report along with a more detailed review of the typical costs to develop a pharmaceutical product.

This report will be useful for anyone involved in R&D to better understand the important research tools being developed through the biotechnology industry. Additionally, it provides insights into the cost of typical R&D activities and also describes the direction that the technology has taken in life sciences applications.
CONTENTS

1 INTRODUCTION .......................................................................................................1-1

2 SUMMARY ................................................................................................................2-1
   PHARMACEUTICALS ..............................................................................................2-2
   AGRICULTURE .................................................................................................2-2
   CHEMICALS ...................................................................................................2-3

   PLATFORM TECHNOLOGIES ...........................................................................2-4
       Combinatorial Chemistry ........................................................................2-4
       High-Throughput Screening ...................................................................2-5
       Genomics ..............................................................................................2-6
       Bioinformatics .....................................................................................2-9

3 INDUSTRY STATUS ..............................................................................................3-1
   PHARMACEUTICALS ......................................................................................3-5
   AGRICULTURAL ............................................................................................3-7
   CHEMICALS ...............................................................................................3-12

4 RESEARCH AND DEVELOPMENT COSTS ...............................................................4-1
   INDUSTRY STATISTICS ................................................................................4-1
   EMPLOYEE COSTS ......................................................................................4-2
   COMPANY STATISTICS ..............................................................................4-7
   R&D PRODUCTIVITY ..................................................................................4-10
       Patent Activity ........................................................................................4-10
       New Chemical Entities ........................................................................4-13
       Knowledge Capital ...............................................................................4-14

5 RESEARCH AND DEVELOPMENT COST ELEMENTS .............................................5-1
   DRUG DISCOVERY AND DEVELOPMENT .......................................................5-1
   DISCOVERY AND DEVELOPMENT TIME ......................................................5-4
   DISCOVERY AND DEVELOPMENT COSTS ..................................................5-6
   REGULATORY RELIEF ..............................................................................5-8
   CROP PROTECTION ..................................................................................5-10
   ENZYMES .................................................................................................5-12

6 COMBINATORIAL CHEMISTRY AND HIGH-THROUGHPUT SCREENING ..........6-1
   INTRODUCTION ..........................................................................................6-1
   COMBINATORIAL CHEMISTRY SYNTHESIS STRATEGIES .........................6-2
       Solid-Phase Synthesis ...........................................................................6-4
       Deconvolution and Encoding ................................................................6-8
## CONTENTS (continued)

- Spatially Addressed Arrays ................................................................. 6-11
- Solution-Phase Combinatorial Chemistry ........................................... 6-11
- Automation of the Combinatorial Laboratory ...................................... 6-14
- High-Throughput Screening ................................................................. 6-16

### NONPHARMACEUTICAL APPLICATIONS OF COMBINATORIAL METHODS .......................... 6-19
- Materials ............................................................................................ 6-20
- Catalyst Discovery and Optimization .................................................... 6-21
- Reaction Optimization ................................................................. 6-23

### COMMERCIAL VENTURES IN COMBINATORIAL CHEMISTRY ................................ 6-24

### CONCLUSIONS ..................................................................................... 6-25

### GENOMICS ............................................................................................... 7-1
- STRUCTURAL GENOMICS ................................................................. 7-1
  - Sequencing ....................................................................................... 7-1
  - Mapping ......................................................................................... 7-5
- FUNCTIONAL GENOMICS ............................................................... 7-6
- APPLICATION GENOMICS .............................................................. 7-8
  - Plant Genomics .............................................................................. 7-9
  - Industrial Genomics ................................................................. 7-20

### SOCIAL, REGULATORY, AND LEGAL ISSUES .................................................. 7-21
- Social Issues ...................................................................................... 7-21
- Regulatory Issues ............................................................................. 7-23
- Legal Issues ...................................................................................... 7-23

### BIOINFORMATICS .................................................................................... 8-1
- BIOINFORMATICS: FROM DATA TO KNOWLEDGE DISCOVERY .................. 8-1
  - Bioinformatics and the Genomics Industry ...................................... 8-2
  - Industry Overview ........................................................................ 8-3
  - Commercial Ventures .................................................................... 8-4
  - Industry Challenges ...................................................................... 8-7
  - Future of Bioinformatics in Life Sciences Research .......................... 8-9

### DATABASES AND TOOLS ......................................................................... 8-9
- Databases and Tools: Organization .................................................. 8-10
- Proposed Standards for Distributed Computing .................................. 8-11
- The Common Object Request Broker Architecture (CORBA) ........... 8-11
- Extensible Markup Language (XML) .............................................. 8-12
CONTENTS (concluded)

Bioinformatics Software .......................................................................................... 8-12
THE KNOWLEDGE DISCOVERY PROCESS ................................................................. 8-16
  Data Retrieval and Integration ............................................................................. 8-16
  Knowledge Representation in Higher-Order Models ................................................. 8-17
APPLICATION OF BIOINFORMATICS ....................................................................... 8-19
  Pharmaceutical Applications of Bioinformatics ....................................................... 8-20
    Informatics and the Pharmaceutical Research Process .......................................... 8-20
    Pharmaceutical Research Bioinformatics Strategies ............................................... 8-21
    The Bioinformatics Impact on Pharmaceutical Research ...................................... 8-23
    The Future of the Drug Discovery Process ........................................................... 8-27
  Nonpharmaceutical Applications of Bioinformatics .................................................. 8-27
    Informatics and the Agricultural Research Process ................................................ 8-27
    Impact of Bioinformatics on Agricultural Research ............................................... 8-29
    Impact of Informatics on Microbial Research ....................................................... 8-29
IMPLICATIONS FOR THE LIFE SCIENCES ................................................................ 8-30

APPENDIXES
  A: PATENT SUMMARY TABLE
  B: COMPANY FINANCIAL DATA
  C: PATENT HISTORY OF PHARMACEUTICAL COMPANIES
  D: COMBINATORIAL CHEMISTRY/HIGH-THROUGHPUT SCREENING VENTURES
  E: BIOINFORMATICS COMPANIES
  F: SEQUENCE AND NONSEQUENCE KNOWLEDGE BASES
  G: CITED REFERENCES
  H: PATENT REFERENCES BY COMPANY
FIGURES

2.1 GROWTH OF GENBANK ........................................................................................................... 2-1
2.2 THE DNA DOUBLE HELIX ................................................................................................. 2-7
3.1 THE DNA DOUBLE HELIX ................................................................................................. 3-2
3.2 GROWTH OF GENBANK ................................................................................................. 3-5
3.3 BIOPHARMACEUTICALS APPROVED BY THE FDA ..................................................... 3-6
3.4 DRUG DISCOVERY AND DEVELOPMENT SUCCESS RATES ........................................ 3-7
3.5 SALES OF BIOTECH-MODIFIED CROPS ......................................................................... 3-9
4.1 R&D PROGRAM SIZE DISTRIBUTION IN INDUSTRY ..................................................... 4-2
4.2 HISTORICAL COST OF R&D AS A PERCENT OF SALES ................................................ 4-3
4.3 EMPLOYMENT TREND FOR R&D PROFESSIONALS ...................................................... 4-4
4.4 R&D PROFESSIONAL EMPLOYMENT AS A FRACTION OF OVERALL EMPLOYMENT .... 4-4
4.5 HISTORICAL R&D SPENDING PER R&D PROFESSIONAL ........................................... 4-6
4.6 R&D COST TREND FOR U.S. AND EUROPEAN PHARMACEUTICAL COMPANIES ........ 4-8
4.7 PATENT ACTIVITY FOR U.S. AND NON-U.S. CORPORATIONS ..................................... 4-11
4.8 DRUG PATENT ACTIVITY FOR TOP PHARMACEUTICAL COMPANIES ....................... 4-11
4.9 TOTAL PATENT HISTORY FOR TOP PHARMACEUTICAL COMPANIES ....................... 4-12
4.10 R&D COSTS ON PER-DRUG PATENT BASIS FOR TOP U.S. PHARMACEUTICALCOMPANIES ........................................................................................................... 4-12
4.11 R&D COSTS ON TOTAL PATENT BASIS FOR TOP U.S. PHARMACEUTICALCOMPANIES ........................................................................................................... 4-13
5.1 MOLECULAR TARGETS USED IN DRUG DISCOVERY ...................................................... 5-2
5.2 AVERAGE TIME FROM FIRST PHARMACOLOGICAL TESTING TO NEW DRUGAPPROVAL ............................................................................................................. 5-4
5.3 TYPICAL DRUG DEVELOPMENT TIMELINE .................................................................... 5-6
5.4 BREAKDOWN OF R&D COSTS BY FUNCTION FOR SELECTED YEARS ......................... 5-7
5.5 GROWTH IN OVERSEAS CLINICAL STUDIES BY U.S.-BASED PHARMACEUTICAL FIRMS ........................................................................................................... 5-9
6.1 COMBINATORIAL SYNTHESIS STRATEGIES .................................................................... 6-3
6.2 METHOD FOR PRODUCING COMBINATORIAL LIBRARIES OF SMALL ORGANICMOLECULES ............................................................................................................. 6-5
6.4 DECONVOLUTION TECHNIQUES ....................................................................................... 6-10
6.5 ENCODING TECHNOLOGIES ......................................................................................... 6-12
7.1 THE DNA DOUBLE HELIX ................................................................................................. 7-2
7.2 GROWTH OF GENBANK ................................................................................................. 7-5
7.4 GENETIC ENGINEERING IN PLANTS .............................................................................. 7-10
FIGURES (concluded)

8.1 BIOINFORMATICS PARADIGM...................................................................................... 8-2
8.2 IMPROVEMENT IN RESEARCH METHODOLOGY ........................................................... 8-23
8.3 EMERGING MODEL OF THE NEW DRUG DISCOVERY PROCESS................................. 8-24
8.4 BIOINFORMATICS AND THE PHARMACEUTICAL RESEARCH PROCESS.................... 8-25
8.5 THE IMPACT OF BIOINFORMATICS ON THE DRUG DISCOVERY PROCESS.............. 8-26
TABLES

3.1 MAJOR EVENTS IN THE COMMERCIALIZATION OF BIOTECHNOLOGY ...........................................3-3
3.2 FIFTEEN LEADING BIOTECHNOLOGY DRUGS ...........................................................................3-8
3.3 MAJOR GLOBAL HERBICIDE-TOLERANT CROPS .....................................................................3-10
3.4 ALLIANCES AND ACQUISITIONS AMONG AGRICULTURAL CHEMICAL COMPANIES ........3-11
3.5 SELECTED ENZYME APPLICATIONS AND PROCESS BENEFITS .............................................3-13
4.1 COST OF R&D AS A PERCENT OF SALES ..................................................................................4-3
4.2 R&D STAFF COSTS ....................................................................................................................4-6
4.3 TOP LIFE SCIENCE COMPANIES ..............................................................................................4-7
4.4 GROWTH IN BIOTECHNOLOGY PATENTS FOR TOP 10 CROP-PROTECTION COMPANIES ....4-9
4.5 CHEMICAL COMPANY VS. PHARMACEUTICAL COMPANY R&D PRODUCTIVITY ...............4-15
5.1 DURATION OF TOXICITY STUDIES TO SUPPORT CLINICAL TRIALS ..................................5-3
5.2 CLINICAL TRIAL PHASES ........................................................................................................5-3
5.3 U.S. PHARMACEUTICAL R&D COSTS BY FUNCTION .............................................................5-8
5.4 TESTS REQUIRED TO OBTAIN EPA REGISTRATION FOR PESTICIDE COMMERCIALIZATION ..........................................................5-12
6.1 POSSIBLE COMPOUNDS WITH DIFFERENT OLIGOMERS .....................................................6-4
6.2 REPRESENTATIVE FRITTED VESSEL SYSTEMS .......................................................................6-7
6.3 COMBINATORIAL BIOCATALYSIS REACTIONS .......................................................................6-14
6.4 FLUORESCENT DETECTION TECHNOLOGIES .......................................................................6-18
6.5 MAJOR SYMYX COLLABORATIONS .......................................................................................6-21
6.6 SCREENING TECHNOLOGIES PATENTED BY SYMYX ........................................................6-22
6.7 COMBINATORIAL COLLABORATIONS .....................................................................................6-23
6.8 FINANCIAL CHARACTERISTICS OF COMBINATORIAL CHEMISTRY ALLIANCES ..........6-25
7.1 GENE-SEQUENCING METHODS ...............................................................................................7-4
7.2 MODEL ORGANISM DATABASES ..............................................................................................7-9
7.3 MAJOR AGROCHEMICAL COMPANY ALLIANCES .................................................................7-11
7.4 IMPORTANT COMMERCIAL TRANSGENIC CROPS AND TRAITS ........................................7-15
7.5 AGRICULTURE AND BIOTECHNOLOGY PATENTS ................................................................7-16
7.7 BIOCHIP COMPANIES .............................................................................................................7-20
8.1 BIOINFORMATICS NEEDS IDENTIFICATION ............................................................................8-4
8.2 PATENTS IN BIOINFORMATICS ...............................................................................................8-6
8.3 WEB SITES MAINTAINING LISTS OF DATABASES ..................................................................8-10
8.4 DATABASES AND LISTS OF BIOINFORMATICS SOFTWARE ..................................................8-11
8.5 GENE-SEQUENCING AND SEQUENCE-ANALYSIS SOFTWARE COMPANIES ..........8-15
8.6 SPECIALIZED SEQUENCE AND NONSEQUENCE KNOWLEDGE BASES ..................8-18
8.7 GENOME SIZE COMPARISONS OF ECONOMICALLY IMPORTANT PLANTS TO ARABIDOPSIS AND RICE .................................................................8-28