Postprint version. Original publication in:
Molecular Biology (USSR) (1985) 19, 1, 226 - 235
Molekulyarnaya Biologiya (1985) 19, 1, 267 - 277

GENETIC ENGINEERING OF PEPTIDE HORMONES

P. M. Rubtsov, B. K. Chernov, V. G. Gorbulev,
A. Sh. Parsadanyan, P. S. Sverdlova, V. V. Chupeeva,
Yu. B. Golova, N. V. Batchikova, G. S. Zhvirblis,
K. G. Skryabin, and A. A. Baev

UDC 577.212.175.3

Peptide and polypeptide hormones represent an extensive group of biologically active compounds of important significance for medicine and agriculture. In recent years genetic engineering methods have been used to create strains of microorganisms synthesizing eukaryotic proteins, including hormones and their precursors. The first stage of such developments is the isolation of DNA coding the desired product. We have accomplished the cloning of the cDNA of a number of polypeptide and peptide hormones of the pituitary of man and domestic animals. The model gene of human calcitonin has also been synthesized and cloned. The obtained genes are being used to develop methods for the microbiological synthesis of human and animal hormones.

One of the new possibilities opened by the development of genetic engineering is the production of strains of microorganisms programmed for the synthesis of polypeptides of great practical significance. The genetic-engineering production of polypeptides is a practical problem but also greatly expands the potential for basic research, since the construction of systems of expression of foreign genes in pro- and eukaryotic cells and the comparison of their effectiveness permits a more through analysis of the mechanisms of gene function and regulation. Among the products of great interest from a practical viewpoint are many peptides and polypeptide hormones. It is not surprising that the first studies of the expression of foreign genes in *Escherichia coli* were devoted to the synthesis of such hormones as somatostatin [1], growth hormone [2], insulin [3], and others [4].

The Laboratory of Functional Enzymology, Institute of Molecular Biology, Academy of Sciences of the USSR, has for a number of years conducted studies of the cloning and expression in heterologous systems of the genes of polypeptide and peptide hormones. The present article will describe some of the results of these investigations.

The successful realization of a project for the genetic-engineering production of a needed peptide or polypeptides requires the solution of a whole series of problems [5]. The first stage is the production of DNA coding of the intact product. There are two means for solving this problem: The isolation of a natural DNA fragment or the chemicoenzymatic synthesis of polydeoxynucleotide. In the first case the method of reverse mRNA transcription is generally used, resulting in the production of a DNA fragment containing complete information on the polypeptide structure. The chemicoenzymatic method permits the synthesis of an artificial sequence of nucleotides. Being more laborious, this approach at the same time possesses the advantage that it permits the introduction of various changes into the coding sequence. The experimentor can thus vary the structure of the obtained product, creating analogs of the natural polypeptide or peptide.

Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow. Translated from Molekulyarnaya Biologiya, Vol. 19, No. 1, pp. 267-277, January-February, 1985.

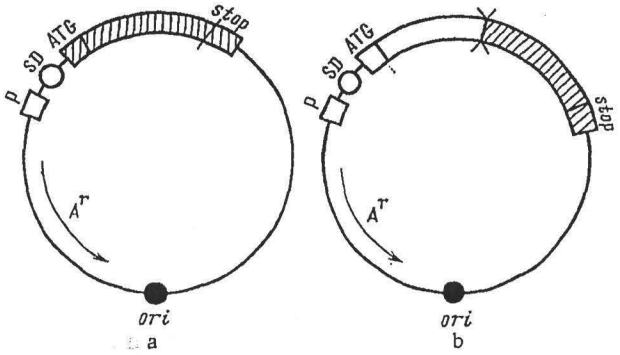

Fig. 1. Scheme of organization of plasmid vectors ensuring direct expression (a) and expression of hybrid protein (b). P) Promotor; SD) Shine–Dalgarno sequence; ATG) initiating codon; Stop) terminating codon; Ori) replication initiation site; A^r) antibiotic resistance gene. Region occupied by nucleotide sequence coding desired product is hatched.

The next problem is to ensure the expression of the cloned gene in the cell. There are two approaches to solving this problem. The first is the synthesis of a hybrid protein and the subsequent excision from it of the necessary polypeptide; the second is the direct expression of this polypeptide. The first technique was first used to obtain the hormone somatostatin [1]; it is usually used for the synthesis of short peptides that are rapidly degraded in microorganismic cells. The second method is used in the microbiological synthesis of polypeptides that are stable within the cell.

Figure 1 presents schematically the organization of vectors ensuring the expression of hybrid proteins and the direct expression of foreign polypeptides in bacterial cells [5]. As apparent from the scheme, during direct expression the gene is placed under the control of the bacterial promotor; inserted in front of the initiating codon is a region providing for translation initiation (SD, the Shine–Dalgarno sequence). For the expression of the hybrid protein the sequence coding the desired product is attached (with the preservation of the read frame) to the gene for any buffer protein. The β-galactosidase gene is most frequently used for this in the case of *E. coli* [1, 6-10]. At the same time, the initiation of transcription and translation is ensured by the signal elements of the bacterial gene.

The last stage of development is the isolation of the synthetic product from the cell, its purification and characterization. In general, the development of methods for the microbiological synthesis of valuable polypeptides is a complex problem, requiring the use of modern methods of molecular biology and genetics, biochemistry, bioorganic chemistry, microbiology, and other related disciplines. The present article will illustrate the separate stages of this approach for the production of certain peptide and polypeptide hormones as exemplified by studies at the Laboratory.

Cloning of cDNA of Polypeptide and Peptide Hormones. A universal means for obtaining DNA coding of one or another polypeptide is the method of cloning cDNA synthesized on the mRNA of the given polypeptide. A decisive role in the success of cloning is played by the selection of the source from which the mRNA is isolated. Enriched mRNA preparations can be obtained by the use for mRNA isolation of specialized tissues or tumors producing increased quantities of the desired product. The synthesis of DNA complementary to this mRNA and its cloning permit the creation of banks of clones enriched with the necessary sequences. This approach was used in our laboratory to construct clone banks containing DNAs complementary to the template RNAs of human and domestic-animal pituitaries. The guanidine thiocyanate method [11] was used for the isolation of mRNA from pituitaries (pituitary adenomas obtained at the Institute of Neurosurgery, Academy of Medical Sciences of the USSR, during the surgery of patients suffering acromegaly were used as the starting material); the poly(A)-containing RNA fraction was isolated by chromatography on a column with oligo(dT) cellulose [12]; double-stranded cDNA was synthesized by the standard method [13] and fractionated with the help of electrophoresis in polyacrylamide gel. Earlier it was shown that the complete length of the mRNA of mammalian growth prehormone is 800-900 nucleotides, not counting the poly(A) sequences [14]; therefore, an additional enrichment of the mixture with full-sized cDNA copies of growth

Fig. 2. Hybridization of colonies on filter: a, b) bank of bovine pituitary cDNA. Trials) cloning of cDNA of human growth prehormone (a) and rat prolactin (b). c) Bank of swine pituitary cDNA. Trial) cloned cDNA of bovine growth prehormone.

prehormone was attained by extracting a fraction of about 800 bp size from the gel. This fraction was cloned in the HindIII region of the plasmid pBR322 using the linker technique [15].

The next stage of the work was the isolation from the banks of individual clones containing hormone genes. A clone was first isolated containing a full-sized copy of the cDNA of human growth hormone precursor. This was done using the methods of hybridization of colonies with highly labeled cDNA [16], restrictase analysis, and the partial sequencing of the insertions [17]. In one plasmid detected in this way (phGH18) the cDNA insertion consisted of 809 bp and contained a 5'-nontranslatable region (30 bp), sequences coding a signal peptide (78 bp) and mature growth hormone (573 bp), a 3'-nontranslatable region (108 bp), and a poly(dA) sequence (20 bp). The primary structure of the cDNA we cloned was indistinguishable from that published earlier [18].

An analysis of the banks containing insertions of the cDNA of domestic-animal pituitaries by hybridization of the colonies with labeled cloned cDNA made it possible to detect clones bearing plasmids with insertions of the cDNA of bovine and swine growth prehormones (Fig. 2a, c). The plasmid pbGH18, containing an insertion of bovine growth prehormone cDNA of about 830 bp size, was investigated in detail using restrictase analysis and subsequent sequencing (Figs. 3b and 4). The insertion from this plasmid included 33 bp of the 5'-nontranslatable region, 651 bp of the prehormone-coding sequence, 104 bp of the 3'-nontranslatable region, and a poly(dA) sequence. The cloned cDNA of bovine growth prehormone from plasmid pbGH18 had four substitutions in the coding part and an insertion of two base pairs in the 5'-nontranslatable region compared with the nucleotide sequence published by Miller et al. [19]. All substitutions in the coding region were in the third position of the codons and did not influence the amino acid sequences of the hormone.

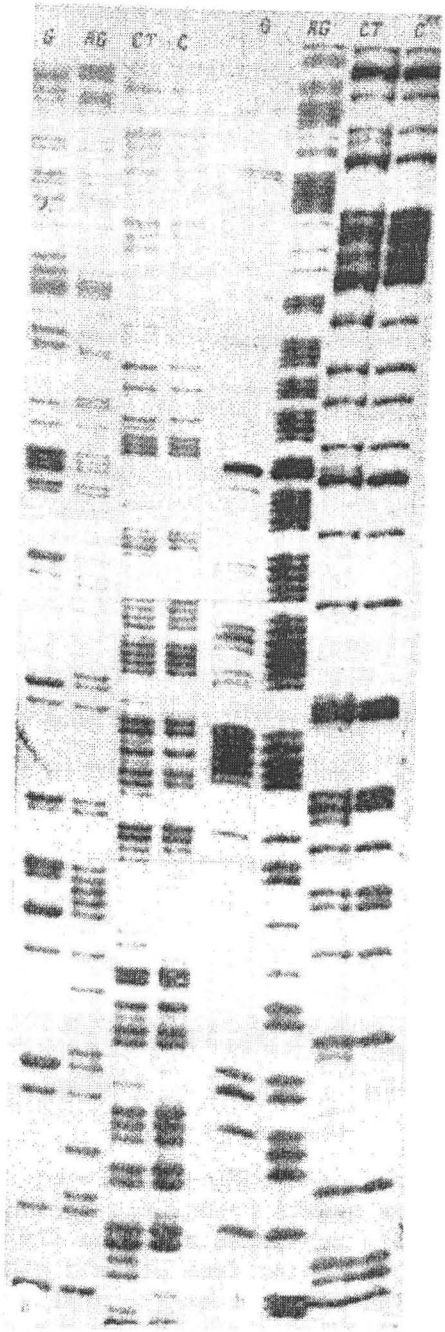

Fig. 3. Autoradiograms of structural gels obtained after the sequencing of the cloned cDNAs of human (a) and bovine (b) growth prehormones.

Also isolated from the bank of pituitary cDNA clones was the plasmid pbLT19, containing the cDNA of the precursor of the adrenocorticotropic hormone (ACTH) and β-lipoprotein. The cloned cDNA, as demonstrated by the determination of its complete nucleotide sequence (Figs. 5 and 6), represents an incomplete copy of the mRNA coding the polypeptide precursor of ACTH and β-lipoprotein. The first 47 base pairs of the insertion represent the result of a reorganization that probably occurred during cloning; nevertheless, the plasmid pbLT19 contains sequences coding all known peptides formed from the initial polypeptide precursor.

Also isolated from the obtained banks of agricultural-animal clones were plasmids with insertions of prolactin-precursor DNA (Fig. 2b). Thus, the cloning of cDNA is a sufficiently effective means of obtaining starting material for the subsequent construction of recombinant plasmids ensuring hormone expression in microorganismic cells.

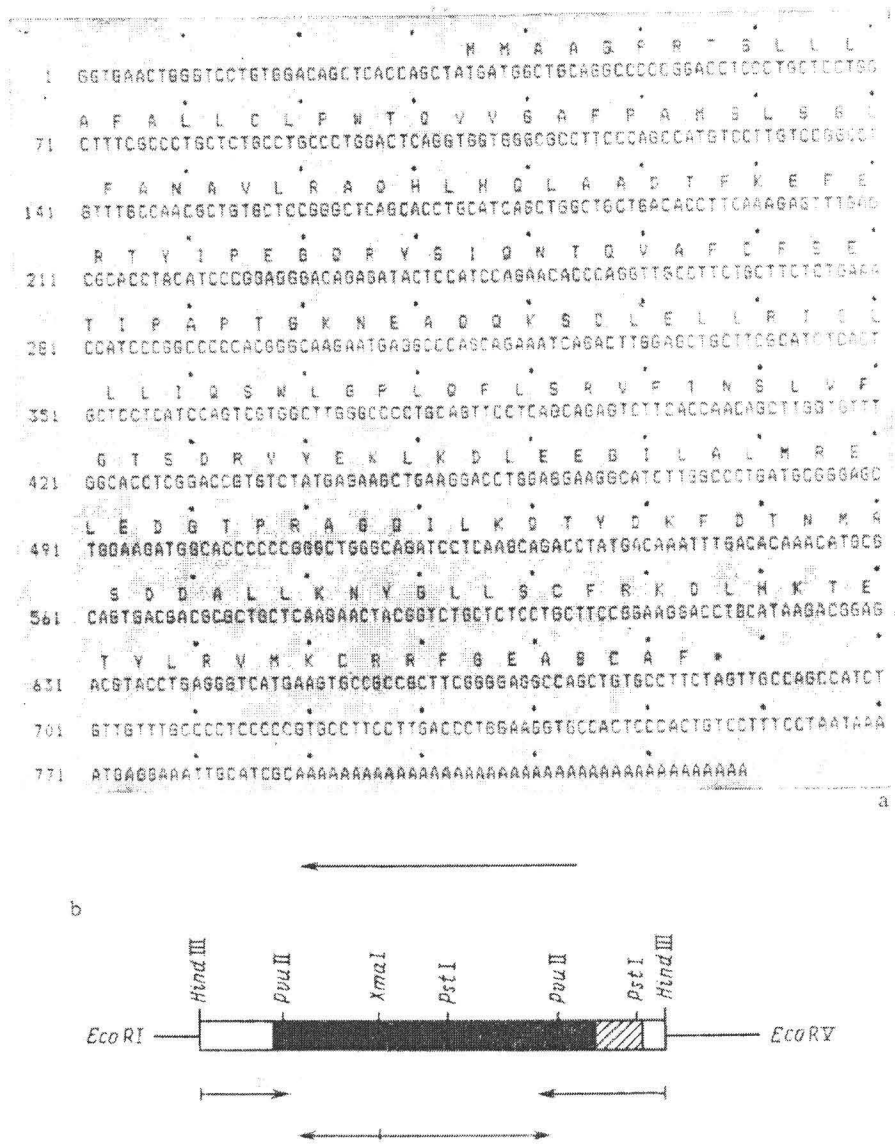

Fig. 4. a) Nucleotide sequence of cloned cDNA and amino acid sequence of bovine growth prehormone derived from it. Differences compared with published sequence [19] indicated. b) Restrictase map of insertion from plasmid pbGH18. Arrow shows orientation of insertion. Sequence coding signal peptide is hatched; sequence coding mature hormone is darkened.

Chemicoenzymatic Synthesis of Peptide-Hormone Genes. An alternative means of obtaining DNA fragments coding the required product is the chemicoenzymatic synthesis of polydeoxynucleotides, the strategy of which was developed by Khorana's group in the late 1960s [20, 21]. At the present time the phosphotriester method is the most widely used for the synthesis of oligodeoxynucleotides [22]. In our laboratory this method was used for the synthesis of a model gene of human calcitonin — a thyroid gland peptide hormone consisting of 32 amino acid residues [23] and regulating calcium metabolism in the organism [24]. Taken as a basis for planning the synthesis of the gene was the natural nucleotide sequence published in 1982 [25]. The following changes were introduced into this sequence: 1) The triplet ATG, coding methionine-8, was substituted by the valine codon, GTT; 2) the point substitutions of G for T and C for A were made in positions 38 and 41, respectively, of the nucleotide sequence (Fig. 7). Methionine was substituted by valine so that the whole product might be isolated from hybrid proteins by means of their treatment with cyanobromide. This substitution should not affect the immunologic properties and biological activity of calcitonin, since according to the current concepts of the structural organization of this hormone its active part is the fragment 10-32 [26]. Point substitutions in the 38th and 41st positions of the nucleo-

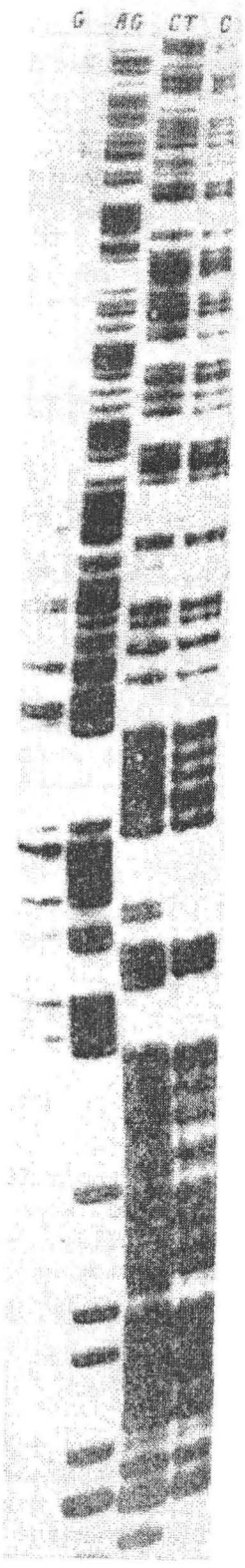

Fig. 5. Autoradiogram of structural gel obtained during the sequencing of cloned cDNA of ACTH–β-lipotropin precursor.

tide sequence, which do not alter the structure of the coded peptide, result in the appearance of a restrictase KpnI cleavage site (GGTACC), which greatly expands the possibilities of subsequent manipulations with the aim of constructing different variants of the calcitonin gene.

The general structure of the synthetic gene is presented in Fig. 7. It includes the translation-initiating codon ATG, the coding nucleotide sequence, and the two terminating codons TGA and TAG. The ends of the genes are provided with the 5'-protruding single-stranded regions 5'-AATT-3' and 5'-GATC-3', which permits the insertion of the gene into the vector at the EcoRI and BamHI restrictase sites.

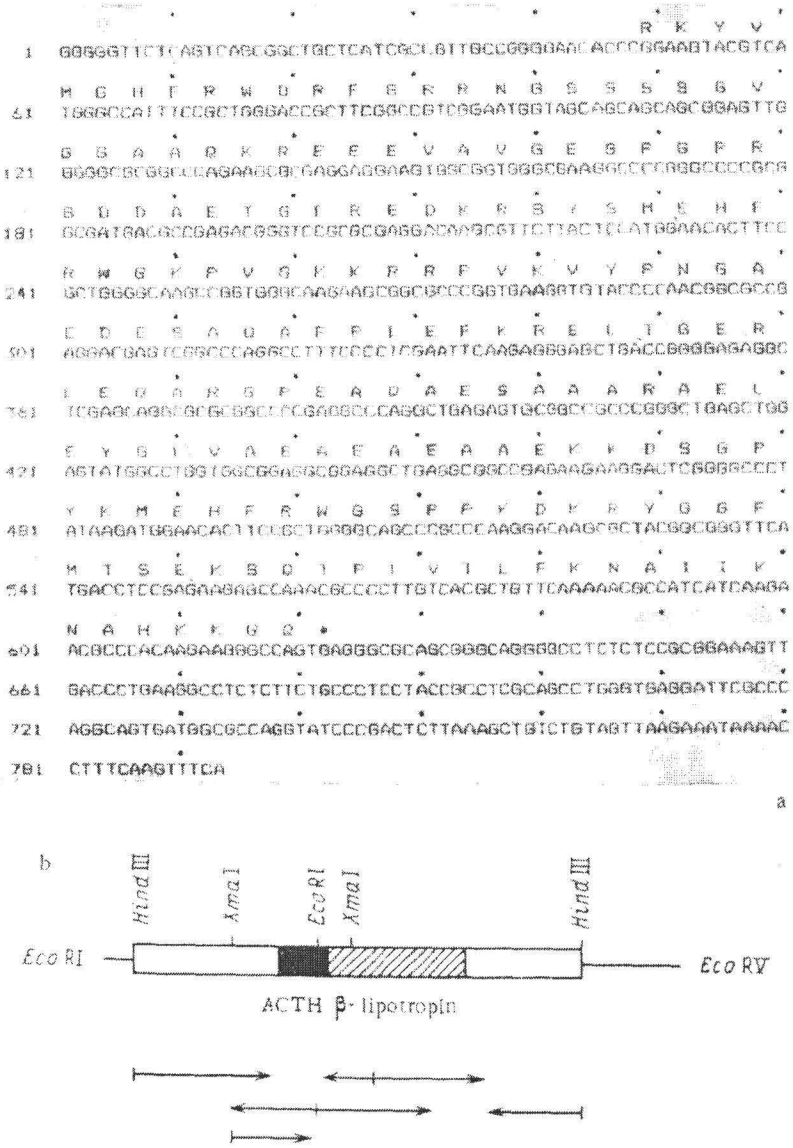

Fig. 6. a) Nucleotide sequence of cloned cDNA and amino acid sequence of bovine ACTH-β-lipotropin precursor derived from it. b) Restrictase map of insertion from plasmid pbLT19. Sequence coding β-lipotropin is hatched; sequence coding ACTH is darkened.

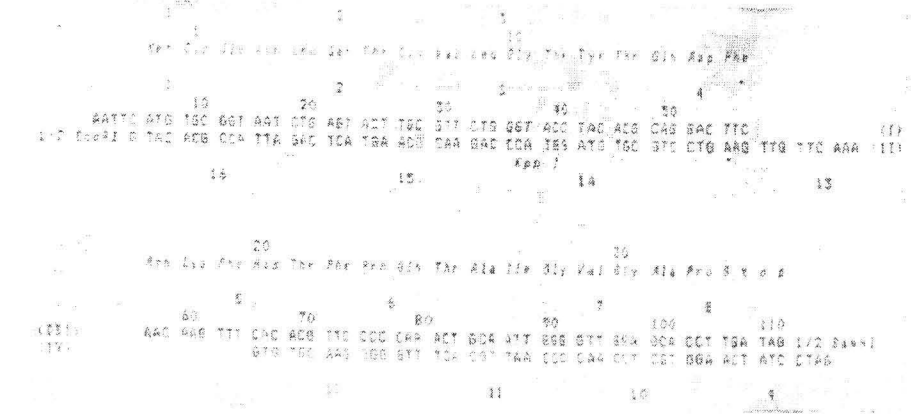

Fig. 7. Scheme of organization of human-calciton in synthetic gene and its corresponding amino acid sequence. 1-16) Initial synthetic oligonucleotides; I-IV) intermediate products of ligase splicing.

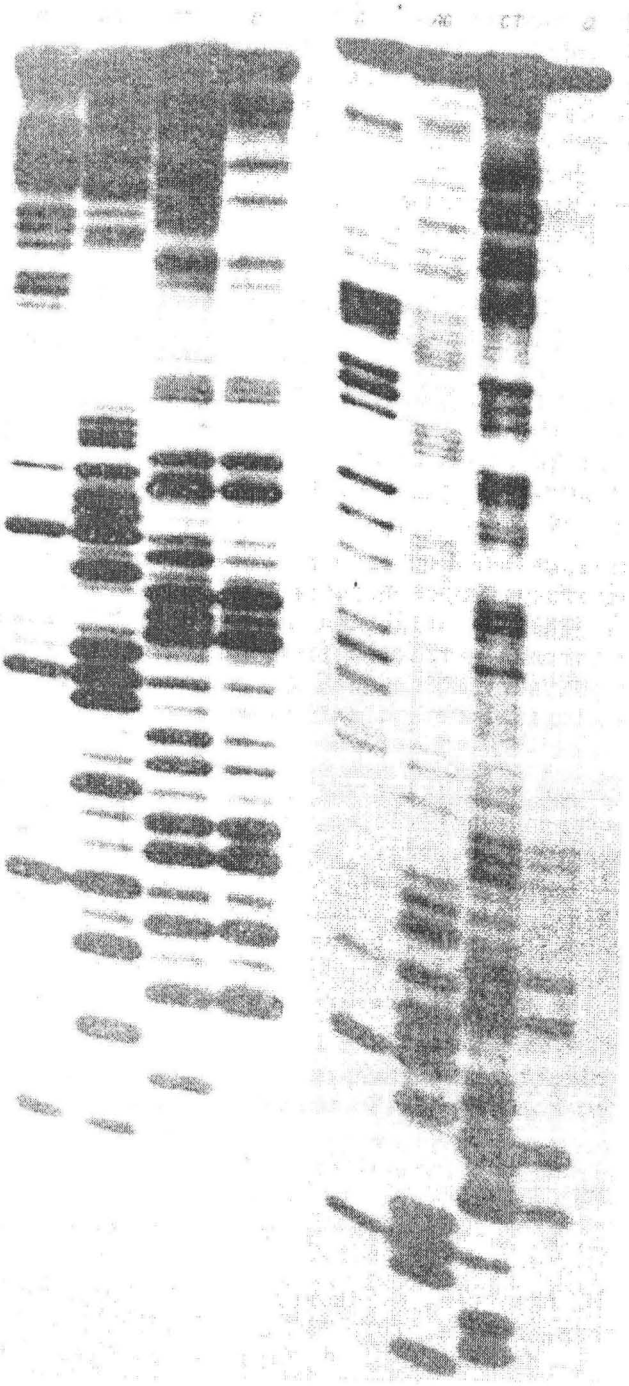

Fig. 8. Autoradiogram of structural gels obtained during the sequencing of cloned human-calcitonin model gene.

In planning the synthesis, the nucleotide sequence of the gene was broken up into 16 fragments (eight fragments each in the translated and complementary chains) so that the overlap of complementary sequences would comprise from six to nine nucleotides when the fragments were joined. The enzymatic splicing of the synthetic gene from oligonucleotide fragments synthesized by the block triester method in solution [27] with certain modifications was performed in several stages. Chains I-IV were first assembled (Fig. 7), which, after isolation from the incubation mixtures and verification of their nucleotide sequence, were spliced to obtain the synthetic model gene of human calcitonin. The structure of the complete gene was verified by the Maxam–Gilbert method [17]. We cloned the synthetic human-calcitonin model gene in the vector pBR322 (Fig. 8) and will use it in the Laboratory's subsequent studies.

Expression of Growth Hormone Gene in *E. coli* Cells and Characterization of Product. The next stage in the isolation of polypeptides by genetic-engineering methods is the expression of the gene in the recipient cell and the isolation and characterization of the synthesized product. One of the first polypeptide products obtained by genetic-engineering is human

growth hormone. The firm Genentech Inc. (USA) has synthesized a "quasisynthetic" hormone gene consisting of an artificial DNA sequence coding 23 amino acid residues of the N-terminal part of the hormone and a cDNA fragment bearing information on the rest of the molecule [2]. The introduction in front of the gene of a DNA fragment containing the lac UV5 promotor and the translation initiation signal of the lac Z gene of *E. coli* ensured the effective expression of the gene and resulted in the synthesis of hormone in a quantity reaching 185,000 molecules per cell [2]. Later, the method of the direct expression of polypeptides in *E. coli* was used to obtain many proteins important in medicine and agriculture, such as interferons [28-30], blood proteins [31], enzymes [32], etc.

We have employed this method to construct *E. coli* strains to produce human growth hormone. Using as the starting material the cloned cDNA of human growth prehormone, we removed from it the 5'-nontranslatable region and the sequence coding the signal peptide and introduced in its place a translation initiation codon. The reconstructed human growth hormone gene that was obtained was placed in *E. coli* under the control of various promotors. The *E. coli* strains transformed by such plasmids synthesized human growth hormone, as demonstrated by double radial immunodiffusion in agar. The highest level of synthesis comprised (by an approximate estimate) more than 550,000 molecules per cell.

An analysis was also made of the physicochemical and biological properties of the hormone synthesized in the cells of the producer strain. Growth hormone was extracted from the biomass following described schemes [2, 33, 34]. The investigation of a highly purified hormone preparation showed that it is identical to hormone extracted from cadaver pituitaries according to data electrophoresis in denaturing polyacrylamide gel, gel filtration on Sephadex G-100, double radial immunodiffusion, and the biological determination of activity using the tibia test [35].

Human growth hormone is one of the key hormones regulating various aspects of metabolism. It possesses fat-mobilizing, renotropic, and lactogenic activities, acts on carbohydrate and mineral metabolims, and increases tissue regeneration [36, 37]. Because of the shortage of growth hormone preparation, it is used in medicine only to treat pituitary dwarfism caused by a disturbance of pituitary hormone biosynthesis. Data are available on the positive effect of growth hormone in the treatment of complex fractures, burns, and ulcers [37]. The development of methods for the genetic engineering of strains producing human growth hormone opens prospects for the increased production of the medical preparation of this hormone and its clinical use.

The cloning of growth-hormone cDNA from agricultural animals performed in the Laboratory is an important step toward the production of these compounds by means of microbiological synthesis.

LITERATURE CITED

1. K. Itakura, T. Hirose, R. Crea, A. D. Riggs, H. L. Heyneker, F. Bolivar, and H. Boyer, Science, 198, 1056-1063 (1977).
2. D. V. Goeddel, H. L. Heyneker, T. Hozumi, R. Arentzen, K. Itakura, D. G. Yansura, M. J. Ross, G. Moizzari, R. Crea, and P. H. Seeburg, Nature, 281, 544-548 (1979).
3. D. V. Goeddel, D. G. Kleid, F. Bolivar, H. L. Heyneker, D. G. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, and A. D. Riggs, Proc. Natl. Acad. Sci. USA, 76, 106-110 (1979).
4. K. G. Skryabin, Zh. Vses. Khim. Ova. im. D. I. Mendeleeva, 29, 195-204 (1984).
5. R. Wetzel and D. V. Goeddel, in: The Peptides. Analysis, Synthesis, Biology, E. Gross and J. Meienhofer (eds.), Vol. 5, Academic Press, New York–London (1983), pp. 1-64.
6. S. I. Gorodetskii, T. V. Kapelinskaya, A. F. Lisenkov, A. G. Slyusarenko, and N. P. Dubinin, Dokl. Akad. Nauk SSSR, 250, 208-212 (1980).
7. V. P. Kumarev, M. I. Rivkin, V. S. Bogachev, L. V. Baranova, V. M. Merkulov, V. N. Rybakov, E. I. Solenov, and V. I. Fedorov, Dokl. Akad. Nauk SSSR, 252, 1506-1510 (1980).
8. M. F. Chemyakin, A. V. Chestukhin, G. M. Dolganov, E. M. Khodkova, G. S. Monastyrskaya, and E. D. Sverdlov, Nucleic Acids Res., 8, 6163-6174 (1980).
9. J. Shine, I. Fettes, N. C. Y. Lan, J. L. Roberts, and J. D. Baxter, Nature, 285, 456-461 (1980).
10. S. Tanaka, T. Oshima, K. Ohsue, T. Ono, S. Oikawa, I. Takano, T. Noguchi, N. Kangawa, N. Minamino, and H. Matsuo, Nucleic Acids Res., 10, 1741-1744 (1982).
11. J. M. Chirgwin, A. E. Przybyla, R. J. McDonald, and W. J. Rutter, Biochemistry, 18, 5294-5299 (1979).

12. H. Aviv and P. Leder, Proc. Natl. Acad. Sci. USA, 69, 1408-1412 (1972).
13. P. H. Seeburg, J. Shine, J. A. Martial, and J. D. Baxter, Nature, 270, 486-494 (1977).
14. W. G. Roscam and F. Rougeon, Nucleic Acids Res., 7, 305-320 (1979).
15. F. Bolivar, R. L. Rodriguez, P. Y. Greene, M. C. Betlach, H. L. Heyneker, and H. W. Boyer, Gene, 2, 95-113 (1977).
16. M. Grunstein and D. S. Hogness, Proc. Natl. Acad. Sci. USA, 72, 3961-3965 (1975).
17. A. M. Maxam and W. Gilbert, Proc. Natl. Acad. Sci. USA, 74, 560-564 (1977).
18. J. A. Martial, R. A. Hallewell, J. D. Baxter, and H. M. Goodman, Science, 205, 602-607 (1979).
19. W. L. Miller, J. A. Martial, and J. D. Baxter, J. Biol. Chem., 255, 7521-7524 (1980).
20. K. L. Agarwal, H. Buchi, M. H. Caruthers, N. Gupta, U. L. Raj Brandary, J. H. Van de Sande, V. Sgaramella, H. Weber, and T. Yamada, Nature, 227, 27-34 (1970).
21. V. Sgaramella, J. H. Van de Sande, and H. G. Khorana, Proc. Natl. Acad. Sci. USA, 67, 1468-1475 (1970).
22. K. Itakura, N. Katagiri, C. P. Bahl, R. H. Wightman, and S. A. Narang, J. Am. Chem. Soc., 97, 7327-7332 (1975).
23. R. Neher, B. Riniker, W. Rittel, and H. Zuber, Helv. Chim. Acta, 51, 1900-1905 (1968).
24. I. MacIntyre, T. R. Arnett, D. J. Brown, F. Galan Galan, S. I. Girgis, R. M. Rogers, E. Spanos, J. C. Stevenson, and Q. Bone, in: Molecular Endocrinology, I. MacIntyre and M. Szelke (eds.), Vol. 1, Elsevier/North-Holland Biomedical Press, Amsterdam-New York (1979), pp. 193-201.
25. R. K. Craig, L. Hall, M. R. Edbrooke, J. Allison, and I. MacIntyre, Nature, 295, 345-347 (1982).
26. F. M. Dietrich and W. Rittle, Nature, 225, 75-76 (1970).
27. E. Ohtsuka, M. Ikehara, and D. Söll, Nucleic Acids Res., 10, 6553-6570 (1982).
28. T. Taniguchi, L. Guarente, T. M. Roberts, D. Kimelman, J. Douhan III, and M. Ptashne, Proc. Natl. Acad. Sci. USA, 77, 5230-5233 (1980).
29. D. V. Goeddel, H. M. Shepard, E. Yelverton, D. Leung, R. Crea, A. Sloma, and S. Pestka, Nucleic Acids Res., 8, 4057-4074 (1980).
30. Yu. A. Ovchinnikov, E. D. Sverdlov, S. A. Tsarev, E. M. Khodkova, G. S. Monastyrskaya, I. S. Solomatina, V. A. Efimov, O. G. Chakhmakhcheva, V. D. Solov'ev, V. L. Kuznetsov, V. M. Zhdanov, A. S. Novokhatskii, and A. D. Aspetov, Dokl. Akad. Nauk SSSR, 265, 238-241 (1982).
31. R. M. Lawn, J. Adelman, S. C. Bock, A. E. Franke, C. M. Houck, R. C. Najarian, P. H. Seeburg, and K. L. Wion, Nucleic Acids Res., 9, 6103-6128 (1981).
32. D. Pennica, W. E. Holmes, W. J. Kohr, R. N. Harkins, G. A. Vehar, C. A. Ward, W. E. Bennett, E. Yelverton, P. H. Seeburg, H. L. Heyneker, D. V. Goeddel, and O. Collen, Nature, 301, 214-221 (1983).
33. T. A. Bewley and C. H. Li, Enzymology, 43, 73-166 (1975).
34. M. Schleyer and K. H. Voigt, Hoppe-Seyler's Z. Physiol. Chem., 358, 1557-1564 (1977).
35. P. M. Rubtsov, A. Sh. Parsadanyan, P. S. Sverdlova, V. V. Chupeeva, L. V. Lashas, K. G. Skryabin, and A. A. Baev, Dokl. Akad. Nauk SSSR, 276, 762-764 (1984).
36. N. A. Yudaev, S. A. Afinogenova, and A. A. Bulatov, Biochemistry of Hormones and Hormonal Regulation [in Russian], Nauka, Moscow (1976), pp. 63-71.
37. L. Lashas and D. Lashene, Human Somatotropin [in Russian], Mokslas, Vil'nyus (1981).