

Perspectives

Anecdotal, Historical and Critical Commentaries on Genetics

Edited by James F. Crow and William F. Dove

The *hisB463* Mutation and Expression of a Eukaryotic Protein in *Escherichia coli*

Kevin Struhl¹

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

CERTAINTY in science is never achieved and can only be approached asymptotically. How close to the asymptote and what kind of evidence is necessary to demonstrate a scientific point beyond reasonable doubt? How important is missing information that may be difficult or impossible to obtain with the available technology?

Jeff Strathern just determined the sequence of the *Escherichia coli hisB463* allele, the end to a story that began 33 years ago. The *hisB463* mutation was critical for the first demonstration of functional expression of a eukaryotic protein in *E. coli* and the cloning of the first yeast gene (STRUHL *et al.* 1976). The evidence for such functional expression relied on genetic analysis, with the molecular basis of the *hisB463* mutation being inferred, but unknown until now. Here, I provide a personal history of *hisB463* and its role in the early days of recombinant DNA technology and yeast molecular biology.

With the advent of recombinant DNA technology ~35 years ago, it was imagined that expression of eukaryotic proteins in *E. coli* would be of great practical value. Nevertheless, it was widely believed that this would be difficult due to a high functional barrier between eukaryotes and prokaryotes arising from different molecular mechanisms of gene regulation. In addition, the issue of functional expression of proteins across species barriers was very controversial and garnered extensive press coverage. It engendered fears of public safety and ethical qualms about creating new forms of life with combinations of genetic material completely unlike those existing in nature. In this respect, the public and political response toward recombinant DNA technology was a harbinger of more recent issues such as

therapeutic and reproductive cloning via embryonic stem cells, genetic testing for human disease, and genetic modification of plants and animals for food production. For biological research, recombinant DNA technology also initiated the transition from an academic discipline to the development of a major industry.

I joined the Department of Biochemistry at Stanford Medical School as a graduate student in the fall of 1974. This decision was made ~6 months earlier, during a visit (actually a day of interviews) when I first heard about the pioneering work on recombinant DNA technology that was going on in the department. Noteworthy achievements in the department included the construction of the first hybrid DNA molecules (JACKSON *et al.* 1972; LOBBAN 1972; LOBBAN and KAISER 1973), the discovery that restriction endonucleases generated cohesive ends suitable for ligation (MERTZ and DAVIS 1972), and generations of collections of recombinant molecules containing bacteriophage λ and eukaryotic DNA (THOMAS *et al.* 1974). As an undergraduate at the Massachusetts Institute of Technology studying molecular biology and working in Boris Magasanik's laboratory, I was introduced to the field of prokaryotic gene regulation, which was still in its heyday. Very little was known about gene regulation in eukaryotes at the time, but that was what I wanted to pursue, even if I did not have a clear idea about how. Recombinant DNA technology was the answer.

During my first visit to Stanford, I was convinced that I wanted to do my Ph.D. with Ron Davis, and I joined his laboratory shortly after my arrival. Ron had been at Stanford for only 3 years, but he was already a major figure in the fledgling field of recombinant DNA technology (MERTZ and DAVIS 1972; THOMAS *et al.* 1974). This included the creation of collections of recombinant DNA molecules with eukaryotic DNA that, amusingly, were called "pools" in California, "banks" in Switzerland, and "libraries" at Harvard (the last term

¹Address for correspondence: Department of Biological Chemistry, Harvard Medical School, 200 Longwood Ave., Boston, MA 02115. E-mail: kevin@hms.harvard.edu

eventually won). Prior to my arrival, Ron's first graduate student, John Cameron, had cloned the *E. coli* DNA ligase gene from a hybrid pool using a genetic selection based on phage growth (CAMERON *et al.* 1975). My first project was to extend this work by cloning genes from other bacteria, and this resulted in the isolation of the DNA polymerase I gene from *Klebsiella aerogenes* and *K. pneumoniae* (STRUHL and DAVIS 1980).

My real goal, however, was to clone a eukaryotic protein-coding gene, both for its own sake and for initiating molecular analysis of transcriptional regulation. As my undergraduate research involved a heavy dose of P1 transductions in *Klebsiella*, I decided to do this by functional complementation of an *E. coli* auxotroph. Such functional complementation was viewed as a long shot (or worse), because of the large evolutionary distance and mechanistic differences between prokaryotes and eukaryotes. However, I thought these mechanistic differences were largely irrelevant and did not constitute a functional barrier between prokaryotes and eukaryotes. Instead, the common genetic code and the ability of *E. coli* to initiate translation at internal AUG codons would permit synthesis of the correct eukaryotic protein, provided there was any reasonable level of transcription throughout the gene. My thinking would undoubtedly have been different if I had known about introns, but their discovery was 2 years later (BERGET *et al.* 1977; CHOW *et al.* 1977).

Even with the belief that there was no fundamental barrier against expressing a eukaryotic gene in *E. coli*, there were practical concerns. It was virtually certain that significant gene-specific differences would occur at any of several levels: (1) transcription through a eukaryotic protein-coding region; (2) translational initiation at the relevant AUG codon; (3) stability, folding, post-translational modification of a eukaryotic protein in *E. coli*; (4) the ability of a eukaryotic protein to function under the physiological condition of an *E. coli* cell (*e.g.*, pH, temperature, ionic strength); or (5) the amount of functional eukaryotic gene product needed to support the growth of an *E. coli* auxotroph. In addition, the few vectors available in early 1975 all involved the cloning of *EcoRI* fragments. As a consequence, some eukaryotic genes would not be functionally expressed owing to the presence of an *EcoRI* site in the protein-coding region, whereas others would be present on very large *EcoRI* fragments that could not be easily cloned. Thus, it was impossible to predict which, or even what percentage of, eukaryotic genes could be isolated by functional complementation.

Given these practical concerns, the strategy was to introduce collections of recombinant DNA molecules into multiple *E. coli* auxotrophs. If the basic idea was correct, then some of the attempts should have worked, even if many failed for the reasons above. Implementation of this strategy was greatly facilitated by the use of bacteriophage λ -vectors (MURRAY and MURRAY 1974;

THOMAS *et al.* 1974) as opposed to plasmid vectors. Aside from the ability to generate larger and hence more comprehensive libraries, λ -vectors permitted one to generate high-titer stocks of the hybrid phage easily and to introduce these hybrid phages into *E. coli* efficiently by simple infection. In contrast, introduction of plasmid DNAs into *E. coli* by transformation was far less efficient on a per-cell basis. As the standard assay for auxotrophic complementation required cells to form colonies on a medium lacking the required nutrient, the hybrid phages needed to be stably integrated into the *E. coli* genome as prophages.

In principle, any auxotrophic *E. coli* strain capable of being infected by λ -phages could be used to search for a yeast DNA segment capable of complementation. Ideally, the auxotrophic mutation would be nonrevertible, so that colonies emerging on the selection plate would likely involve functional expression of eukaryotic DNA rather than reversion or suppression of the original mutation. While I was prepared to try many auxotrophic strains, the initial experiments were biased by the potential for future work, should functional complementation be successful. I was most interested in histidine auxotrophs, because the work of Gerry Fink over the previous decade had determined the gene-enzyme relationships of the histidine pathway in yeast (FINK 1964) and had identified genes that positively or negatively regulate the general control pathway in which many enzymes involved in the biosynthesis of histidine and other amino acids are coordinately expressed (WOLFNER *et al.* 1975).

Through the pioneering work of Bruce Ames starting in the 1950s, regulation of histidine biosynthesis was one of the classical paradigms of prokaryotic gene regulation, but virtually all of the work had been done in *Salmonella typhimurium*. In fact, there were only two articles in which *E. coli* histidine auxotrophs were classified according to which enzymes were inactivated (GARRICK-SILVERSMITH and HARTMAN 1970; GOLDSCHMIDT *et al.* 1970). I obtained a set of *E. coli* histidine auxotrophs from Phil Hartman and infected them with a pool of $\sim 10,000$ λ -yeast hybrid phages.

In the very first experiment, two His⁺ colonies appeared near the edge of the *hisC463* plate. My first reaction was that this was too good to be true and that these were revertants or contaminants. However, there were no colonies on the control *hisC463* plate that was infected with the λ -vector, and the *hisC463* mutation had been characterized as nonrevertible (GARRICK-SILVERSMITH and HARTMAN 1970). In addition, His⁺ colonies were not observed in parallel experiments involving hybrid pools containing DNA from the slime mold *Dictyostelium discoideum*. More importantly, curing the His⁺ colonies of the lysogenic phages rendered the strain His⁻, and conversely, reinfection by the phage recovered from the original His⁺ colonies into new *hisC463* cells resulted in a large number of His⁺ colonies.

Thus, the phage recovered from the initial His⁺ colonies contained a segment of DNA that could functionally complement or suppress the *hisC463* mutation.

A crucial issue at this point was whether the complementing segment of DNA was actually derived from yeast. A contaminating piece of DNA from *E. coli* or some other prokaryotic organism could possibly complement the *hisC463* mutation, thereby providing a trivial explanation. In this regard, the *EcoRI* endonuclease and DNA ligase used to construct the hybrid phage DNAs came from *E. coli*, and the yeast genomic DNA was prepared from spheroplasts generated with a bacterial enzyme preparation. I remember analyzing these various enzyme preparations for DNA by gel electrophoresis and ethidium bromide staining as well as by direct visualization in the electron microscope (in his doctoral work with Norman Davidson, Ron Davis developed methods for electron microscopic analysis of nucleic acids including heteroduplex analysis, and the department had an electron microscope). But, the ultimate test utilized a new technique, now known as a Southern blot (SOUTHERN 1975), and I was fortunate enough to have a preprint describing this revolutionary technology from Ed Southern. The complementing phage contained a 10-kb *EcoRI* fragment, and hybridization of a ³²P-labeled probe derived from this fragment to a Southern blot of *EcoRI*-digested yeast genomic DNA revealed a single band corresponding to a 10-kb *EcoRI* fragment, thereby proving that the complementing fragment was indeed derived from yeast. For a variety of technical reasons, it took 4 days to see a weak hybridization signal in the first experiment; later on, it took <1 hour. Taken together, these experiments demonstrated functional expression of eukaryotic DNA in *E. coli*.

The initial experiments also indicated that the transcription for such functional expression initiated from the yeast DNA segment. In particular, His⁺ complementation occurred when the yeast DNA segment was cloned in either orientation with respect to the λ -vector sequences in the context of a λ -lysogen, a condition in which the major lytic promoters of λ are repressed. Thus, if the promoter did not reside in the yeast DNA segment, two unknown λ -promoters would be required for complementation. Subsequent experiments mapped this "yeast in *E. coli*" promoter to a region just upstream of the yeast protein-coding region (STRUHL *et al.* 1980) that resulted from fortuitous homology between the eukaryotic TATA element and the prokaryotic -10 promoter element (STRUHL 1986).

In addition to its scientific interest, the demonstration of functional expression of eukaryotic DNA in *E. coli* fueled further the raging debate on the safety of recombinant DNA technology. Just before I arrived at Stanford, a group of distinguished scientists had pointed out the potential biohazards of recombinant DNA molecules and proposed a voluntary moratorium on such experiments (BERG *et al.* 1974). This morato-

rium ended with a complex set of recombinant DNA guidelines, first voluntary, then required by the National Institutes of Health under penalty of being shut down. Some individuals thought that the creation of completely new organisms by a man-made process outside of natural selection was potentially dangerous, even cataclysmic, in ways that could not be predicted or even coherently described. Those of us actually doing recombinant DNA experiments thought the risks of such experiments to be very remote and believed the biohazard guidelines to be excessive, arbitrary, and confusing; hence, we followed them grudgingly. Nevertheless, expression of eukaryotic DNA in *E. coli* meant that organisms generated by recombinant DNA technology would not only be a source of cloned DNA, but would also have new functional properties.

Given the excitement of functional expression of eukaryotic DNA in *E. coli*, I started to prepare a manuscript for publication. The key question at this point became whether the complementation was caused by production of the yeast enzyme corresponding to the *hisC* gene product (histidinol phosphate aminotransferase, which was encoded by the yeast *HIS5* gene) or by expression of an activity that suppressed the *hisC463* mutation. The main argument favoring expression of the yeast enzyme was that the *hisC463* mutation was nonrevertible, as confirmed by my inability to obtain His⁺ suppressors (frequency < 10¹¹) even after a variety of mutagenic treatments. This issue would certainly need to be discussed. To directly address the issue, I began the next phase of the project, namely biochemical characterization of the enzyme produced in *E. coli* expressing the eukaryotic DNA.

Unlike Captain Renault's response upon his discovery of gambling in Rick's nightclub in the cinematic classic *Casablanca*, I was shocked to find that the wild-type and *hisC463* mutant strain had comparable levels of histidinol phosphate aminotransferase activity. Regrowing the strains from different isolates and repeating the assays did not change the results. This immediately halted the writing, as I certainly could not publish an article on functional complementation if the mutant strain did not have the expected biochemical defect. The strain was a histidine auxotroph, and it was one of the few histidine auxotrophs in the published collection that was nonrevertible (GARRICK-SILVERSMITH and HARTMAN 1970), so it seemed likely that it was simply mischaracterized or misrecorded with respect to which *his* gene was affected by the mutation.

Fortuitously, a fellow graduate student, John Scott, had a complete collection of Salmonella strains from which individually mutated *his* genes, present on an F factor containing the entire *his* operon, could be transferred to *E. coli* through mating. He had these strains from his undergraduate research at Berkeley with John Roth, a major contributor to understanding regulation of the *his* operon (his son Fritz Roth is now my

departmental colleague and a collaborator on several publications). Standard complementation analysis using these *Salmonella* strains quickly revealed that the supposed *hisC463* strain actually contained a mutation in the *hisB* gene. Given the proximity of B and C in the alphabet, it is easy to imagine a recording error occurred when processing a large number of histidine auxotrophs.

Although the fact that the *E. coli hisC463* strain actually contained a *hisB* mutation (which I renamed *hisB463*) easily explained the normal levels of the *hisC* gene product, it posed a new problem. The *hisB* gene encodes a bifunctional protein with two enzymatic activities, imidazole glycerol phosphate (IGP) dehydratase and histidinol phosphate phosphatase, that, respectively, mediate the sixth and eighth steps of histidine biosynthesis. In yeast, these two enzymes are encoded by separate genes (*HIS3* and *HIS2*, respectively) located on different chromosomes; thus, a single segment of yeast DNA could not possibly express both enzymes in *E. coli*. If the yeast DNA segment expressed one of these enzymes, as opposed to having some bizarre suppressor activity, the *hisB463* mutation had to be specifically defective for one of the two *hisB* activities.

I was pleased to find that the *hisB463* strain indeed had normal levels of histidine phosphate phosphatase activity, implying a specific defect in IGP dehydratase. However, the assay for IGP dehydratase was more difficult, because the substrate (IGP) was not commercially available and had to be chemically synthesized and purified. The main source of IGP was Bruce Ames himself, and his laboratory at Berkeley was less than an hour's drive from Stanford. I arranged to go there and perform IGP dehydratase assays, and it was gratifying to see that the *hisB463* strain lacked detectable IGP dehydratase activity. I was of course relieved, as it would have been very embarrassing to publish the first article on functional expression of eukaryotic DNA in *E. coli* (and my first scientific article) with the wrong gene being expressed!

Given the new view of *E. coli hisB463*, the experiments indicated that the yeast DNA segment expressed yeast IGP dehydratase, the *HIS3* gene product, in *E. coli*. The alternative view that the yeast DNA segment encoded a factor that suppressed the *hisB463* mutation was highly implausible. First, as mentioned above, *hisB463* was completely nonrevertible to His⁺ even after treatment with a variety of mutagens, suggesting that the mutation was a deletion that inactivated IGP dehydratase without affecting histidinol phosphate phosphatase activity. The lack of His⁺ revertants excluded informational suppression (*e.g.*, of nonsense or frameshift mutations), because such suppressors arise at easily detectable frequencies. More generally, it was difficult to conceive of how the product of a yeast gene could suppress a mutation that could not be reverted or suppressed by any *E. coli* mechanism. Second, the yeast DNA segment

also complemented *hisB2404*, a revertible *hisB* allele selectively defective for IGP dehydratase activity, whereas it was unable to complement a *hisB* mutation lacking both enzymatic activities or any other allele in other *his* genes (these *E. coli* strains were generated via mating with the *Salmonella* derivatives described above for the complementation analysis). Third, the yeast *HIS3* gene product is the only enzyme with IGP dehydratase activity, because *his3* mutants are histidine auxotrophs and some *his3* mutants are nonrevertible to His⁺. Hence, it seemed extremely unlikely that functional complementation of *hisB463* and *hisB2404* was mediated by expression of an enzyme that fortuitously catalyzes IGP dehydratase activity and is encoded by a locus distinct from *HIS3*.

I first presented this work at a symposium held in Keystone, Colorado, in March 1976, and it generated both a great deal of excitement and a certain amount of skepticism. It had actually been submitted to the Proceedings of the National Academy in late 1975, and it was published in May, 1976 (STRUHL *et al.* 1976). A commentary in *Nature News & Views* appearing shortly thereafter found the genetic evidence convincing, and it described this work as the first example of functional expression of a eukaryotic protein in *E. coli* (ATKINS 1976).

However, as was clear at Keystone and from discussions at meetings I subsequently attended, as well as through comments I heard through the grapevine, this was not a universal reaction. Not everyone was convinced by the genetic evidence, in part due to the checkered history and unknown molecular basis of the *hisB463* allele. On multiple occasions over the next year, I heard comments that suppression of *hisB463* was not excluded as an alternative explanation, as well as rumors that "there was something wrong" with the article and that the mutation was really in the *hisC* gene. Such skepticism about functional expression of a eukaryotic protein in *E. coli* gradually dissipated with subsequent examples involving yeast (RATZKIN and CARBON 1977), *Neurospora* (VAPNEK *et al.* 1977), and mammalian (ITAKURA *et al.* 1977) proteins. Skepticism concerning functional expression of yeast IGP dehydratase was finally eliminated by showing that the equivalent DNA fragments from two yeast *his3* mutant strains were unable to complement *hisB463*, but could be recombined via a phage cross to generate a complementing fragment (STRUHL and DAVIS 1977). Furthermore, the equivalent DNA fragment from a strain with a *his3* amber mutation could not complement the *hisB463* mutation unless the strain also contained an appropriate suppressor tRNA (STRUHL *et al.* 1979). Lastly, the enzymatic properties of the *HIS3*-encoded enzyme produced in *E. coli* were similar to the enzyme found in yeast cells (STRUHL and DAVIS 1977).

Scientific proof is approached asymptotically, but never reached. In my opinion, then and now, the initial genetic analysis demonstrated the expression of a eu-

karyotic protein in *E. coli*. Alternative explanations of the data required a combination of highly unlikely scenarios. Nevertheless, there is often a reluctance to accept molecular conclusions solely on the basis of genetic analysis. First, it is commonly believed that genetic experiments provide only inferences for molecular understanding, whereas biochemical experiments provide direct evidence. However, interpretation of biochemical data (*e.g.*, a band on a gel) also relies on a chain of inferences and abstractions, so it is an illusion that biochemical experiments provide direct molecular information. Second, even when there is a straightforward interpretation of the genetic observations, it is routine to consider alternative explanations that are not formally excluded by the data. Nevertheless, it is important to consider the likelihood that such formal alternatives fit all the available data and whether they are significant enough to cast reasonable doubt on the main conclusion. Third, genetic analysis often deals with, and indeed selects for, very low ($<10^{-9}$) frequency events, whereas biochemical assays are much less sensitive, thereby making it difficult to identify or measure infrequent events.

As a consequence, conclusions or models generated through genetic analysis are often considered hypothetical, perhaps even ephemeral, until there is confirmatory molecular evidence. But, it does not take much molecular evidence to convert a hypothetical model to a molecular dogma. A classic example of this is the “cassette” model for the control of mating type in *Saccharomyces cerevisiae*, which was largely developed in the laboratory of Ira Herskowitz (HICKS *et al.* 1977) with important earlier contributions from Yasuji Oshima (OSHIMA and TAKANO 1971), Gennadi Naumov (NAUMOV and TOLSTORUKOV 1973), and Don Hawthorne (HAWTHORNE 1963; see HERSKOWITZ 1988). The cassette model involves an active mating-type locus that expresses either **a** or α -information and two silent loci that contained the same **a** or α -information that is not expressed. In this model, mating-type switching occurs by replacing the active cassette with a silent cassette containing information from the opposite mating type. The genetic evidence for the cassette model was overwhelming (HICKS and HERSKOWITZ 1977; KLAR and FOGEL 1979; KLAR *et al.* 1979; STRATHERN *et al.* 1979a,b), but many were not fully convinced until the genes were cloned and shown by Southern blotting to exist in three copies (HICKS *et al.* 1979). The molecular confirmation and much of the previous genetic evidence was the work of Jim Hicks, Amar Klar, and Jeff Strathern, who was mentioned at the beginning of this historical account.

As mentioned above, the checkered history and unknown molecular nature of the *hisB463* allele contributed to the initial skepticism about functional expression of the yeast *HIS3* gene in *E. coli*. Based on its nonrevertible nature even upon treatment with mutagens, *hisB463* was suggested to be a small deletion

that inactivated IGP dehydratase, but not histidinol phosphate phosphatase (STRUHL *et al.* 1976), and this description persisted for decades without any additional evidence. DNA sequencing was not available at the time I did this work, and the matter was never pursued until Jeff Strathern contacted me several months ago. Jeff asked if I knew the molecular nature of *hisB463*, because he wanted to confirm the genotype of *E. coli* strains he was going to use. Of course I did not, and this led Deanna Gotte in Jeff’s laboratory to sequence the relevant PCR fragments from wild-type and *hisB463* strains.

To my delight, the *hisB463* allele in *E. coli* is an in-frame 6-bp deletion that removes two amino acids, V232 and E233, from the *hisB* coding region. These residues are located within the C-terminal IGP dehydratase domain (residues 168–356), which is structurally independent of the N-terminal histidine phosphate phosphatase domain (residues 1–167) (CARLOMAGNO *et al.* 1988). IGP dehydratases are highly conserved through evolution (BRILLI and FANI 2004), and X-ray crystal structures of fungal (SINHA *et al.* 2004) and plant (GLYNN *et al.* 2005) enzymes have been reported. IGP dehydratase is composed of 24 identical subunits, with each subunit containing an internal α/β repeat. Assembly of the active 24-mer from an inactive trimer depends on a dimanganese cluster involving two metal-binding motifs. The deleted residues in the protein encoded by *hisB463* lie within one of these metal-binding motifs, and E233 is directly involved in coordinating one of the manganese ions.

So, it is now clear why *hisB463* is a nonrevertible mutation, why the encoded protein has histidinol phosphate phosphatase activity, but is completely defective for IGP dehydratase activity, and why functional complementation of *hisB463* by a segment of yeast DNA could occur only by functional expression of yeast IGP dehydratase and not by suppression of the mutant allele. The sequence of the *hisB463* mutation brings us asymptotically closer to scientific certainty.

I thank Jeff Strathern and Deanna Gotte for sequencing the *hisB463* allele and for many enjoyable conversations with Jeff over the past 30 years. I also thank Marjorie Oettinger and Zarmik Moqtaderi for useful comments on the manuscript. The work described in this history was supported by a grant to Ron Davis from the National Institutes of Health (GM 21891). My subsequent work on the yeast *HIS3* gene was supported by a postdoctoral fellowship from the Jane Coffin Childs Foundation and, since 1982, by a grant from the National Institutes of Health (GM30186).

LITERATURE CITED

- ATKINS, J., 1976 Expression of a eukaryotic gene in *Escherichia coli*. *Nature* **262**: 256–257.
- BERG, P., D. BALTIMORE, H. W. BOYER, S. M. COHEN, R. W. DAVIS *et al.*, 1974 Letter: potential biohazards of recombinant DNA molecules. *Science* **185**: 303.
- BERGET, S. M., C. MOORE and P. A. SHARP, 1977 Spliced segments at the 5’ terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. USA* **74**: 3171–3175.

- BRILLI, M., and R. FANI, 2004 Molecular evolution of *hisB* genes. *J. Mol. Evol.* **58**: 225–237.
- CAMERON, J. R., S. M. PANASENKO, I. R. LEHMAN and R. W. DAVIS, 1975 In vitro construction of bacteriophage lambda carrying segments of the *Escherichia coli* chromosome: selection of hybrids containing the gene for DNA ligase. *Proc. Natl. Acad. Sci. USA* **72**: 3416–3420.
- CARLOMAGNO, M. S., L. CHIARIOTTI, P. ALIFANO, A. G. NAPPO and C. B. BRUNI, 1988 Structure and function of the *Salmonella typhimurium* and *Escherichia coli* K-12 histidine operons. *J. Mol. Biol.* **203**: 585–606.
- CHOW, L. T., R. E. GELINAS, T. R. BROKER and R. J. ROBERTS, 1977 An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **12**: 1–8.
- FINK, G. R., 1964 Gene-enzyme relationships in histidine biosynthesis in yeast. *Science* **146**: 525–527.
- GARRICK-SILVERSMITH, L., and P. E. HARTMAN, 1970 Histidine-requiring mutant of *Escherichia coli* K-12. *J. Bacteriol.* **66**: 231–244.
- GLYNN, S. E., P. J. BAKER, S. E. SEDELNIKOVA, C. L. DAVIES, T. C. EADSFORTH *et al.*, 2005 Structure and mechanism of imidazoleglycerolphosphate dehydratase. *Structure* **13**: 809–817.
- GOLDSCHMIDT, E. P., M. S. CATER, T. S. MATNEY, M. A. BUTLER and A. GREENE, 1970 Genetic analysis of the histidine operon in *Escherichia coli* K12. *Genetics* **66**: 219–229.
- HAWTHORNE, D. C., 1963 A deletion in yeast and its bearing on the structure of the mating type locus. *Genetics* **48**: 1727–1729.
- HERSKOWITZ, I., 1988 The Hawthorne deletion twenty-five years later. *Genetics* **120**: 857–861.
- HICKS, J. B., and I. HERSKOWITZ, 1977 Interconversion of yeast mating types II. Restoration of mating ability to sterile mutants in homothallic and heterothallic strains. *Genetics* **85**: 373–393.
- HICKS, J. B., J. N. STRATHERN and I. HERSKOWITZ, 1977 Cassette model for mating-type interconversion, pp. 457–462 in *DNA Insertions, Plasmids, and Episomes*, edited by A. BUKHARI, J. SHAPIRO and S. ADHYA. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- HICKS, J. B., J. N. STRATHERN and A. J. KLAR, 1979 Transposable mating type genes in *Saccharomyces cerevisiae*. *Nature* **282**: 478–483.
- ITAKURA, K., T. HIROSE, R. CREA, A. D. RIGGS, H. L. HEYNEKER *et al.*, 1977 Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* **198**: 1056–1063.
- JACKSON, D. A., R. H. SYMONS and P. BERG, 1972 Biochemical method for inserting new genetic information into DNA of simian virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **69**: 2904–2909.
- KLAR, A. J., and S. FOGEL, 1979 Activation of mating type genes by transposition in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **76**: 4539–4543.
- KLAR, A. J., S. FOGEL and D. N. RADIN, 1979 Switching of a mating-type a mutant allele in budding yeast *Saccharomyces cerevisiae*. *Genetics* **92**: 759–776.
- LOBBAN, P. E., 1972 An enzymatic method for end-to-end joining of DNA molecules. Ph.D. Dissertation, Stanford University, Palo Alto, CA.
- LOBBAN, P. E., and A. D. KAISER, 1973 Enzymatic end-to-end joining of DNA molecules. *J. Mol. Biol.* **78**: 453–471.
- MERTZ, J. E., and R. W. DAVIS, 1972 Cleavage of DNA by RI restriction endonuclease generates cohesive ends. *Proc. Natl. Acad. Sci. USA* **69**: 3370–3374.
- MURRAY, N. E., and K. MURRAY, 1974 Manipulation of restriction targets in phage lambda to form receptor chromosomes for DNA fragments. *Nature* **251**: 476–481.
- NAUMOV, G. I., and I. I. TOLSTORUKOV, 1973 Comparative genetics of yeast. X. Reidentification of mutators of mating types in *Saccharomyces*. *Genetika* **9**: 82–91.
- OSHIMA, Y., and I. TAKANO, 1971 Mating types in *Saccharomyces*: their convertibility and homothallism. *Genetics* **67**: 327–335.
- RATZKIN, B., and J. CARBON, 1977 Functional expression of cloned yeast DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **74**: 487–491.
- SINHA, S. C., B. N. CHAUDHURI, J. W. BURGNER, G. YAKOVLEVA, V. J. DAVISSON *et al.*, 2004 Crystal structure of imidazole glycerolphosphate dehydratase: duplication of an unusual fold. *J. Biol. Chem.* **279**: 15491–15498.
- SOUTHERN, E. M., 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**: 503–517.
- STRATHERN, J. N., L. C. BLAIR and I. HERSKOWITZ, 1979a Healing of mat mutations and control of mating type interconversion by the mating type locus in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **76**: 3425–3429.
- STRATHERN, J. N., C. S. NEWLON, I. HERSKOWITZ and J. B. HICKS, 1979b Isolation of a circular derivative of yeast chromosome III: implications for the mechanism of mating type interconversion. *Cell* **18**: 309–319.
- STRUHL, K., 1986 Yeast *HIS3* expression in *Escherichia coli* depends upon fortuitous homology between eukaryotic and prokaryotic promoter elements. *J. Mol. Biol.* **191**: 221–229.
- STRUHL, K., J. R. CAMERON and R. W. DAVIS, 1976 Functional expression of eukaryotic DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **73**: 1471–1475.
- STRUHL, K., and R. W. DAVIS, 1977 Production of a functional eukaryotic enzyme in *Escherichia coli*: cloning and expression of the yeast structural gene for imidazoleglycerolphosphate dehydratase (*his3*). *Proc. Natl. Acad. Sci. USA* **74**: 5255–5259.
- STRUHL, K., and R. W. DAVIS, 1980 Conservation and DNA sequence arrangement of the DNA polymerase I gene region from *Klebsiella aerogenes*, *Klebsiella pneumoniae*, and *Escherichia coli*. *J. Mol. Biol.* **141**: 343–368.
- STRUHL, K., R. W. DAVIS and G. R. FINK, 1979 Suppression of a yeast amber mutation in *Escherichia coli*. *Nature* **279**: 78–79.
- STRUHL, K., D. T. STINCHCOMB and R. W. DAVIS, 1980 A physiological study of functional expression in *Escherichia coli* of the cloned yeast imidazoleglycerolphosphate dehydratase gene. *J. Mol. Biol.* **136**: 291–307.
- THOMAS, M., J. R. CAMERON and R. W. DAVIS, 1974 Viable molecular hybrids of bacteriophage lambda and eukaryotic DNA. *Proc. Natl. Acad. Sci. USA* **71**: 4579–4583.
- VAPNEK, D., J. A. HAUTALA, J. W. JACOBSON, N. H. GILES and S. R. KUSHNER, 1977 Expression in *Escherichia coli* K-12 of the structural gene for catabolic dehydroquinase of *Neurospora crassa*. *Proc. Natl. Acad. Sci. USA* **74**: 3508–3512.
- WOLFNER, M., D. YEP, F. MESSENGUY and G. R. FINK, 1975 Integration of amino acid biosynthesis into the cell cycle of *Saccharomyces cerevisiae*. *J. Mol. Biol.* **96**: 273–290.