Proc. Natl. Acad. Sci. USA
Vol. 76, No. 1, pp. 106–110, January 1979
Biochemistry

Expression in *Escherichia coli* of chemically synthesized genes for human insulin

(plasmid construction/*lac* operon/fused proteins/radioimmunoassay/peptide purification)

DAVID V. GOEDDEL*†, DENNIS G. KLEID*, FRANCISCO BOLIVAR*, HERBERT L. HEYNEKER*, DANIEL G. YANSURA*, ROBERTO CREA*‡, TADAAKI HIROSE‡, ADAM KRASZEWSKI‡, KEIICHI ITAKURA‡, AND ARTHUR D. RIGGS†‡

*Division of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080; and ‡Division of Biology, City of Hope National Medical Center, Duarte, California 91010

Communicated by Ernest Beutler, October 3, 1978

ABSTRACT Synthetic genes for human insulin A and B chains were cloned separately in plasmid pBR322. The cloned synthetic genes were then fused to an *Escherichia coli* β-galactosidase gene to provide efficient transcription and translation and a stable precursor protein. The insulin peptides were cleaved from β-galactosidase, detected by radioimmunoassay, and purified. Complete purification of the A chain and partial purification of the B chain were achieved. These products were mixed, reduced, and reoxidized. The presence of insulin was detected by radioimmunoassay.

Recently improved methods of DNA chemical synthesis, combined with recombinant DNA technology, permit the design and relatively rapid synthesis of modest-sized genes that can be incorporated into prokaryotic cells for gene expression. The feasibility of this general approach was first demonstrated by the synthesis, and expression in *Escherichia coli*, of a gene for the mammalian peptide somatostatin (1).

Following the precursor protein approach used for somatostatin (1), the experimental design for this work was such that the insulin peptide chains would be made *in vivo* as short tails joined by a methionine to the end of β-galactosidase. After synthesis, the insulin chains, which contain no methionine, can be cleaved off efficiently by treatment with cyanogen bromide. We deliberately chose to construct two separate bacterial strains, one for each of the two peptide chains of insulin: the 21-amino-acid A chain and the 30-amino-acid B chain. In native insulin, the two chains are held together by two disulfide bonds, and methods have been available for years for joining the chains correctly, *in vitro*, by air oxidation (2). The efficiency of correct joining has been variable and often low. However, by using *S*-sulfonated derivatives and an excess of A chain, 50–80% correct joining has been obtained (3).

The synthetic plan and chemical synthesis of the DNA fragments coding for the A and B chains of human insulin were described in a previous paper (4) and were summarized in Fig. 1 of that paper. In this communication, we describe the assembly and cloning of the genes for the A and B chains, their insertion into the carboxy terminus of the *E. coli* β-galactosidase structural gene, the expression and purification of the separate A and B chains, and their joining to form native human insulin.

MATERIALS AND METHODS

Bacterial Strains. *E. coli* K-12 strain 294 (*endA*, *thi*$^-$, *hsr*$^-$, *hsm*$_k^+$) (5) was provided by K. Backman. *E. coli* K-12 strain D1210, a *lac*$^+$ ($i^Qo^+z^+y^+$) derivative of HB101, was constructed by J. Betz and J. Sadler and obtained from J. Sadler.

Enzymes and DNA Preparations. T4 DNA ligase and T4 polynucleotide kinase were purified as described (6). Restriction endonuclease *Eco*RI was purified by the procedure of Greene *et al.* (7); *Hin*dIII was purified by a method developed by D. Goeddel (unpublished). Restriction endonuclease *Bam*HI was purchased from Bethesda Research (Rockville, MD); *E. coli* alkaline phosphatase was purchased from Worthington.

Plasmids, including pBR322 (8), were isolated by a published procedure (9) with some modifications. The chemical synthesis of the deoxyoligonucleotides (figure 1 of ref. 4) has been described (4). λ*plac5* DNA was isolated as described (10).

The following reaction buffers were used: kinase buffer, 60 mM Tris·HCl, pH 8/15 mM 2-mercaptoethanol/10 mM $MgCl_2$; ligase buffer, 20 mM Tris·HCl, pH 7.5/10 mM dithiothreitol/10 mM $MgCl_2$; *Bam*HI buffer, 20 mM Tris·HCl, pH 7.5/7 mM $MgCl_2$/2 mM 2-mercaptoethanol; *Eco*RI-*Hin*dIII buffer, *Bam*HI buffer containing 50 mM NaCl; and phosphatase buffer, 50 mM Tris·HCl, pH 8/10 mM $MgCl_2$.

Assembly of Insulin Genes. The assembly of the right (BB) half of the B-chain gene (see figure 1 of ref. 4) will be described in detail. Oligonucleotides B_2–B_9 were phosphorylated individually. Fifty microcuries of [γ-^{32}P]ATP (≈2000 Ci/mmol, New England Nuclear) was evaporated to dryness in a 1.5-ml polypropylene tube, then incubated with the oligonucleotide (10 μg) and 8 units of T4 polynucleotide kinase in 60 μl of kinase buffer. After 20 min at 37°C, 10 nmol of ATP and 10 units of T4 kinase were added and the reaction was continued for an additional hour. The kinase was inactivated by heating at 90°C for 5 min.

Phosphorylated fragments B2, B3, B6, and B7 (2.5 μg each) were combined with 2.5 μg of 5′-OH fragment B1 and dialyzed for 2 hr against 1 liter of ligase buffer. ATP was added to a concentration of 0.2 mM, the reaction mixture (60 μl) was cooled to 12°C, and T4 ligase (50 units) was added. A separate ligation reaction involving phosphorylated fragments B4, B5, B8, and B9 and unphosphorylated B10 was performed identically. After 12 hr at 12°C, the two ligation reaction mixtures were combined, additional T4 ligase (40 units) was added, and the mixture was incubated at 12°C for 4 hr. The mixture was extracted with phenol/chloroform and precipitated with ethanol, and the DNA fragments were purified by electrophoresis on a 10% acrylamide gel (11). The most slowly migrating band was sliced from the gel and the DNA was extracted (11).

A similar procedure, with the following exceptions, was used to assemble the left (BH) half of the B-chain gene. All eight

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: BB, left half of insulin B gene; BH, right half of insulin B gene; A(SSO_3^-), *S*-sulfonated derivatives of the insulin A-chain peptide; B(SSO_3^-), *S*-sulfonated derivatives of the insulin B-chain peptide.

† To whom reprint requests should be addressed.

oligonucleotides (H1–H8 in figure 1 of ref. 4, 30 μg each) were phosphorylated. Therefore, after complete ligation and before purification by gel electrophoresis, the reaction mixture was treated with 400 units of *Eco*RI and 400 units of *Hin*dIII for 2 hr at 37°C. The BH band migrating at 46 base pairs was eluted from a 10% acrylamide gel.

The procedure used to construct the A-chain gene was also similar to that described for the BB fragment. The only major difference was that, after assembly, the 5′ ends of the complete A gene were phosphorylated.

Construction and Characterization of *lac*–Insulin Hybrid Plasmids. The BB fragment was cloned as follows: 1 μg of pBR322 was treated with 5 units of *Bam*HI in *Bam*HI buffer for 1 hr at 37°C. After addition of NaCl to 50 mM, *Hin*dIII (5 units) was added and the reaction was continued for 1 hr. The enzymes were inactivated by heating at 70°C for 10 min. The prepared pBR322 was ligated to the BB fragment for 3 hr at 12°C in 25 μl of ligase buffer (containing 0.16 mM ATP) by using 20 units of T4 ligase. Half of the resulting DNA was used to transform *E. coli* 294 by a published procedure (12). The BH fragment and the A-chain gene were cloned similarly, with the appropriate restriction endonucleases to cut pBR322.

Construction of the plasmids for expression of the synthetic insulin genes is described in the legend of Fig. 1. The separate chains in insulin are biologically inactive (2) and were synthesized attached to large precursor proteins. Therefore, the containment level of P2-EK1, recommended by the National Institutes of Health guideline, was used.

DNA Sequences. The method of Maxam and Gilbert (11) was used to determine DNA sequences. Sequence data are not included, but will be provided upon request.

Preparation of Insulin Reagents. Porcine and bovine insulin were purchased from Sigma. The *S*-sulfonated derivatives (SSO_3^-) of their A and B chains were prepared and purified as described (13). ^{35}S-Labeled A(SSO_3^-) and B(SSO_3^-) were prepared similarly except that 5 mCi of sodium [^{35}S]sulfite (69 mCi/mol, New England Nuclear) was substituted for unlabeled sodium sulfite. After separation of the chains, the specific activity was 92,000 cpm/μg and 32,000 cpm/μg, respectively, for A and B chains. The radioimmunoassay for the insulin chains is described in the legend of Fig. 3.

Purification of B Chain of Human Insulin. *E. coli* D1210/pIB1 was grown to late logarithmic phase in 7 liters of LB medium (10) containing 20 mg of ampicillin per liter. Isopropyl-β-D-thiogalactoside was added to a final concentration of 2 mM, and the cells were grown for one more doubling. Wet cell paste (24 g) was suspended in 30 ml of BB buffer (10) and cells were lysed by one passage through a French press at 4000 lb/inch^2 (27.6 MPa). The cell debris was pelleted by centrifugation at 15,000 rpm for 30 min. The pellet was dissolved in 40 ml of 6 M guanidinium chloride/1% 2-mercaptoethanol and centrifuged at 30,000 rpm for 1 hr. The supernatant was dialyzed overnight against 20 liters of H_2O. The precipitate, containing about 1 g of protein, was dissolved in 25 ml of 70% formic acid. Cyanogen bromide (1.3 g) was added and the mixture was allowed to react overnight at room temperature. Formic acid and cyanogen bromide was removed by rotary evaporation and the residue was dissolved in 50 ml of 8 M guanidinium chloride. *S*-Sulfonated derivates of the peptide mixture were prepared by adding 1 g of sodium tetrathionate and 2 g of sodium sulfite, adjusting the pH to 9 with NH_4OH, and stirring the mixtures at room temperature for 24 hr. The pH was then adjusted to 5 with acetic acid and the mixture was dialyzed twice against 3 liters of H_2O. The resulting white precipitate (≈0.6 g of protein) was pelleted by centrifuging at 10,000 rpm for 10 min.

Purification of A Chain of Human Insulin. *E. coli* 294/pIA1 was grown to A_{550} of 2 in 5 liters of LB medium containing 20 mg of ampicillin per liter. This strain is constitutive for β-galactosidase and so was not induced. The 15 g (wet weight) of cells obtained were processed by the same procedure used for the B chain up through the preparation of the *S*-sulfonated derivatives. After the pH was adjusted to 5 and the mixture was dialyzed against H_2O to an ionic strength of about

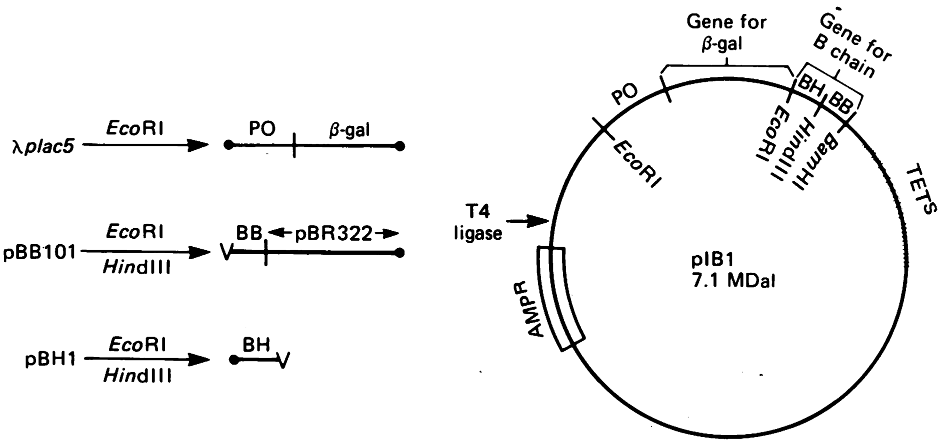

FIG. 1. Construction of *lac*–insulin plasmids. pBB101 (2 μg) (pBR322 containing the BB sequence) was treated with *Eco*RI and *Hin*dIII (20 units each), and the large fragment was purified on a 10% acrylamide gel. pBH1 (8 μg) (pBR322 containing BH sequence) also was treated with *Eco*RI and *Hin*dIII, and the small fragment was purified on a 10% acrylamide gel. These two fragments were ligated to 2 μg of *Eco*RI-digested λ*plac5* in 30 μl of ligase buffer with 20 units of ligase. This mixture was used to transform *E. coli* 294. The configuration of restriction site ends (V represents *Hin*dIII; ● represents *Eco*RI) was such that only correct joining of the two halves of the B gene would lead to viable plasmids. To screen for the presence of the *lac* fragment, we plated the transformed culture on glucose minimal plates (10) containing 40 μg of 5-bromo-4-chloro-3-indoyl-β-galactoside (X-gal) and 20 μg of ampicillin per ml. Plasmids were prepared from β-galactosidase constitutive (blue) colonies. Because the λ*plac5* *lac* operon fragment contains an asymmetrical *Hin*dIII site (14), the orientation of that fragment in the resulting plasmids can be determined. Plasmid samples of 1 μg were digested with *Hin*dIII and sized on 0.7% agarose gels. Plasmids (15-μg samples) having the desired orientation of the *lac* fragment were then treated with *Eco*RI, *Hin*dIII, and *Bam*HI, and sized on a 10% acrylamide gel to verify the presence of both the BH and BB fragments. The diagram of pIB1 (7.1 megadaltons) is not drawn to scale. To construct the *lac*–insulin A plasmid (pIA1, not shown), we ligated 1 μg of *Eco*RI-treated pA11 (pBR322 containing the A gene) and 3 μg of *Eco*RI-treated λ*plac5* for 4 hr at 4°C. Transformants of *E. coli* 294 were selected for resistance to ampicillin on X-gal plates. Orientation of the *lac* fragment was determined by digesting plasmids purified from the blue colonies with *Hin*dIII and *Bam*HI.

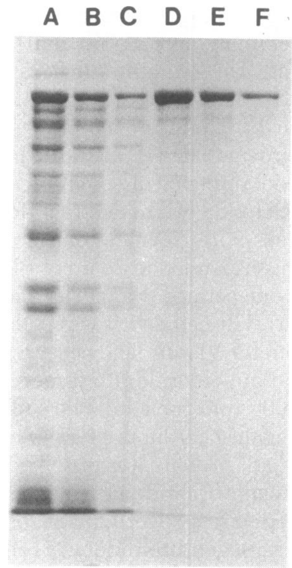

FIG. 2. Sodium dodecyl sulfate/polyacrylamide gel electrophoresis of extracts of strain 294/pIA1. Samples were heated in O'Farrell sample buffer and electrophoresed in a sodium dodecyl sulfate/10% gel as described (15). Lanes A, B, and C: total cells; 20, 10, and 5 μl, respectively. Lanes D, E, and F: insoluble cell debris; 20, 10, and 5 μl, respectively.

0.01 M, the mixture was centrifuged and the supernatant was used for further purification (see *Results and Discussion*).

RESULTS AND DISCUSSION

Assembly and Cloning of B-Chain Gene and A-Chain Gene. The gene for the B chain of insulin was designed to have an *Eco*RI restriction site on the left end, a *Hin*dIII site in the middle, and a *Bam*HI site at the right end. This was done so that both halves, the left *Eco*RI–*Hin*dIII half (BH) and the right *Hin*dIII–*Bam*HI half (BB), could be separately cloned in the convenient cloning vehicle pBR322 (8) and, after their sequences had been verified, joined to give the complete B gene (Fig. 1). The BB half was assembled by ligation from 10 oligodeoxyribonucleotides, labeled B1–B10 in figure 1 of ref. 4, made by phosphotriester chemical synthesis. B1 and B10 were not phosphorylated, thereby eliminating unwanted polymerization of these fragments through their cohesive ends (*Hin*dIII and *Bam*HI). After purification by preparative acrylamide gel electrophoresis and elution of the largest DNA band, the BB fragment was inserted into plasmid pBR322 that had been cleaved with *Hin*dIII and *Bam*HI. About 50% of the ampicillin-resistant colonies derived from the DNA were sensitive to tetracyline, indicating that a nonplasmid *Hin*dIII–*Bam*HI fragment had been inserted. The sequences of the small *Hin*dIII–*Bam*HI fragments from four of these colonies (pBB101 to pBB104) were determined (11) and were correct as designed.

The BH fragment was prepared in a similar manner and inserted into pBR322 that had been cleaved with *Eco*RI and *Hin*dIII restriction endonucleases. Plasmids from three ampicillin-resistant, tetracycline-sensitive transformants (pBH1 to pBH3) were analyzed. The small *Eco*RI–*Hin*dIII fragments had the expected nucleotide sequence.

The A-chain gene was assembled in three parts. The left four, middle four, and right four oligonucleotides (see figure 1 of ref. 4) were ligated separately, then mixed and ligated (oligonucleotides A1 and A12 were unphosphorylated). The assembled A-chain gene was phosphorylated, purified by gel electrophoresis, and cloned in pBR322 at the *Eco*RI–*Bam*HI sites. The

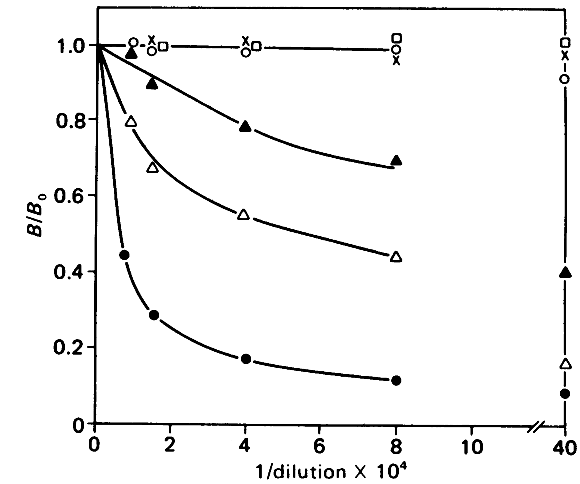

FIG. 3. Reconstitution radioimmunoassay for insulin chains. The *S*-sulfonated A sample was mixed with the *S*-sulfonated B sample in a 1.5-ml conical polypropylene tube and dried under reduced pressure. The dried proteins were suspended in 25 μl of 10 mM sodium acetate (pH 4.5). Two microliters of 10% (vol/vol) 2-mercaptoethanol was added and the mixture was heated (≈95°C) for 10 min. The mercaptoethanol was removed by ethyl acetate extraction. Two microliters of 0.1 M glycine buffer (pH 10.6) was added, and the pH was adjusted, if necessary, to 9.6–10.6. The open tube was placed in a dessicator over moist hyamine hydroxide at room temperature for ≥6 hr. A diluted aliquot of the reaction mixture was assayed for insulin radioimmune activity by use of a commercially available radioimmunoassay kit (Phadebus Insulin Test, Pharmacia). Dilution and assay were done according to the instructions supplied. The ordinate B/B_0 is the cpm in the pellet of the experimental sample divided by the cpm in the pellet obtained with buffer only. ●, 40 μg of porcine A(SSO_3^-) in 10 mM NH_4HCO_3 (pH 9) was mixed with 10 μg of bovine B(SSO_3^-) in the same buffer; ○, porcine A(SSO_3^-) only; □, bovine B(SSO_3^-) only; ▲, 100 μg of porcine A(SSO_3^-) and 93 μg of *E. coli* B-chain fraction F-10 (cleaved by CNBr and *S*-sulfonated, insoluble at pH 5); ×, fraction F-10 only; △, 100 μg of porcine A(SSO_3^-), 93 μg of fraction F-10, and 3 μg of bovine B(SSO_3^-).

*Eco*RI–*Bam*HI fragments from two ampicillin-resistant, tetracycline-sensitive clones (pA10 and pA11) contained the desired A-gene sequence.

Construction of Plasmids for Expression of A and B Insulin Genes. Fig. 1 illustrates the construction of the *lac*-insulin B plasmid (pIB1). Plasmids pBH1 and pBB101 were digested with *Eco*RI and *Hin*dIII endonucleases. The small BH fragment of pBH1 and the large fragment of pBB101 (containing the BB fragment and most of pBR322) were purified by gel electrophoresis, mixed, and ligated in the presence of *Eco*RI-cleaved λ*plac5*. The 4.4-megadalton *Eco*RI fragment of λ*plac5* contains the *lac* control region and the majority of the β-galactosidase structural gene (1, 14). The configuration of the restriction sites ensures correct joining of BH to BB. The *lac* *Eco*RI fragment can be inserted in two orientations; thus, only half of the clones obtained after transformation should have the desired orientation. The orientation of 10 ampicillin-resistant, β-galactosidase-constitutive clones were checked by restriction analysis (see legend of Fig. 1). Five of these colonies contained the entire B-gene sequence and the correct reading frame from the β-galactosidase gene into the B-chain gene. One, pIB1, was chosen for subsequent experiments.

In a similar experiment, the 4.4-megadalton *lac* fragment from λ*plac5* was introduced into the pA11 plasmid at the *Eco*RI site to give pIA1. pIA1 is identical to pIB1 except that the A-gene fragment is substituted for the B-gene fragment. DNA sequence analysis showed that the correct A- and B-chain gene sequences were retained in pIA1 and pIB1, respectively.

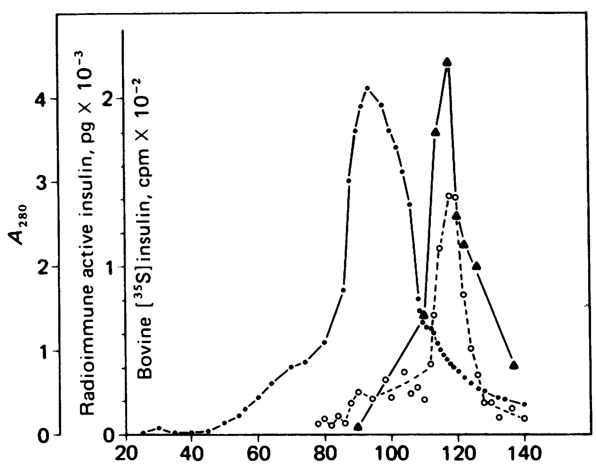

FIG. 4. DEAE-cellulose chromatography of an extract of strain D1210/pIB1. The fraction that was insoluble at pH 5 (580 mg of protein) was dissolved in 20 ml of 10 mM NH_4HCO_3 and adjusted to pH 9. ^{35}S-Labeled bovine B(SSO_3^-) (135,000 cpm; 4.4 μg) was added and the sample was applied to a 2 × 60 cm column of Whatman DE52. Elution was with a 1-liter gradient of 0.01–2.0 M NH_4HCO_3 (pH 9). Fractions of 4 ml were collected. ●, A_{280}; ○, 100-μl aliquots were used to measure radioactivity of ^{35}S-labeled bovine B(SSO_3^-); ▲, 100-μl aliquots were assayed for B-chain radioimmune activity by being mixed with 100 μg of porcine A(SSO_3^-) and using the reconstitution assay (Fig. 3).

Expression. The strains that contain the insulin genes correctly attached to β-galactosidase (D1210/pIB1 and 294/pIA1) both produce large quantities of a protein the size of β-galactosidase (Fig. 2). Approximately 20% of the total cellular protein was this β-galactosidase–insulin A or B chain hybrid. The hybrid proteins are insoluble and were found in the first low-speed pellet in which they constitute ≈50% of the protein (Fig. 2).

To detect the expression of the insulin A and B chains, we used a radioimmunoassay based on the reconstitution of complete insulin from the separate chains. The insulin reconstitution procedure of Katsoyannis et al. (3), adapted to a 27-μl assay volume, provided a very suitable assay. Easily detectable insulin radioimmune activity was obtained after S-sulfonated derivatives of the insulin chains were mixed and reconstituted by the procedure described in the legend of Fig. 3. The separate S-sulfonated chains of insulin do not react significantly, after reduction and oxidation, with the anti-insulin antibody used. Our reconstitution assay, though not extremely sensitive (limits of detection about 1 μg), was specific and suitable for following insulin chain radioimmune activity during purification.

To use the reconstitution assay, we partially purified the β-galactosidase–A or B chain hybrid protein, cleaved it with cyanogen bromide, formed S-sulfonated derivatives, and partially purified the peptides as described in *Materials and Methods*. This procedure was based on our earlier experience with the purification of somatostatin from E. coli (unpublished data) and the known properties of the insulin chains.

Table 1. Amino acid composition of E. coli insulin A chain

Amino acid	Residues per peptide: E. coli A(SSO_3^-)	Porcine A(SSO_3^-)	Predicted
His	0.08	0.08	0
Lys	0.00	0.00	0
Trp	0.00	0.00	0
Arg	0.00	0.00	0
Phe	0.00	0.00	0
Asx	2.38	2.50	2
Thr	0.24	0.28	1
Ser	0.14	0.23	2
H-Ser	0.02	0.00	0
Glx	3.97	4.58	4
Pro	0.00	0.09	0
Gly	1.40	1.48	1
Ala	0.20	0.11	0
Cys	0.55	0.00	0
Val	1.15	1.06	1
Met	0.62	0.43	0
Ile	1.99	1.48	2
Leu	2.33	2.35	2
Tyr	1.89	2.30	2

Approximately 25 μg of porcine A(SSO_3^-) (which is identical in sequence to human A) and 25 μg of E. coli A(SSO_3^-) purified twice by high-performance liquid chromatography were hydrolyzed and analyzed in parallel. The SSO_3^- derivatives of cysteine were destroyed during hydrolysis and do not register as amino acids with the program used. Serine and threonine were also partially destroyed.

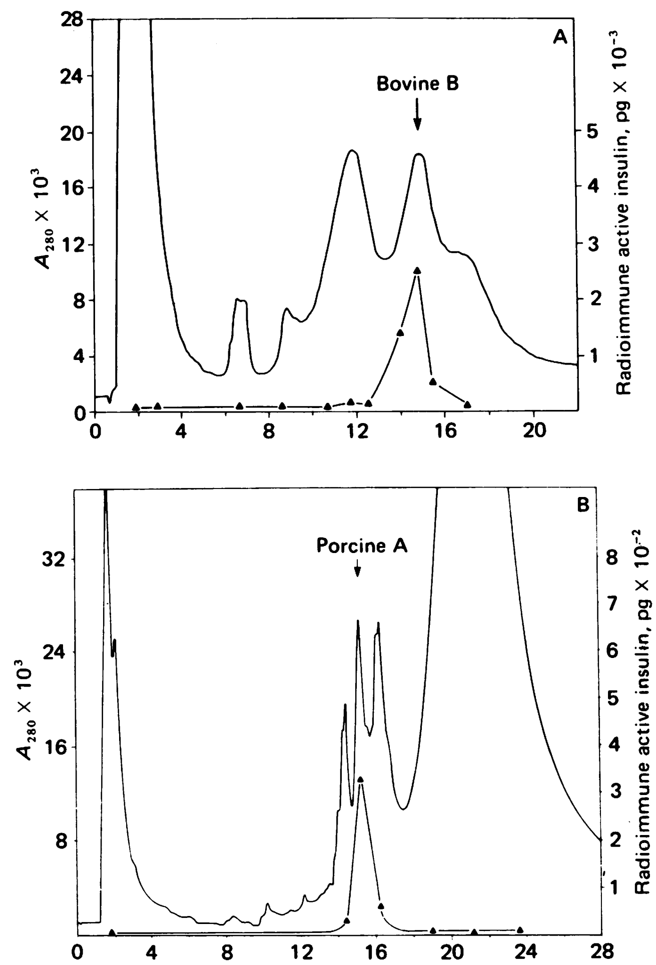

FIG. 5. Reversed-phase high-performance liquid chromatography. (A) A sample (500 μg), purified by DEAE-cellulose chromatography (B chain, Fig. 4), was subjected to high-performance liquid chromatography at room temperature on an SP 3500 liquid chromatograph (Spectra-Physics) equipped with a LiChrosorb RP-8 0.3 × 25 cm column (Merck EM). The eluting buffer was 50 mM NH_4OAc with an acetonitrile gradient of 24–60%. Fractions were collected and dried, and samples were assayed for B-chain radioimmune activity by our reconstitution assay. (B) An A-chain sample (500 μg total protein), partially purified by aminoethyl-cellulose chromatography, was subjected to high-performance liquid chromatography with an acetonitrile gradient of 15–60%. Fractions were collected and samples were assayed for A-chain radioimmune activity. Solid line, A_{280}; ▲, radioimmune activity.

Table 2. Reconstitution of radioimmune human insulin

"A" sample	"B" sample	Radioimmune active insulin, ng
E. coli 58-HPLC*	—	<0.5
—	*E. coli* DE117†	<0.5
Porcine A‡	*E. coli* DE117	74
E. coli 58-HPLC	Bovine B§	45
E. coli 58-HPLC	*E. coli* DE117	20

Our standard reconstitution assay procedure was used (Fig. 3). The results are given as ng of radioimmune active insulin produced per 20 μl of the reaction mixture. HPLC, high-performance liquid chromatography.

* Five hundred microliters of fraction 58 from an aminoethyl-cellulose column was chromatographed on an RP-8 column and the "A" peak was collected. As estimated from the peak height, the sample contained approximately 25 μg of protein.

† Ten microliters of DEAE-cellulose fraction 117 (Fig. 4), concentrated to 1.6 mg of total protein per ml, was used as the "B" sample.

‡ *S*-Sulfonated porcine A (70 μg).

§ *S*-Sulfonated bovine B (10 μg).

Insulin B-chain radioimmune activity was detected first among the *S*-sulfonated cyanogen bromide peptides insoluble at pH 5 (fraction F-10, Fig. 3). The activity was enriched further by chromatography on DEAE-cellulose (Fig. 4). The B-chain radioimmune activity coeluted with *S*-[^{35}S]sulfonated bovine B chain.

A portion of the material purified by DEAE-cellulose chromatography was analyzed by high-performance liquid chromatography on a reversed-phase RP-8 column (Fig. 5*A*). This column separates primarily on the basis of hydrophobic interactions. The insulin B-chain radioimmune activity eluted at a position very close to that of bovine B chain. Good purification was obtained by high-performance liquid chromatography, but the breadth of the peak indicated that the chromatographic fraction was not pure.

Another sample (1 mg total protein) of the material purified by DEAE-cellulose chromatography was subjected to gel filtration on Sephadex G-75 in 50% acetic acid, a system that completely resolves A chain from B chain. The B-chain radioimmune activity eluted at the same position as *S*-sulfonated bovine B chain, indicating similar sizes (data not shown).

Insulin A-chain radioimmune activity was detected first in the total mixture of cyanogen bromide peptide fragments obtained from the partially purified β-galactosidase–A chain hybrid. The activity was enriched by pH 5 precipitation and aminoethyl-cellulose chromatography and purified on a microgram scale by high-performance liquid chromatography (Fig. 5*B*). The insulin A-chain radioimmune activity eluted from the column at a position indistinguishable from that of porcine *S*-sulfonated A chain. Porcine A chain is identical to human A chain (2).

When an excess of porcine A(SSO_3^-) (40 μg) was mixed, reduced, and oxidized with bovine B(SSO_3^-) (10 μg), we usually obtained 10–15% correct joining to yield radioimmune active insulin. Reconstitution in impure mixtures was lower, as expected. Because of this strong and variable competitive inhibition by other peptides, the amount of insulin chains in the extracts can best be estimated by adding to the extract a known amount of the chain to be assayed. This type of experiment (illustrated in Fig. 3) indicates that the yield of insulin chains is high (approximately 10 mg from 24 g wet weight of cells) and consistent with the amount of insoluble β-galactosidase protein obtained (at least 10^5 molecules per cell). This estimated yield is 10 times higher than that reported for somatostatin (1).

The evidence that we have obtained correct expression from chemically synthesized genes for human insulin can be summarized as follows. (*i*) Radioimmune activity has been detected for both chains. (*ii*) The DNA sequences obtained after cloning and plasmid construction have been directly verified to be correct as designed. Because radioimmune activity is obtained, translation must be in phase. Therefore, the genetic code dictates that peptides with the sequences of human insulin are being produced. (*iii*) The *E. coli* products, after cyanogen bromide cleavage, behave as insulin chains in three different chromatographic systems that separate on different principles (gel filtration, ion exchange, and reversed-phase high-performance liquid chromatography). (*iv*) The A chain produced by *E. coli* has been purified on a small scale by high-performance liquid chromatography and has the correct amino acid composition (Table 1).

Table 2 illustrates that insulin radioimmune activity can be obtained entirely from *E. coli* products. Easily detectable radioimmune insulin activity is produced when purified *E. coli* A chain is mixed and reconstituted with partially purified ($\approx$10% pure) *E. coli* B chain.

We gratefully acknowledge the excellent technical assistance of Louise Shively, Rochelle Sailor, Frances Fields, and Mark Backer. We give special thanks to Herbert W. Boyer for his encouragement and scientific consultation and to Robert A. Swanson for making this work possible. Work at the City of Hope was supported by contracts from Genentech, Inc.

1. Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F. & Boyer, H. W. (1977) *Science* **198,** 1056–1063.
2. Humbel, R. E., Bosshard, H. R. & Zahn, H. (1972) in *Handbook of Physiology,* eds. Steiner, D. F. & Freinkel, N. (Williams & Wilkins, Baltimore, MD), Vol. 1, pp. 111–132.
3. Katsoyannis, P. G., Trakatellis, A. C., Johnson, S., Zalut, C. & Schwartz, G. (1967) *Biochemistry* **6,** 2642–2655.
4. Crea, R., Hirose, T., Kraszewski, A. & Itakura, K. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 5765–5769.
5. Backman, K., Ptashne, M. & Gilbert, W. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 4174–4178.
6. Panet, A., van de Sande, J. H., Loewen, P. C., Khorana, H. G., Raae, A. J., Lillenhaug, J. R. & Kleppe, K. (1973) *Biochemistry* **12,** 5045–5049.
7. Greene, P. J., Betlach, M., Bolivar, F., Heyneker, H. L., Tait, R. & Boyer, H. W. (1978) *Nucleic Acids Res.* **12,** 2373–2380.
8. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2,** 95–119.
9. Clewell, D. B. (1972) *J. Bacteriol.* **110,** 667–676.
10. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
11. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
12. Hershfield, V., Boyer, H. W., Yanofsky, C., Lovett, M. A. & Helinski, D. R. (1974) *Proc. Natl. Acad. Sci. USA* **71,** 3455–3459.
13. Katsoyannis, P. G., Tometsko, A., Zalut, C., Johnson, S. & Trakatellis, A. C. (1967) *Biochemistry* **6,** 2635–2642.
14. Polisky, B., Bishop, R. J. & Gelfand, D. H. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 3900–3904.
15. O'Farrell, P. H. (1975) *J. Biol. Chem.* **250,** 4007–4021.