Am J Hum Genet 31:531–538, 1979

Synthetic DNA and Medicine

ARTHUR D. RIGGS[1] AND KEIICHI ITAKURA

Synthetic DNA chemistry is no longer an esoteric discipline without obvious practical applications. On the contrary, the combination of synthetic DNA chemistry, recombinant DNA techniques, and molecular cloning already has resulted in useful products—somatostatin [1] and insulin [2, 3]—and promises much more. In this review, we will first discuss the results and methods of our recent work on insulin [3] and mutation correction [4] and then follow with speculation on additional potential applications.

THE INSULIN PROJECT

Bacterial production of human insulin

Figure 1 illustrates the overall scheme that we used [3]. The insulin chains (21 amino acid A chain and 30 amino acid B chain) are made in separate bacterial strains as tails on a rather large precursor protein, the enzyme β-galactosidase. The insulin chains are efficiently clipped from the precursor protein by treatment with cyanogen bromide, a methionine-specific cleavage reagent. Because synthetic DNA was used to make the insulin genes, we were able to arrange that the insulin tails are attached to β-galactosidase by a methionine linkage (see next section).

The yields of the separate chains are extremely good; about 20% of the total bacterial protein is made as the insulin-β-galactosidase precursor protein [3], and even higher yields should be obtained soon. Almost the entire protein-synthesizing machinery of the bacterial cell can be turned to the production of the desired peptide product. Milligrams of the insulin chains are made per liter of bacterial culture. The individual chains can be joined in good yields (up to 80%) by air oxidation [5] and active insulin obtained, as ascertained by chemical, radioimmune, and biological activity [3, and unpublished data]. With our present results, the commercial production of human insulin by bacteria seems practical, and two firms, Genentech, Inc. (South San Francisco, Calif.), and Eli Lilly, Inc. (Indianapolis, Ind.), are trying to develop procedures for large-scale production.

The techniques we used are quite general; thus we are confident that bacteria can be engineered to produce any peptide hormone that does not contain methionine. By using

Received April 25, 1979.

[1] Both authors: Division of Biology, City of Hope National Medical Center, Duarte, CA 91010.

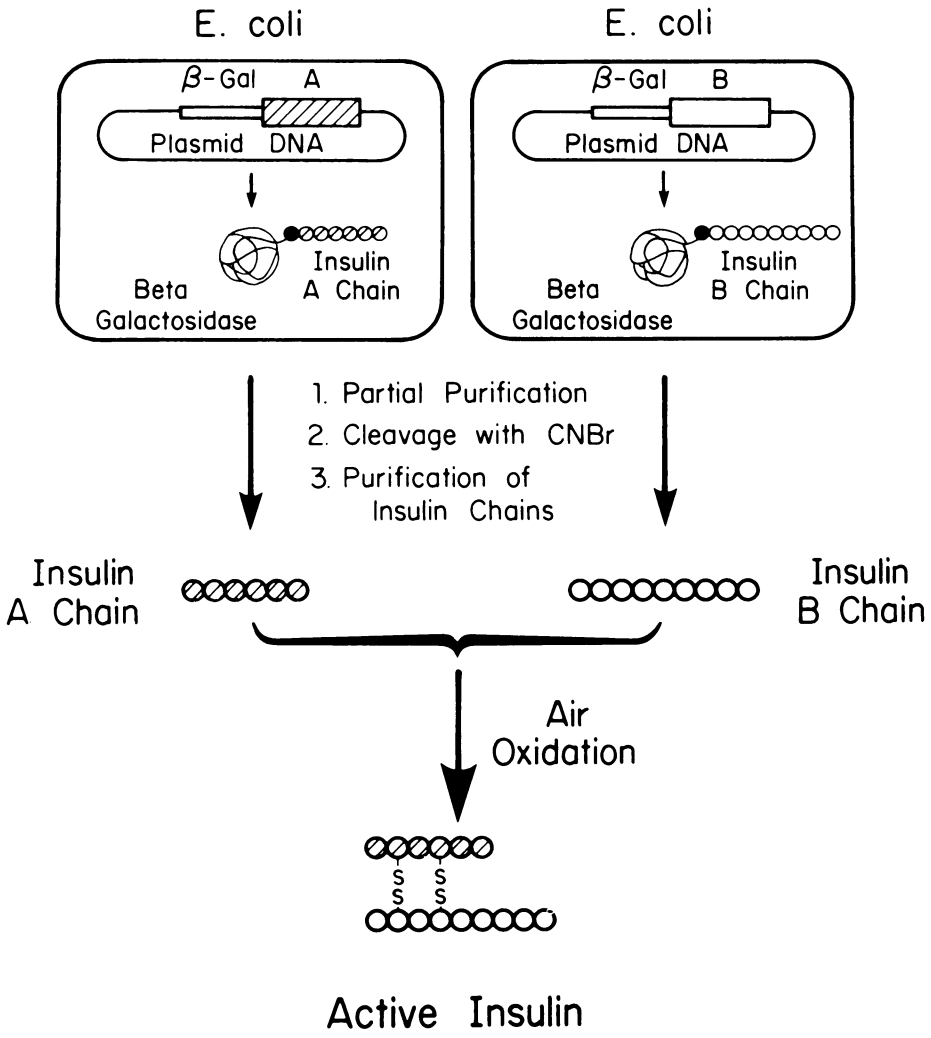

FIG. 1.—Schematic overview of strains and procedures for production of human insulin by bacteria. Two *E. coli* strains were constructed having chemically synthesized insulin A or B chain genes inserted into the β-galactosidase gene of a plasmid cloning vector. In vivo, a fused protein is made, mostly β-galactosidase but with an insulin tail joined by a methionine. In vitro, insulin peptide chain is clipped off by treatment with cyanogen bromide. After separate purification, insulin A and B chains are joined by air oxidation.

other cleavage tricks, or accepting lower yields, even peptides that contain methionine can probably be made.

With chemical DNA synthesis, it is not necessary to copy the natural nucleotide sequence of a gene, because, given the sequence of amino acids in the desired peptide product, one can use the genetic code to design an "artificial" gene carrying the necessary information. This is the approach that we used first for somatostatin [1] and

then for insulin [2]. Techniques have developed rapidly, so that the genes necessary for altering the bacteria can be made and inserted with relatively modest expenditures of time and money. In the next section, we will describe in more detail how genes can be made and inserted into bacteria.

Chemical DNA synthesis

Although there are alternative methods [6], the fastest way to make DNA is by the phosphotriester method [2, 7] illustrated in figure 2. We will not go into details beyond those given in the legend of figure 2, but recent improvements in the method (such as the rapid synthesis of trimers, together with the extensive use of high performance liquid chromatography for analysis and purification of the oligodeoxyribonucleotides) have dramatically reduced the time necessary for the construction of DNA fragments [2]. A library of trimers has been established, and longer oligonucleotides can be assembled quickly from the trimer units (which correspond to amino acid codons).

To make the insulin A and B chain genes, it was necessary to make 29 oligonucleotides which were assembled and joined by ligation to make a total of 181 base pairs of duplex DNA. Starting from the trimer library, the DNA fragments were made in about 3 months. The next stages, cloning and expression, took somewhat longer. DNA synthesis no longer is the rate limiting step.

Molecular cloning and expression

Figure 3 illustrates how the insulin A chain gene was assembled, cloned, and positioned at the end of β-galactosidase. Step 1 (fig. 3) was joining the small (13 base average) oligonucleotides. Because they were designed to have complementary overlaps, they assemble themselves, and were joined to give duplex DNA by the action of the T4 DNA ligase. The gene was designed to have restriction enzyme sites at each end (Eco R1 on the left and Bam H1 on the right). Step 2 was preparation of the plasmid DNA cloning vector pBR322. Preparation included treatment with Eco R1 and Bam H1 restriction enzymes, which cuts out a small piece of the plasmid and provides a site for insertion of the synthetic A gene. In step 3, the prepared plasmid and synthetic DNA are mixed and joined by T4 DNA ligase, followed by transformation of *E. coli* and molecular cloning. A clone was obtained that contained a correct insulin A gene, as verified by direct DNA sequencing. Next, a DNA fragment containing most of the *E. coli* lac operon, including the lac promoter, operator, and the first 1006 amino acid codons of β-galactosidase, was inserted (steps 4, 5, and 6; Z symbolizes the β-galactosidase gene). This led to a clone making insulin-β-galactosidase fused protein.

MUTATION CORRECTION

DNA changes directed by synthetic DNA

Several techniques have been used recently for site-localized mutagenesis of DNA in vitro [4, 8–10]. However, we think that the use of synthetic DNA provides the most specific and general approach to making directed changes in DNA [4, 10]. It should be possible to repair or create mutations, convert a gene of one species to the same gene of

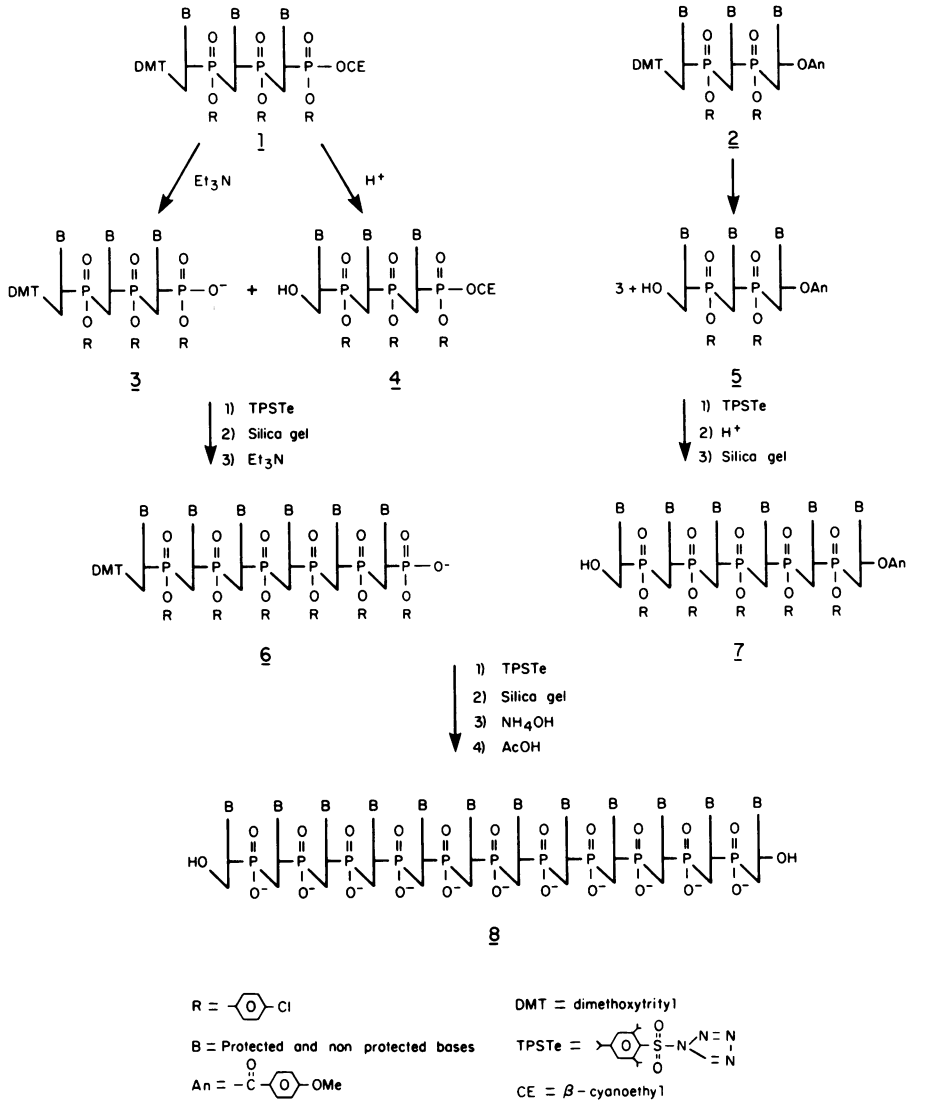

FIG. 2.—Chemical synthesis of oligodeoxyribonucleotides by improved triester method [2, 7]. Starting with nucleosides, a library of fully protected triester trimers, such as *1* and *2* are made. The type *2* trimer, with the anisole 3′-protecting group will become the 3′ end of the oligonucleotide. The type *1* trimer is bifunctional, and depending on treatment (either mild acid or base) will be either the 5′-end component or an internal sequence component. Because of the chlorophenyl protecting groups attached by ester linkage to the phosphate groups (forming phosphotriesters), the trimers and intermediate oligonucleotides (e.g., *3*, *4*, *5*, *6*, *7*) are not water soluble. Therefore, all condensations and purifications are done in nonaqueous solvents such as chloroform. Trimers can be condensed to yield hexamers (e.g., *3* + *4* yields *6*) and hexamers can be condensed to yield dodecamers (e.g., *6* + *7* yields *8*, still in fully protected triester form). The next-to-last step in a typical synthesis is the removal of all protecting groups by treatment with acetic acid and NH_4OH, generating the desired water soluble single stranded DNA fragment. The last step is a careful purification of the DNA fragment by high performance liquid chromatography.

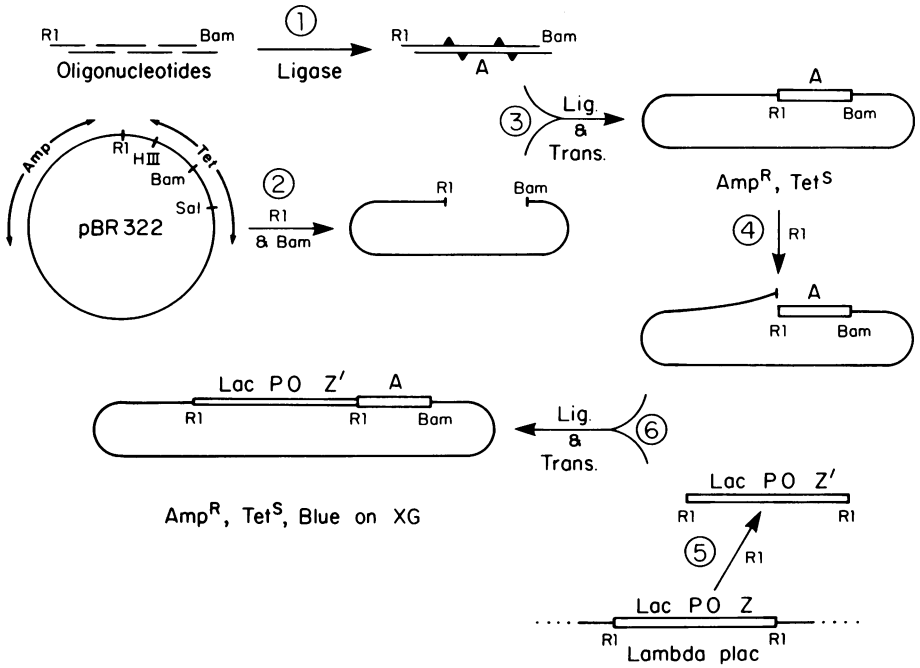

FIG. 3.—The construction of a plasmid DNA containing a synthetic insulin A chain gene inserted at the end of a β-galactosidase gene. The procedures are described in the text, and details are given in Goeddel et al. [3], so only an explanation of the symbols is given here. The symbol *A* represents the synthetic A chain gene. *pBR322* is a well-characterized plasmid cloning vector containing two antibiotic resistance genes, ampicillin (*Amp*) and tetracycline (*Tet*), and several convenient restriction endonuclease sites including Eco R1 (*R1*) and Bam H1 (*Bam*). *Lambda plac* is a lambda transducing phage carrying the entire *E. coli* operon, which includes the lac promoter (*P*), the lac operator (*O*), and the entire β-galactosidase structural gene (*Z*). There is an Eco R1 endonuclease site to the left of the operon and also one near the end of the β-galactosidase gene; thus, the lac operon DNA fragment can be readily obtained. The phenotype of the bacterial strains successfully infected with the desired plasmid are shown. For example, the A chain producing strains would be ampicillin resistant (Amp^R), tetracycline sensitive (Tet^S), and the colonies would be *Blue* on a special indicator agar called *Xg*.

another species, make genes for peptide analogues, create restriction sites, etc. Although practical applications have not yet been made, the feasibility of the approach has been demonstrated [4, 10].

Figure 4 illustrates how we have efficiently corrected a mutation in ϕX174 bacteriophage DNA, a favorable test system because the mature viral DNA is a single stranded circle, and the complete DNA sequence is known [11]. We made a synthetic primer with wild type sequence. This synthetic DNA hybridized with a one-base pair mismatch to mutant viral DNA and served as a primer for in vitro DNA replication by *E. coli* DNA polymerase. Complete heteroduplex circles were made by in vitro DNA replication, and these were used to infect *E. coli*. In vivo replication led to homoduplex progeny, a high percentage of which were converted to wild type [4]. The efficiency of the directed change is high enough that even sequence changes for which there is no method of selection can be made. Direct DNA sequencing techniques can be used to identify the converted molecule.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.