BIOTECHNOLOGY
The Renewable Frontier

Edited by Daniel E. Koshland, Jr.

The American Association for the Advancement of Science

Biotechnology
The Renewable Frontier

Edited by Daniel E. Koshland, Jr.

The American Association for the Advancement of Science

Library of Congress Cataloging-in-Publication Data
Biotechnology: The Renewable Frontier
Includes index.
1. Biotechnology. 2. Biology—Research.
I. Koshland, Daniel.
TP248.2.B546 1986 660'.6 86-3378
ISBN 0-87168-314-8
ISBN 0-87168-283-4 (pbk.)

This material originally appeared in ***Science***, the official journal of the American Association for the Advancement of Science.

Distributed to bookstores and overseas through Westview Press, Boulder, Colorado 80301

AAAS Publication No. 85-26

Other titles in this series of ***Science*** volumes include:
Biotechnology & Biological Frontiers
Edited by Philip H. Abelson

Neuroscience
Edited by Philip H. Abelson, Eleanore Butz, and Solomon H. Snyder

Astronomy & Astrophysics
Edited by Morton S. Roberts

AIDS: Papers from ***Science***, 1982–1985
Edited by Ruth Kulstad

Frontiers in the Chemical Sciences
Edited by William Spindel and Robert M. Simon

Biotechnology
The Renewable Frontier

Introduction

In Pursuit of the Renewable Frontier

Daniel E. Koshland, Jr.

The endearing feature of intellectual frontiers is that they are in endless supply. Explorers of continents fight their way through wildernesses until they arrive at the water's edge, and then sigh that there are no new mountains to conquer. Researchers, on the other hand, are part of an ever-expanding universe, inevitably creating new territories to explore as they complete the maps begun by the discoveries of the past.

No area of research illustrates this phenomenon more clearly than modern biology, which many believe is in its Golden Age. This book presents a collection of articles that contain illustrative—but certainly not exhaustive—areas that have great potential for the future.

The articles are grouped together under a number of sections, but simply organizing the articles was a revelation. Almost any of these papers could have been listed in another section equally well. For example, Sherie Morrison describes chimeric antibodies which could be listed under "New Techniques" or under "Immunology;" Smith, Duncan, and Moir's article could be listed under "Biotechnology" or under "New Techniques;" Goodenow, Vogel, and Linsk's article could be under "Immunology" or "Developmental Biology and Cancer." The list could go on, but it illustrates the extraordinary cross-disciplinary aspects of modern biology. Discoveries in one area almost inevitably generate ideas or techniques applicable to a widely different area.

The same interdisciplinary phenomenon applies to the line between a new biological frontier and biotechnology. We have listed two articles in the latter category because they deal directly with industry, but in effect all the articles could be listed under that one heading. "Safety Concerns and Genetic Engineering in Agriculture" is both a frontier in biology and of key interest in biotechnology. "Genetic Engineering of Novel Genomes of

Large DNA Viruses" is of great theoretical interest and involves new techniques, but it also may have vast practical applications. The line between basic and applied research becomes ever more fuzzy. The new frontiers in biology today thus become the frontiers of biotechnology tomorrow.

For many years biology seemed a study in pure science. Biologists were rarely consulted by industry, unlike their colleagues in chemistry, engineering, and law. Today all that has changed—and there is anguished ethical study of the dangers to academic research of the new practicality of biological research. The dangers are real, but no different from those found for other research areas at an earlier time. The finding that the discoveries in the modern biology laboratory are of great practical importance in industry, as they have been in medicine for many years, should only add interest and excitement to the frontiers of biology and biotechnology being uncovered each day.

As this book goes to press and we plan new volumes in other areas for the future, an editor cannot help but reflect on the personality of individuals who are satisfied by such an unceasing quest. Are we scientists just curious children who have never grown up? Are we the most idealistic of people, bravely confronting the ultimate challenges for the good of humankind? Or are we the most selfish of its citizens, who have discovered the ideal way of life: solving nature's crossword puzzles while being subsidized in our happiness? Whatever the answer, we are all, in the words of the poet, "emperors of the endless dark, even in seeking."

Biotechnology
The Renewable Frontier

Contents

III. DEVELOPMENTAL BIOLOGY AND CANCER

IV. HORMONES AND METABOLISM

V. BIOTECHNOLOGY

VI. VIROLOGY

VII. PLANT SCIENCES

VIII. BEHAVIOR AND SENSORY PHENOMENA

Part I

New Techniques

1

Strategies and Applications of in Vitro Mutagenesis

David Botstein and David Shortle

Biochemists have begun to use mutations to probe the relationship between the structure and activity of proteins; cell biologists are using mutations to define the roles of particular proteins and protein assemblies in the cell; developmental biologists are using mutations to determine the logic and order of molecular events during differentiation and morphogenesis; and neurobiologists are beginning to turn to mutations to try to understand the way in which neural networks are formed and, eventually, how they function. These new and expanding applications of mutations in many disciplines of biology represent one of the most important consequences of the revolution in the life sciences that has resulted from the development of recombinant DNA technology. The objective of this review article is to sketch the outlines of the many mutagenesis strategies made possible by the availability of cloned genes. Our emphasis on general principles and applications has required that we gloss over many ideas and technical accomplishments in the field; such information, though, has been reviewed (*1*, *2*).

Before we address the question of how best to isolate or construct mutations, it is important to review the fundamental logic behind the use of gene mutations to analyze biological phenomena. The primary reason for isolating and characterizing a mutation is to assess its consequences, or in genetic terminology, its phenotype. In the ideal mutation experiment, two organisms are submitted to careful comparison, one being mutated at a single known site and the other lacking the mutation. Any observed difference between the two is then attributable to the mutation; by characterization

Science 229, 1193–1201 (20 September 1985)

of this difference, inferences can be made about the function of the corresponding wild-type gene, regulatory signal, or nucleotide pair. In other words, the properties of a normal function can be learned from the consequences of perturbing or eliminating a single gene or genetic element.

One approach for designing the ideal mutation experiment is to begin with a phenotype of interest (such as failure to complete mitosis at a given temperature) and search, after classical mutagenesis, for mutant organisms that exhibit this phenotype. The mutations obtained by this strategy would then be characterized by genetic mapping, followed by cloning and sequencing the wild-type and mutant genes. A second approach is to begin with the idea that a particular protein (such as tubulin) is likely to play an important role in mitosis. To establish this point genetically, it would be necessary to clone the gene, induce mutations in the cloned copy in a recombinant DNA molecule, and return these mutant genes to the organism in a way that would allow assessment of phenotype.

Both these approaches might lead to the same experiment, although with unequal likelihood. The first approach, in which the phenotype determines the selection of mutations for analysis, has the advantages that no detailed hypotheses are required to obtain material for study and that each new mutation has a high probability of contributing to an understanding of the phenomena underlying the phenotype. The major advantage of the second approach, in which the gene for a particular protein is cloned initially, is that very specific hypotheses, often arrived at by earlier biochemical or biological studies, can be put to definitive genetic tests.

Although, in principle, every gene that encodes information relevant to a given biological process should be found by the first approach if the "correct" phenotype has been used to search for mutations, the results, in practice, are different. For example, extensive screening for cell-division–cycle mutants in yeast failed to produce mutations in the β-tubulin gene, despite the fact that mutations in that gene isolated by the second approach (*3*) exhibited precisely the same cell-division–cycle phenotype used in the unsuccessful screening. Perhaps the most important feature of in vitro mutagenesis is the ability to efficiently and predictably introduce mutations into a gene of interest.

Classical in Vivo Mutagenesis

The only mutations available for study by the earliest geneticists were natural variants and occasional spontaneous mutations. The discoveries that organisms exposed to x-rays (*4*) and certain chemical compounds (*5*) yield much higher frequencies of mutant progeny led to a revolution in experimental genetics. Because these agents gave geneticists partial control over the process of mutagenesis, they could systematically study biological phenomena by collecting large numbers of different mutations displaying a characteristic phenotype, which could then be classified on the basis of complementation behavior, map position, and various other criteria. In this way the genetic loci relevant to the phenomena could be identified, counted, positioned on a genetic map, and then functional interactions between loci defined. With the classical studies of eye pigmentation mutations in *Drosophila*

melanogaster (*6*) and of auxotrophic mutations in *Neurospora crassa* (*7*), it became apparent that the detailed phenotypic characterization of mutations could be extended to the level of individual molecules.

The power of classical in vivo mutagenesis to provide material for genetic and biochemical analysis is attested to by the progress made over the past few decades in the genetics of eukaryotes such as *D. melanogaster* and the molecular genetics of *Escherichia coli* and other microbes. Nevertheless, the requirement for a specific phenotype in order to identify rare mutations in a mixture of wild-type and irrelevant mutations imposed serious limits on the range of phenomena to which genetics could be applied. It was not feasible to study more than a few mutations in a single gene without the special circumstances of a unique phenotype that is easy to score plus an efficient genetic crossing system for weeding out secondary mutations in irrelevant genes caused by the general mutagenesis. Isolation of mutations in a gene of special biochemical interest by simple assay ("brute force," in the jargon of the geneticist) required not only heroic amounts of labor to screen thousands of mutagenized organisms but also major (even lucky) assumptions about the phenotype. A good example is the isolation of a mutation in DNA polymerase I of *E. coli* (*8*).

In bacteria, some improvements in mutagenesis were afforded by the concept of "localized mutagenesis" (*9*). New mutations could be limited to individual segments of a genome by mutagenizing transducing particles that carry only a small fragment of the bacterial chromosome, followed by generalized transduction with selection for a marker known to be closely linked to the genes targeted for mutagenesis. The development of specialized transducing phages and episomes carrying only a portion of the bacterial genome also allowed the mutagenesis of specific genes without exposure of the entire genome of the host cell to the action of a mutagen. Despite having some of the advantages of the new methods of in vitro mutagenesis, mutagenesis of these naturally occurring recombinants had only limited applicability because many genes of interest in *E. coli* could not be readily isolated on specialized transducing phage or small F′ episomes.

Transposon Mutagenesis

A second revolutionary development in mutagenesis came with the realization that insertion mutations could be induced in virtually any gene of interest by the appropriate manipulation of naturally occurring transposable elements. Although these mobile segments of DNA had been discovered many years earlier in maize and *Drosophila*, it was in bacteria that the advantages of insertion mutagenesis were first systematically exploited (*10*).

Unlike chemical or radiation mutagenesis, transposon insertion mutagenesis results in a single, unique physical alteration in the gene that has been mutated. Mutagenesis can frequently be carried out in such a way as to limit the number of transposons inserted to essentially one per genome. Most importantly, the insertion event can often be selected by means of drug resistance or phage immunity carried by the transposon. Thus, a population of organisms, each of which had a transposon insertion within a gene

or intergenic segment, can easily be generated for phenotypic screening. Mutagenesis is therefore extremely efficient, the level of secondary mutations is very low, and most mutations lead to total inactivation of the gene. If the transposon inserts randomly, the probability of finding a mutation in a given gene is the fraction of the total genome that the gene of interest occupies. Thus, for bacteria, about 1 of 3000 to 5000 organisms with a single insertion are expected to be mutated in a given gene. Transposon mutagenesis can be applied to very small genomes, such as plasmids, and is a simple alternative to the use of deletion mutations for defining the extent of a gene of interest (or a small cluster of genes) after it has been cloned [for examples, see (*11*, *12*)].

Probably the single most important advantage of insertion mutations is that they contain an insertion of a known DNA element, the transposon. As a result, transposon mutagenesis can be used to isolate genes by first identifying an insertion mutation in or very near the gene of interest, on the basis of phenotype or genetic linkage, and then cloning a fragment of DNA from the mutant genome that harbors the nucleotide sequences of the transposon. This transposon-tagging technique has made possible the efficient identification and cloning of interesting genes in *Drosophila* (*13*) and also, by means of an RNA tumor virus as the transposon, in mice (*14*).

In Vitro Mutagenesis

Since the introduction of recombinant DNA methodology, genes can be removed from their normal environment in an intact genome and isolated as DNA fragments on cloning vectors. The availability of purified genes in vitro in microgram amounts has dramatically expanded the potential for inducing mutations. In the controlled environment of the test tube, it is now possible to alter, efficiently and systematically, the sequence of nucleotides in a segment of DNA. In the following sections the new methods of in vitro mutagenesis are divided into three broad categories: (i) methods that restructure segments of DNA, (ii) localized random mutagenesis, and (iii) oligonucleotide-directed mutagenesis. This classification emphasizes the practical aspects of each method's application.

The considerable increases in mutagenic efficiency and specificity attainable with the new methods, however, do exact a price. Because these methods are designed for use on isolated DNA molecules, a gene must almost always be removed from its normal genetic context—a unique locus on a large complex chromosome inside a living cell (or virus)—and inserted into the abnormal context of a small cloning vector propagated in *E. coli*. Unlike classical in vivo mutagenesis, in which all mutations are isolated in situ, in vitro mutagenesis invariably yields gene mutations out of their normal context. This is the most radical and most troublesome difference between the classical methods and the powerful in vitro methods.

For some applications, this change in genetic context is relatively unimportant. For other applications, though, inferences about a wild-type gene, on the basis of the phenotype of a mutation construction in vitro, can only be made after the mutant allele has been restored to its normal genetic context. In such situations, the genetic manipulations re-

quired to assess the consequences of a mutation in its proper context become the primary challenge to the molecular biologist. Therefore, the last two sections of this review article are devoted to a discussion of the variety of available solutions, some partial and some complete, to this important problem.

Restructuring of DNA Segments

After a gene of interest has been isolated, it is usually necessary to reduce to a minimum the size of the cloned DNA segment carrying the gene and to move it into a small, circular cloning vector. By a form of deletion analysis, extragenic flanking sequences are systematically eliminated from the initially cloned DNA segment and each new deletion is tested, with some assay of structure or function, to determine that the gene is still intact. In this way the ends or boundaries of the gene are roughly defined while, at the same time, smaller subclones are isolated that will simplify subsequent manipulations of the gene, such as DNA sequencing and mutagenesis.

Further reductions in size, which permit very precise definition of the functional boundaries at the 5′ and 3′ ends of the gene, are achieved by generating terminal deletions with one end point located outside of the gene and the other positioned progressively closer to the gene. Collections of deletions, which have been extensively used to identify such regulatory sequences as transcriptional promoters, can be readily constructed by using an exonuclease [either Bal 31 (*15*) or exonuclease III plus S1 nuclease (*16*)] to remove nucleotides starting from a unique site just outside the gene. This site is generated by restriction enzyme cleavage of a circular DNA molecule; in some cases, nuclease digestion can be confined to one of the two ends of the linear DNA molecule, leading to deletions that extend in a single direction (*17, 18*).

Insertion of a small synthetic oligonucleotide that encodes a unique restriction enzyme cleavage site at the end points of these deletions (*16*) or at positions randomly distributed across the cloned DNA segment (*19*) provides a readily available restriction site for use in further restructuring of the cloned DNA segment. These oligonucleotide insertions, termed linker mutations, are especially versatile for the modification of circular DNA molecules. For instance, cleavage at a linker insertion introduced near or within the gene provides a site at which additional nucleotides can be inserted or deleted. It may also provide ends by which the cloned DNA segment itself can be inserted into other types of DNA vectors for procedures such as nucleotide sequencing, production of large quantities of the gene product, and transformation of other types of cells. In addition, by joining restriction fragments isolated from pairs of linker insertion mutations of the appropriate size and position, mutations consisting of four to eight tightly clustered base substitutions can be generated, making it possible to efficiently "scan" a small region of DNA for regulatory sites (*20*).

When changes in the level of expression of a gene are difficult to monitor because of the lack of a convenient assay for the gene product, a common strategy is to construct a gene fusion in which the regulatory elements of the gene of interest control the expression of a gene product that can be readily quantitated. Typically, a second gene specifying an