(12) R. E. Dickerson and I. Geis, "The Structure and Action of Proteins", Harper and Row, New York, N.Y., 1969.
(13) This possibility arises because the backbone carbon atoms of the first and third strands of β structure actually come quite close together (i.e., 8–10 Å), even though the side-chain atoms of these strands are not in contact.
(14) J. C. Wang, *J. Mol. Biol.*, **43**, 263 (1969).
(15) H. C. Watson, *Prog. Stereochem.*, **4**, 299 (1968).
(16) F. S. Mathews, P. Argos, and M. Levine, *Cold Spring Harbor Symp. Quant. Biol.*, **38**, 387 (1972).
(17) R. Huber, D. Kukla, A. Ruhlmann, and W. Steiglmann, *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 141 (1972).

[Sar^1-A]Insulin, a Biologically Active Analog[1,2]

Yoshio Okada and Panayotis G. Katsoyannis*

Contribution from the Department of Biochemistry, Mount Sinai School of Medicine of the City University of New York, New York, New York 10029. *Received February 19, 1975*

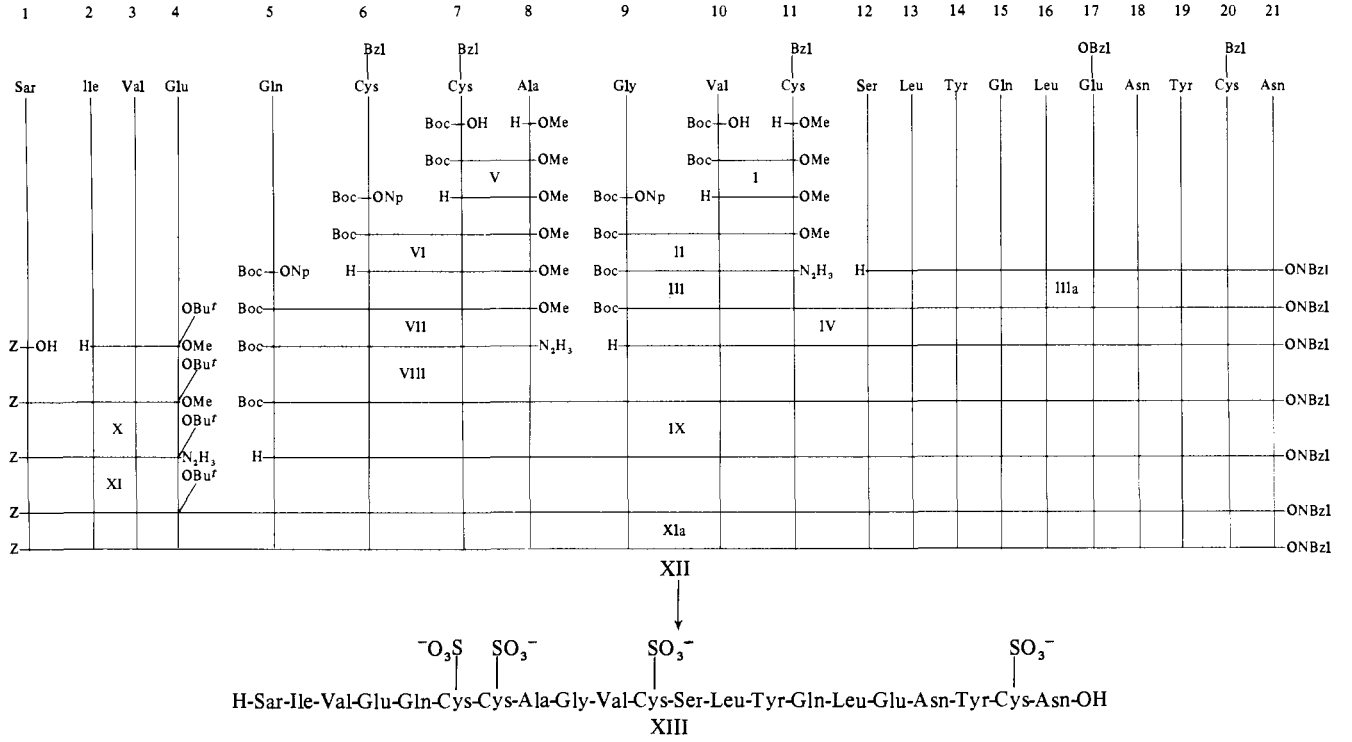
Abstract: The synthesis and isolation in purified form of [Sar^1-A]insulin, a biologically active analog of sheep insulin, are described. This analog differs from the parent molecule in that the N-terminal amino acid residue of the A chain, glycine, has been replaced with sarcosine. This change which in essence is the replacement of one hydrogen of the amino group of the A^1 residue with a methyl group, with a concomitant increase of the basic character of that residue, results in a small decrease of the biological activity and a more pronounced decrease of the immunoreactivity of insulin (20 and 9 IU/mg, respectively, vs. 23–25 IU/mg for the natural hormone). For the synthesis of this analog, the [Sar^1]A chain was synthesized by the fragment condensation method and isolated as the S-sulfonated derivative. Conversion of the latter compound to the sulfhydryl form and combination of the reduced product with the S-sulfonated derivative of the B chain of sheep (bovine) insulin afforded the [Sar^1-A]insulin. Isolation and purification of the insulin analog was accomplished by chromatography on a carboxymethylcellulose column with an exponential sodium chloride gradient.

The importance of the amino acid residues located at the amino terminal regions of the A and B chains of insulin to the biological activity has become evident by studies carried out in several laboratories.[3] Our own investigations have shown that removal of the N-terminal tetrapeptide sequence from the A chain results in complete inactivation of the molecule.[4] Of particular interest was our finding that replacement of the α-amino group of the N-terminal glycine by hydrogen results in a substantial decrease of the biological activity of the hormone[5] (7–10 IU/mg vs. 23–25 IU/mg for the natural compound). Similarly, removal of the N terminus of the A chain, glycine, leads to a 90% loss of the biological activity of insulin.[6] On the other hand, elimination of the N-terminal amino acid residue of the B chain does not cause any appreciable change in the biological activity of the hormone.[7] These data therefore clearly demonstrate that the A^1 residue is critically involved in the maintenance of high biological activity of insulin. It appears, however, that the amino group at A^1 is not per se involved with the manifestation of high biological activity of insulin. Indeed, modification of the amino group of glycine A^1 with a variety of acyl groups does not strikingly affect the biological potency and immunoreactivity of the hormone.[3,8,9] The preparation of numerous analogs of insulin with the A^1 amino group modified indicates that the relative size of the modifying acyl group affects the biological profile of the hormone and that large acyl groups decrease the biological activity of insulin.[9] These results led to the speculation[9] that the N-terminal residue of the A chain is involved in the stabilization of the "active site" of insulin and that this site is affected either by conformational changes and/or a steric effect of the large acyl groups at A^1.

Characteristic of almost all insulin analogs involving modifications at A^1 that have been prepared thus far is that they no longer possess a basic group at A^1. It was therefore of interest to investigate the effect on the biological activity and immunogenicity of insulin of modifications at A^1 that increase the basic character of that residue. This investigation describes the synthesis, isolation in pure form, and biological evaluation of such an analog, [Sar^1-A] sheep insulin, which differs from the parent molecule in that the N terminus of the A chain, glycine (pK_2 = 9.6), has been replaced with sarcosine (pK_2 = 10.00). By the mouse convulsion assay method this analog was found to possess a potency of 20 IU/mg. By the radioimmunoassay method [Sar^1-A]insulin possesses a potency of 9 IU/mg (vs. 23–25 IU/mg for the natural hormone). It thus becomes apparent that replacement of one hydrogen of the amino group of the A^1 residue with a methyl group with a concomitant increase of the basic character of that residue results in but a modest decrease of the biological activity and a more pronounced decrease of the immunogenicity of insulin. It is interesting to note that a parallel situation exists with another insulin analog, the [Arg-Gly^{A1}]insulin, where one hydrogen of the amino group at the A^1 residue is replaced with the basic arginine residue.[10] In this case, too, there is a more pronounced decrease of the immunogenicity than of the biological activity of the molecule (60 vs. 40%). The synthesis of several insulin analogs modified at the A^1 residue is, of course, required before a firm conclusion regarding such a relationship between chemical structure, biological activity, and immunogenicity is established.

General Aspects of the Synthesis. [Sar^1-A] sheep insulin was prepared by combination of the S-sulfonated form of the B chain of bovine (sheep) insulin with the sulfhydryl form of the [Sar^1]A chain of sheep insulin. The S-sulfonated bovine B chain, which is identical with the respective chain of sheep insulin,[11,12] was prepared by oxidative sulfitolysis of bovine insulin, followed by separation of the resulting S-sulfonated derivatives of the A and B chains by continuous flow electrophoresis. The sulfitolysis of insulin and the separation of the resulting A and B chain derivatives have been described in previous communications from this laboratory.[13] The synthesis of the [Sar^1]A chain with its functional groups protected (XII) was accomplished by the classical methods of peptide synthesis, namely, a combination of the "stepwise elongation" and "fragment conden-

Chart I



sation" approaches, and was patterned after that of the natural sheep and human A chains.[14,15] The overall scheme which was used for the construction of the protected chain XII is summarized in Chart I.

The C-terminal decapeptide fragment[16] IIIa (sequence 12–21) was allowed to react with the tripeptide derivative III (sequence 9–11) by the azide method to give the protected dodecapeptide fragment IV (sequence 9–21). Deblocking of the latter compound and coupling of the resulting derivative with the adjacent tetrapeptide fragment VIII (sequence 5–8) by the azide method afforded the protected heptadecapeptide derivative IX (sequence 5–21). Selective deblocking at the amino end of this derivative and azide coupling of the resulting product with the N-terminal tetrapeptide fragment XI (sequence 1–4) gave the fully protected heneicosapeptide XIa which, on exposure to trifluoroacetic acid, was converted to the desired product XII. Removal of the blocking groups, namely, benzyloxycarbonyl, *S*-benzyl, γ-benzyl, and *p*-nitrobenzyl ester, from the protected chain derivative XII was accomplished by the sodium–liquid ammonia method.[17] The deblocked heneicosapeptide was converted to the S-sulfonated derivative XIII by oxidative sulfitolysis.[13,18] The S-sulfonated [Sar^1]A chain was purified by chromatography on Sephadex G-15.

Amino acid analysis of the purified S-sulfonated chain after acid hydrolysis gave a composition expressed in molar ratios in agreement with the theoretically expected values (Table I). Digestion of this material with leucine aminopeptidase and amino acid analysis of the digest showed that the constituent amino acids were present in the theoretically expected ratios and thus established the stereochemical homogeneity of the synthetic chain within the limits of error of the enzymatic technique (Table I). On thin-layer chromatography in two solvent systems (Figures 1 and 2) and on high-voltage thin-layer electrophoresis at two pH values (Figures 3 and 4) the synthetic product exhibited a single Pauly-positive spot.

Isolation of the Insulin Analog. The conversion of the [Sar^1]A chain S-sulfonate to its sulfhydryl form and its combination with the S-sulfonated B chain to produce the [Sar^1-A] sheep insulin was carried out by the procedures developed in this laboratory.[19-21] Isolation of the insulin analog from the combination mixture in purified form was accomplished by chromatography on a CM-cellulose column using an acetate buffer (pH 3.3) and an exponential NaCl gradient, as we have described previously.[19,21] The chromatographic profile obtained is shown in Figure 5. The insulin analog was isolated from the effluent via picrate as the hydrochloride by the procedure described previously.[19,21]

Amino acid analysis of the purified [Sar^1-A]insulin after acid hydrolysis gave a composition expressed in molar ratios in agreement with the theoretically expected values (Table II). On thin-layer electrophoresis the synthetic analog moved as a single component (Figure 6).

Table I. Amino Acid Composition[a] of an Acid Hydrolysate and a Leucine Aminopeptidase Digest of the S-Sulfonated [Sar^1] A Chain of Sheep Insulin

	Acid hydrolysis		Enzymatic Hydrolysis	
Amino acid	Theory	Found	Theory	Found
Aspartic acid	2.0	1.9	0	0
Sarcosine	1.0	1.1	1.0	Emerge on the same position and not determined
Glutamine	0	0	2.0	
Asparagine	0	0	2.0	
Serine	1.0	0.8[b]	1.0	0.9[c]
Glutamic acid	4.0	4.2	2.0	2.1
Glycine	1.0	1.1	1.0	1.0
Alanine	1.0	1.1	1.0	1.0
Half-cystine	4.0	2.9[b]	0	0
Valine	2.0	1.7	2.0	2.0
Isoleucine	1.0	0.7	1.0	1.1
Leucine	2.0	1.9	2.0	2.0
Tyrosine	2.0	1.6[b]	2.0	1.9
S-Sulfocysteine	0	0	4.0	4.2

[a] Number of amino acid residues per molecule. [b] Uncorrected for destruction. [c] Separated from glutamine, asparagine, and sarcosine in a 30° chromatographic run.

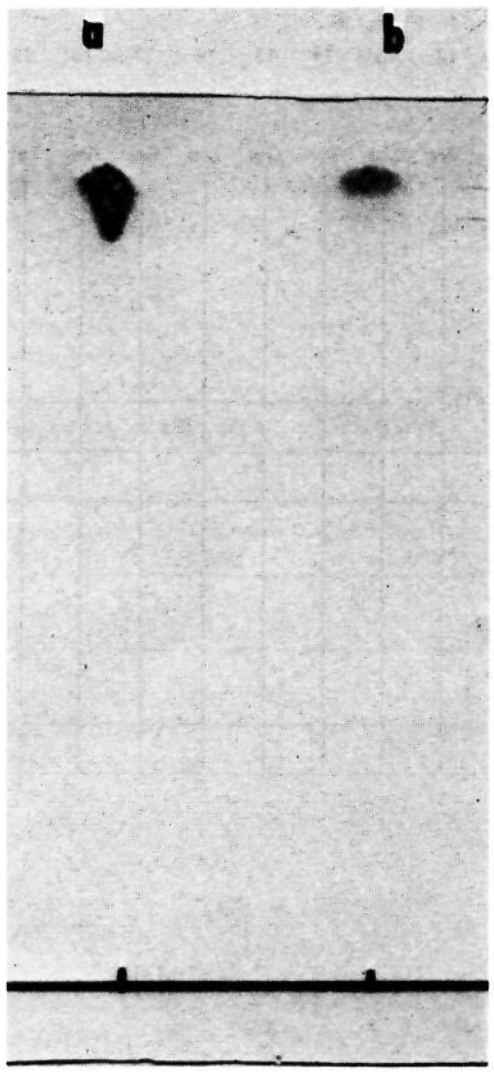


Figure 1. Thin-layer chromatography of natural bovine A chain *S*-sulfonate (a) and synthetic [Sar^1]A chain *S*-sulfonate (b) in the system 1-butanol–pyridine–acetic acid–water (30:20:6:24).

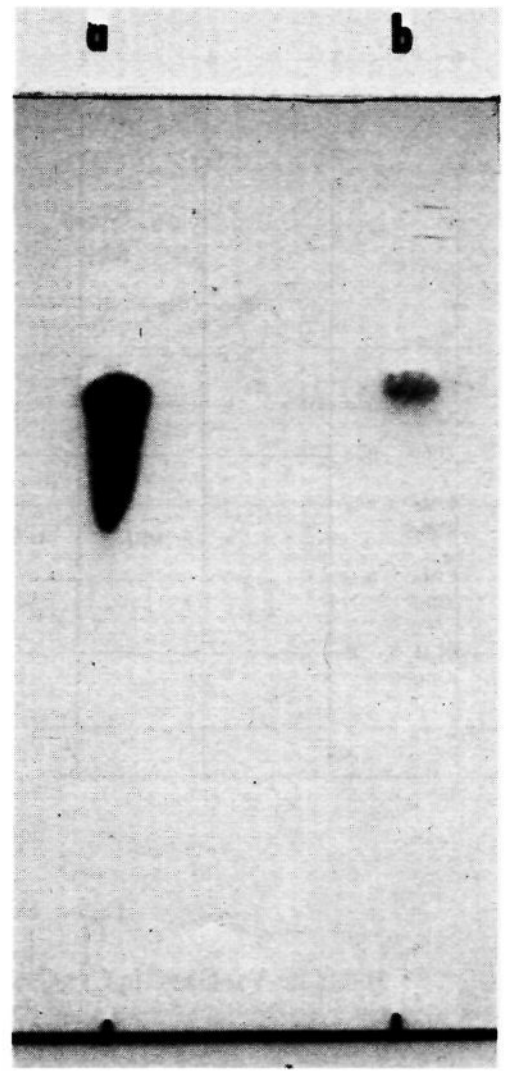


Figure 2. Thin-layer chromatography of the *S*-sulfonates of natural bovine A chain (a) and synthetic [Sar^1]A chain (b) in the system 1-butanol–pyridine–acetic acid–water (4:1:1:2).

Table II. Amino Acid Composition[a] of an Acid Hydrolysate of the [Sar^1-A] Sheep Insulin

Amino acid	Theory	Found
Lysine	1	1.0
Histidine	2	1.8
Arginine	1	1.0
Aspartic acid	3	2.7
Threonine	1	1.0
Serine	2	2.1
Sarcosine	1	1.2
Proline	1	1.0
Glutamic acid	7	6.9
Glycine	4	4.3
Alanine	3	3.2
Half-cystine	6	4.0[b]
Valine	5	4.7
Isoleucine	1	0.8
Leucine	6	5.9
Tyrosine	4	3.3[b]
Phenylalanine	3	3.0

[a] Number of amino acid residues per molecule. [b] Uncorrected for destruction.

Experimental Section

Capillary melting points were determined for all peptide derivatives and are not corrected. Optical rotations were taken with a Zeiss photoelectric precision polarimeter. Thin-layer chromatography was performed on 6060 silica gel (Eastman Chromagram Sheet, Eastman Kodak Co., Rochester, N.Y.). The following solvent systems were used: (A) 1-butanol–pyridine–acetic acid–water (30:20:6:24) and (B) 1-butanol–pyridine–acetic acid–water (4:1:1:2). The homogeneity of all intermediate peptide derivatives, deblocked at the amino end, was ascertained by thin-layer chromatography in both the above-mentioned solvent systems. In all synthetic steps, coupling of the fragments was followed by detection of the amino component present with ninhydrin; completion of the reaction was indicated by a negative ninhydrin test. Thin-layer electrophoresis was performed by a method developed in this laboratory[22] and was carried out with a Wieland-Pfleiderer pherograph (Brinkmann Instruments, Westbury, N.Y.). Amino acid analyses were performed in a Beckman-Spinco amino acid analyzer (Model 120C), equipped with a digital readout system, according to the method of Spackman et al.[23a] Acid hydrolysis and calculations of molar ratios were carried out as described previously.[13] For the enzymatic digestion with leucine aminopeptidase, the method of Hill and Smith[23b] was employed with a chromatographically purified enzyme from Worthington Biochemical Corp., Freehold, N.J. Preswollen microgranular CM-cellulose (Whatman CM 52/1) and Sephadex G-15 (Pharmacia Uppsala) were used. The washing of the resins and the preparation of the columns and of the buffers used were described previously.[13,19] The column eluates were continuously monitored with a Gilford recording spectrophotometer at 278 nm. Sodium tetrathionate was prepared as described by Gilman et al.[24] *tert*-Butoxycarbonylamino acids were purchased from Cyclo Chemical Corp. Protein determinations were carried out by the method of Lowry et al.[25] Biological assays were carried out by the mouse convulsion method as has been described previously.[20,21] For radioimmunoassays the method of Hales and Randle[26] was employed using an "insulin immunoassay kit" (Amersham/Searle Co.). Crystalline bovine insulin was generously provided by Eli Lilly and Co. The following abbreviations are used: Z, benzyloxycarbonyl; Bzl, benzyl; Boc, *tert*-butoxycarbonyl; Bu^t, *tert*-butyl; Me, methyl; Np, *p*-nitrophenyl; Me_2Fam, dimethylformamide; Me_2SO, dimethyl sulfoxide.

***N-tert*-Butoxycarbonyl-L-valyl-*S*-benzyl-L-cysteine Methyl Ester (I).** To a solution of *N-tert*-butoxycarbonyl-L-valine (10.4 g) in tetrahydrofuran (75 ml) containing triethylamine (7 ml) cooled to −20°, isobutyl chloroformate (6.7 ml) was added. The reaction mixture was stirred for 15 min (−20°) and added to a cold solution (0°) of *S*-benzyl-L-cysteine methyl ester (prepared from 13.1 g of the corresponding hydrochloride[27]) in Me_2Fam (50 ml). After 30 min at 0° and 3 hr at room temperature, the solvent was removed by evaporation in vacuo and the remaining product was mixed with ethyl acetate (500 ml) and water (150 ml). The organic layer was washed successively with 10% citric acid, saturated $NaHCO_3$, and water, dried, and concentrated to dryness in vacuo. Addition of

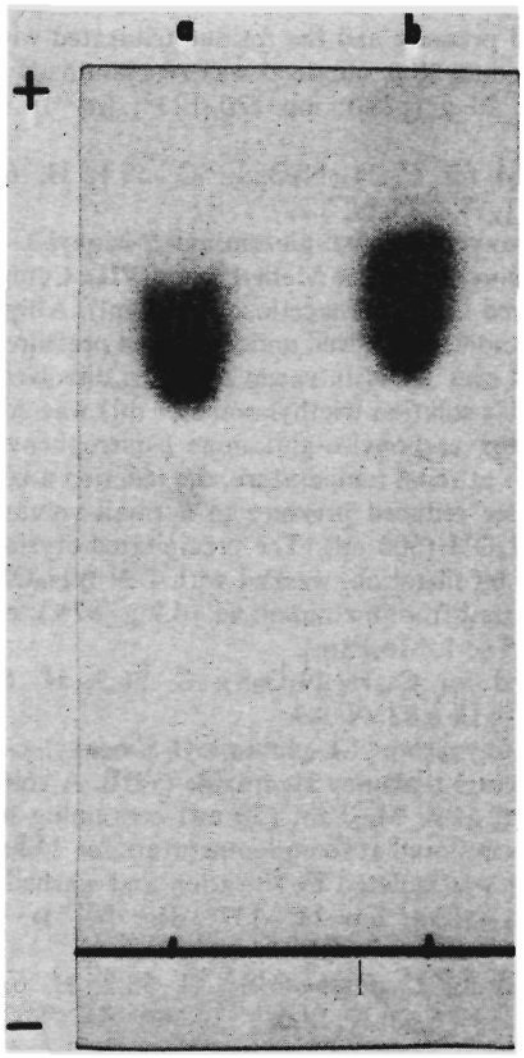


Figure 3. High-voltage thin-layer electrophoresis of natural bovine A chain *S*-sulfonate (a) and synthetic [Sar^1]A chain *S*-sulfonate (b): 0.5 *N* acetic acid, 3500 V, 20 min.

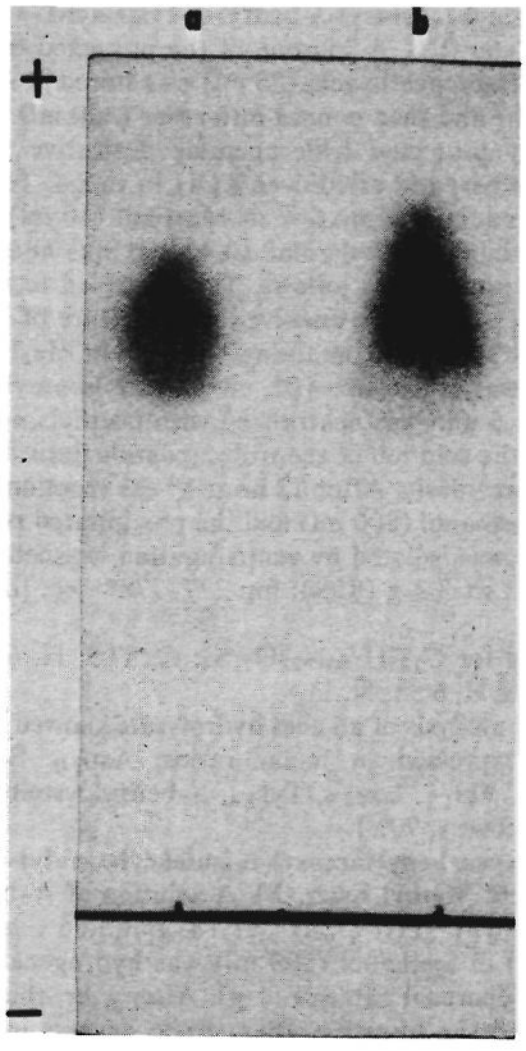


Figure 4. High-voltage thin-layer electrophoresis of the *S*-sulfonates of natural bovine A chain (a) and synthetic [Sar^1]A chain (b): 0.01 *M* NH_4HCO_3, adjusted to pH 10 with NH_4OH, 2900 V, 18 min.

ether and petroleum ether caused the crystallization of the product: wt 13.5 g (64%); mp 112–115°; $[\alpha]^{26}D$ −59.7° (*c* 1, methanol).

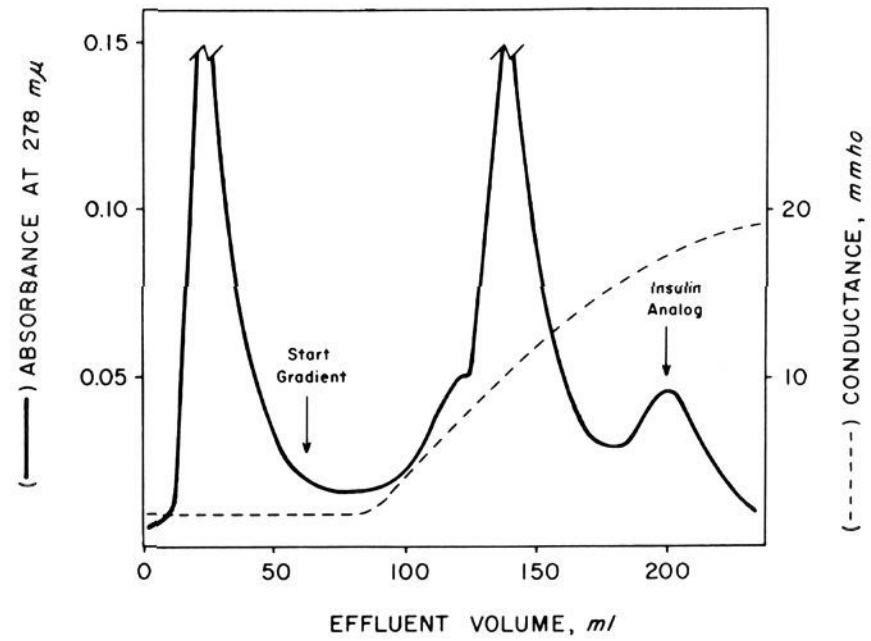


Figure 5. Chromatography of a combination mixture of the sulfhydryl form of sheep [Sar^1]A chain with the S-sulfonated sheep (bovine) B chain on a 0.9 × 23 cm CM-cellulose column with acetate buffer (pH 3.3, [Na^+] 0.024 *M*) and an exponential NaCl gradient. Two combination mixtures (see Experimental Section) were processed. The column effluent was monitored by a Gilford recording spectrophotometer and by a conductivity meter (Radiometer, Copenhagen). The [Sar^1-A] sheep insulin (175–250 ml of effluent) was recovered as the hydrochloride.

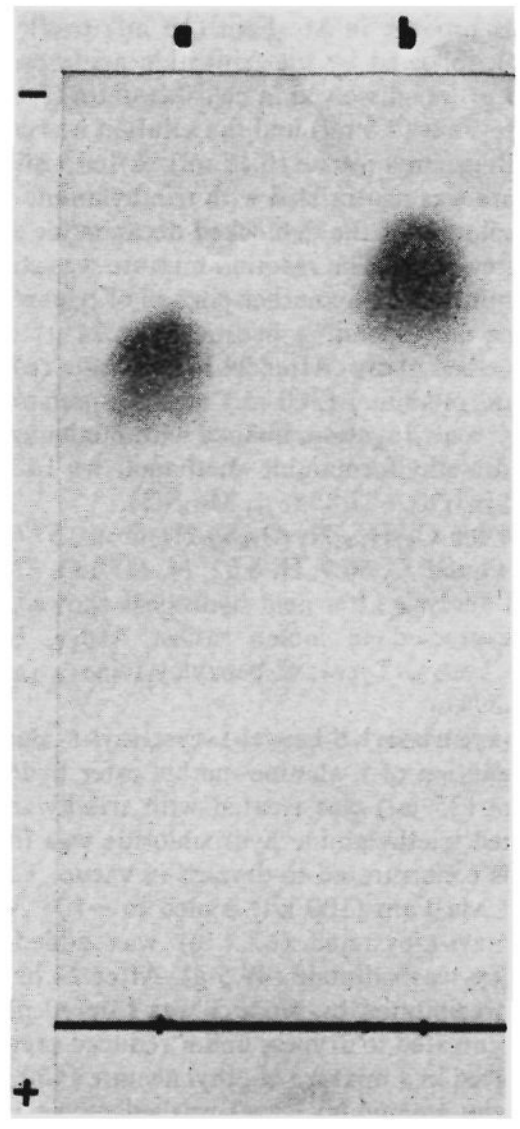


Figure 6. High-voltage thin-layer electrophoresis of natural bovine insulin (a) and synthetic [Sar^1-A] sheep insulin (b): 0.5 *N* acetic acid, 3500 V, 15 min.

Anal. Calcd for $C_{21}H_{32}N_2O_5S$: C, 59.4; H, 7.60; N, 6.6. Found: C, 59.6; H, 7.86; N, 6.8.

***N-tert*-Butoxycarbonylglycyl-L-valyl-*S*-benzyl-L-cysteine Methyl Ester (II).** A solution of compound I (13.4 g) in trifluoroacetic acid (25 ml) was stored at room temperature for 1.5 hr and then concentrated to dryness in vacuo. To a solution of the residue in Me_2Fam (50 ml) cooled to 0°, triethylamine (7 ml) was added followed by *N-tert*-butoxycarbonylglycine *p*-nitrophenyl ester[28]

(10.9 g). After 40 hr at room temperature the solvent was removed by evaporation in vacuo and the residue was treated with 1 *N* NH_4OH. The crystalline product was isolated by filtration, washed successively with 1 *M* NH_4OH and water, and recrystallized from methanol; wt 13.4 g (88%); mp 171–173°; $[\alpha]^{26}D$ −23.9° (*c* 1, Me_2Fam).

Anal. Calcd for $C_{23}H_{35}N_3O_6S$: C, 57.4; H, 7.33; N, 8.7. Found: C, 57.3; H, 7.25; N, 8.7.

***N-tert*-Butoxycarbonylglycyl-L-valyl-*S*-benzyl-L-cysteine Hydrazide (III).** Compound II (10 g) was dissolved in Me_2Fam (30 ml) and treated with hydrazine hydrate (5 ml) at room temperature for 24 hr. Upon removal of the solvent in vacuo and trituration of the residue with methanol, 8.9 g (89%) of product was obtained: mp 161–163°; $[\alpha]^{26}D$ −20.7° (*c* 1, Me_2Fam).

Anal. Calcd for $C_{22}H_{35}N_5O_5S$: C, 54.9; H, 7.33; N, 14.5. Found: C, 54.7; H, 7.36; N, 14.4.

Amino acid analysis of an acid hydrolysate showed the following composition expressed in molar ratios: $Gly_{1.0}$, $Val_{1.0}$, *S*-benzylcysteine$_{0.7}$ (average amino acid recovery, 94%).

***N-tert*-Butoxycarbonylglycyl-L-valyl-*S*-benzyl-L-cysteinyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-γ-benzyl-L-glutamyl-L-asparaginyl-L-tyrosyl-*S*-benzyl-L-cysteinyl-L-asparagine *p*-Nitrobenzyl Ester (IV).** A solution of *N-tert*-butoxycarbonyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-γ-benzyl-L-glutamyl-L-asparaginyl-L-tyrosyl-*S*-benzyl-L-cysteinyl-L-asparagine *p*-nitrobenzyl ester[16] (1.7 g) in trifluoroacetic acid (7 ml) was stored at room temperature for 1 hr. Addition of ether (400 ml) to the reaction mixture caused precipitation of the trifluoroacetate of the decapeptide derivative IIIa which was collected by filtration, washed with ether, and dried over KOH in vacuo. To a cooled (0°) solution of this product in Me_2Fam (10 ml) triethylamine (0.28 ml) was added, followed by the tripeptide azide prepared as follows. III (0.53 g) was dissolved in cold Me_2Fam (7 ml) containing 1 *N* HCl in Me_2Fam (2.5 ml) and the solution was cooled to −15° and mixed with isoamyl nitrite (0.15 ml). After 5 min at −15° the reaction mixture was neutralized with triethylamine (0.31 ml) and added to the solution of the deblocked decapeptide ester prepared as described previously. The reaction mixture was stirred at 4° for 24 hr and then mixed with another portion of tripeptide azide prepared from the corresponding hydrazide (0.24 g) in exactly the same way described above. After 24 hr at 4° the reaction mixture was poured into methanol (500 ml) and the precipitated product was isolated by centrifugation, washed with methanol, and reprecipitated from dimethylformamide–methanol: wt 1.6 g (85%); mp 270–276° dec; $[\alpha]^{26}D$ −27.8° (*c* 1, Me_2SO).

Anal. Calcd for $C_{97}H_{125}N_{17}O_{26}S_2 \cdot 2H_2O$: C, 57.0; H, 6.35; N, 11.6; O, 21.9. Found: C, 56.9; H, 6.23; N, 11.7; O, 22.3.

Amino acid analysis after acid hydrolysis showed the following composition expressed in molar ratios: $Asp_{2.0}$, $Ser_{0.8}$, $Glu_{2.0}$, $Gly_{0.9}$, $Val_{0.9}$, $Leu_{2.0}$, $Tyr_{1.6}$, *S*-benzylcysteine$_{1.6}$ (average amino acid recovery, 87%.)

***N-tert*-Butoxycarbonyl-*S*-benzyl-L-cysteinyl-L-alanine Methyl Ester (V).** A solution of L-alanine methyl ester hydrochloride (28 g) in methanol (35 ml) was treated with triethylamine (28 ml). The precipitated triethylamine hydrochloride was filtered off and the filtrate was concentrated to dryness in vacuo. To a solution of this residue in Me_2Fam (100 ml) cooled to −10° *N-tert*-butoxycarbonyl-*S*-benzyl-L-cysteine (62.2 g) was added followed by *N,N'*-dicyclohexylcarbodimide (49.5 g). After 24 hr at room temperature the precipitated by-product was filtered off and the filtrate was concentrated to dryness under reduced pressure. The residue was dissolved in a mixture of ethyl acetate (400 ml) and water (100 ml) and the organic layer was washed successively with 10% citric acid, saturated $NaHCO_3$, and water. Upon removal of the solvent under reduced pressure and trituration of the residue with petroleum ether, 60 g (76%) of crystalline product was obtained: mp 95–97°; $[\alpha]^{26}D$ −37.4° (*c* 1, methanol).

Anal. Calcd for $C_{19}H_{28}N_2O_5S$: C, 57.6; H, 7.12; N, 7.1. Found: C, 57.6; H, 7.12; N, 7.3.

***N-tert*-Butoxycarbonyl-*S*-benzyl-L-cysteinyl-*S*-benzyl-L-cysteinyl-L-alanine Methyl Ester (VI).** Compound V (15.9 g) was dissolved in trifluoroacetic acid (25 ml). After 1 hr the trifluoroacetic acid was removed under reduced pressure and the residue was dried over KOH in vacuo. To a solution of this product in Me_2Fam (50 ml) triethylamine (5.6 ml) was added followed by *N-tert*-butoxycarbonyl-*S*-benzyl-L-cysteine *p*-nitrophenyl ester[29] (18 g). After 24 hr at room temperature the solvent was removed under reduced pressure and the residue triturated with ether. The crystalline product thus obtained was recrystallized from methanol–ether: wt 20 g (83%); mp 120–123°; $[\alpha]^{26}D$ −53.5° (*c* 1, methanol).

Anal. Calcd for $C_{29}H_{39}N_3O_6S_2$: C, 59.1; H, 6.67; N, 7.1. Found: C, 59.1; H, 6.85; N, 7.2.

***N-tert*-Butoxycarbonyl-L-glutaminyl-*S*-benzyl-L-cysteinyl-*S*-benzyl-L-cysteinyl-L-alanine Methyl Ester (VII).** Compound VI (15 g) was dissolved in trifluoroacetic acid (20 ml). After 1 hr the trifluoroacetic acid was removed under reduced pressure and the residue was dried over KOH in vacuo and then dissolved in Me_2Fam (35 ml). To this solution triethylamine (7 ml) was added followed by *N-tert*-butoxycarbonyl-L-glutamine *p*-nitrophenyl ester[30] (9.6 g). After 24 hr at room temperature, the reaction mixture was concentrated under reduced pressure to a small volume and mixed with 1 *N* NH_4OH (500 ml). The precipitated crystalline product was collected by filtration, washed with 1 *N* NH_4OH and water, and recrystallized from methanol: wt 14.9 g (82%); mp 192–195°; $[\alpha]^{26}D$ −51.6° (*c* 1, Me_2Fam).

Anal. Calcd for $C_{34}H_{47}N_5O_8S_2$: C, 56.9; H, 6.60; N, 9.8. Found: C, 57.0; H, 6.83; N, 9.9.

***N-tert*-Butoxycarbonyl-L-glutaminyl-*S*-benzyl-L-cysteinyl-*S*-benzyl-L-cysteinyl-L-alanine Hydrazide (VIII).** A solution of compound VII (10 g) in Me_2Fam (50 ml) containing hydrazine hydrate (5 ml) was stored at room temperature for 24 hr. The precipitated product was isolated by filtration and washed with methanol: wt 9.8 g (98%); mp 235–237° dec; $[\alpha]^{26}D$ −32.9° (*c* 1, Me_2Fam).

Anal. Calcd for $C_{33}H_{47}N_7O_7S_2$: C, 55.2; H, 6.60; N, 13.7. Found: C, 55.1; H, 6.48; N, 14.0.

Amino acid analysis of an acid hydrolysate showed the following composition expressed in molar ratios: $Glu_{1.0}$, $Ala_{1.0}$, *S*-benzylcysteine$_{1.5}$ (average amino acid recovery, 96%).

***N-tert*-Butoxycarbonyl-L-glutaminyl-*S*-benzyl-L-cysteinyl-*S*-benzyl-L-cysteinyl-L-alanylglycyl-L-valyl-*S*-benzyl-L-cysteinyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-γ-benzyl-L-glutamyl-L-asparaginyl-L-tyrosyl-*S*-benzyl-L-cysteinyl-L-asparagine *p*-Nitrobenzyl Ester (IX).** A solution of the protected tridecapeptide IV (3 g) in trifluoroacetic acid (25 ml) was stored at room temperature for 1.5 hr and then poured into ether (500 ml). The precipitated partially protected tridecapeptide derivative was isolated, washed with ether, and dried over KOH in vacuo. To a cold solution of this product in a mixture of Me_2Fam (10 ml) and Me_2SO (10 ml) containing triethylamine (0.42 ml) was added the tetrapeptide azide prepared as follows. The protected tetrapeptide hydrazide VIII (1.22 g) was dissolved in a mixture of Me_2Fam (10 ml) and Me_2SO (10 ml) containing 1 *N* HCl in Me_2Fam (3.4 ml). This solution was cooled to −10°, mixed with isoamyl nitrite (0.23 ml), and after 5 min was neutralized with triethylamine (0.48 ml) and added to the solution of the tridecapeptide derivative prepared as described previously. After 72 hr at 4° the reaction mixture was poured into methanol (800 ml) and the precipitated protected heptadecapeptide was isolated by centrifugation, washed with methanol, and dried: wt 3.3 g (85%); mp 277–278° dec; $[\alpha]^{26}D$ −31.1° (*c* 1, Me_2SO).

Anal. Calcd for $C_{125}H_{160}N_{22}O_{31}S_4$: C, 57.9; H, 6.21, N, 11.9. Found: C, 57.6; H, 6.33; N, 11.8.

Amino acid analysis of an acid hydrolysate showed the following composition expressed in molar ratios: $Asp_{2.0}$, $Ser_{0.8}$, $Glu_{3.0}$, $Gly_{1.0}$, $Ala_{1.0}$, $Val_{1.1}$, $Leu_{1.9}$, $Tyr_{1.4}$, *S*-benzylcysteine$_{3.5}$ (average amino acid recovery, 90%).

***N*-Benzyloxycarbonylsarcosyl-L-isoleucyl-L-valyl-γ-*tert*-butyl-L-glutamic Acid Methyl Ester (X).** A solution of *N*-benzyloxycarbonyl-L-isoleucyl-L-valyl-γ-*tert*-butyl-L-glutamic acid methyl ester[31] (5.6 g) in methanol (100 ml) was hydrogenated over 10% palladium-on-charcoal catalyst (1 g). After 1 hr, the catalyst was filtered off and the filtrate evaporated to dryness in vacuo. To a cold solution of this residue in Me_2Fam (30 ml) *N*-benzyloxycarbonylsarcosine[32] (2.2 g) was added followed by *N,N'*-dicyclohexylcarbodimide (2.5 g). After 24 hr at room temperature the precipitated by-product was filtered off and the filtrate was concentrated to dryness under reduced pressure. Trituration of the residue with petroleum ether and reprecipitation of the resulting material from ethyl acetate–petroleum ether gave 3.1 g (50%) of product: mp 185–188°; $[\alpha]^{26}D$ −58.7° (*c* 1, methanol).

Anal. Calcd for $C_{32}H_{50}N_4O_9$: C, 60.6; H, 7.94; N, 8.8. Found: C, 60.5; H, 7.87; N, 9.0.

N-Benzyloxycarbonylsarcosyl-L-isoleucyl-L-valyl-γ-*tert*-butyl-L-glutamic Acid Hydrazide (XI). Compound X (2 g) was dissolved in Me_2Fam (15 ml), hydrazine hydrate (1 ml) added, and the solution stirred at room temperature for 48 hr. The precipitated product was collected by filtration and washed with water: wt 1.9 g (95%); mp 245–247° dec; $[\alpha]^{26}D$ −19.4° (*c* 1, Me_2Fam).

Anal. Calcd for $C_{31}H_{50}N_6O_8$: C, 58.7; H, 7.94; N, 13.2. Found: C, 58.8; H, 7.87; N, 13.1.

Amino acid analysis after acid hydrolysis gave the following molar ratios: $Glu_{1.0}$, $Val_{1.0}$, $Ile_{1.0}$, $Sar_{1.0}$ (average amino acid recovery, 93%).

N-Benzyloxycarbonylsarcosyl-L-isoleucyl-L-valyl-L-glutamyl-L-glutaminyl-S-benzyl-L-cysteinyl-S-benzyl-L-cysteinyl-L-alanylglycyl-L-valyl-S-benzyl-L-cysteinyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-γ-benzyl-L-glutamyl-L-asparaginyl-L-tyrosyl-S-benzyl-L-cysteinyl-L-asparagine *p*-Nitrobenzyl Ester (XII). A solution of the protected heptadecapeptide IX (0.52 g) in trifluoroacetic acid (10 ml) was stored at room temperature for 30 min and then poured into ether (300 ml). The precipitated product was collected, washed with ether, and dried over KOH in vacuo. To a solution of this solid in a mixture of Me_2Fam (7 ml) and Me_2SO (7 ml) cooled to 0°, triethylamine (0.03 ml) was added followed by the protected tetrapeptide azide prepared as follows. Tetrapeptide hydrazide XI (0.26 g) was dissolved in a mixture of Me_2Fam (5 ml) and Me_2SO (5 ml). To this solution cooled to −10°, 1 *N* HCl in Me_2Fam (0.8 ml) was added, followed by isoamyl nitrite (0.06 ml). After 5 min the reaction mixture was neutralized with triethylamine (0.14 ml) and subsequently added to the partially protected heptadecapeptide prepared as described previously. The reaction mixture was stirred for 48 hr at 4° and then poured into methanol (300 ml). The precipitated heneicosapeptide XIa was isolated by centrifugation, washed with methanol, and dried. A solution of this solid in trifluoroacetic acid (15 ml) was stored at room temperature for 30 min and then poured into ether (300 ml). The precipitated partially protected heneicosapeptide XII was isolated, washed with ether, and dried: wt 0.43 g (70%); mp 287–291° dec; $[\alpha]^{26}D$ −41.2° (*c* 1, hexamethylphosphoramide).

Anal. Calcd for $C_{147}H_{190}N_{26}O_{37}S_4 \cdot 4H_2O$: C, 56.7; H, 6.41; N, 11.7; O, 21.1. Found: C, 56.6; H, 6.64; N, 12.1; O, 21.1.

An acid hydrolysate gave the following composition expressed in molar ratios: $Asp_{2.0}$, $Ser_{0.8}$, $Sar_{1.0}$, $Glu_{4.1}$, $Gly_{1.1}$, $Ala_{1.1}$, $Val_{1.8}$, $Ile_{0.7}$, $Leu_{1.8}$, $Tyr_{1.2}$, *S*-benzylcysteine$_{3.2}$ (average amino acid recovery, 92%.)

Sarcosyl-L-isoleucyl-L-valyl-L-glutamyl-L-glutaminyl-S-sulfo-L-cysteinyl-S-sulfo-L-cysteinyl-L-alanylglycyl-L-valyl-S-sulfo-L-cysteinyl-L-seryl-L-leucyl-L-tyrosyl-L-glutaminyl-L-leucyl-L-glutamyl-L-asparaginyl-L-tyrosyl-S-sulfo-L-cysteinyl-L-asparagine (Sheep Insulin [Sar^1]A Chain S-Sulfonate) (XIII). The reduction of the partially protected heneicosapeptide XII and the oxidative sulfitolysis of the reduced product was carried out as described previously.[14] Briefly, the thoroughly dry compound XII (300 mg) was dissolved in anhydrous liquid ammonia (250 ml) in a 500-ml round-bottomed flask fitted for magnetic stirring. Reduction was accomplished at the boiling point of the solution, by adding small pieces of sodium into the solution until a faint blue color appeared throughout. The blue color was allowed to persist for 10 sec and was then discharged by the addition of a few crystals of ammonium chloride. The solution was concentrated at atmospheric pressure to dryness and the residue dried in vacuo and dissolved in 8 *M* guanidine hydrochloride (25 ml). To this solution, adjusted to pH 8.9 with 1 *N* NH_4OH or acetic acid (depending on the pH of the solution) sodium sulfite (1.5 g) and freshly prepared sodium tetrathionate (0.75 g) were added. The reaction mixture was stirred at room temperature for 20 hr and then placed in a Visking 18/32 dialysis tubing and dialyzed against four changes of distilled water (4 l. each) at 4° for 20 hr. Upon lyophilization of the dialysate, the crude [Sar^1]A chain *S*-sulfonate (XIII) was obtained as a white powder. For purification the lyophilized material was dissolved in 0.015 *M* ammonium bicarbonate (5 ml) and chromatographed on a Sephadex G-15 column (2.2 × 45 cm), equilibrated, and eluted with 0.015 *M* ammonium bicarbonate at a flow rate of about 40 ml/hr. The elution pattern of this column, as determined by monitoring the effluent by a Gilford recording spectrophotometer, indicated the presence of a single component. Lyophilization of the effluent gave the purified [Sar^1]A chain *S*-sulfonate (XIII) as a white powder: wt 218 mg (82% based on XII used); $[\alpha]^{26}D$ −82.1° (*c* 0.1, water). Amino acid analysis of an acid hydrolysate gave the molar ratios shown in Table I. Digestion of the synthetic chain with leucine aminopeptidase and amino acid analysis of the digest gave the molar ratios shown in Table I. It is apparent that the synthetic product was completely digested by the enzyme indicating that the stereochemical purity of the constituent amino acid was preserved during the synthetic processes. On thin-layer chromatography in systems A (Figure 1) and **B** (Figure 2), the synthetic chain exhibited a single Pauly-positive spot and had a mobility similar to that of the natural bovine A chain *S*-sulfonate. On thin-layer electrophoresis in 0.5 *N* acetic acid and 3500 V (Figure 3) and in 0.1 *M* NH_4HCO_3 (pH 10) and 2900 V (Figure 4), the synthetic chain moved as a single component and had a mobility slightly different than that of the corresponding natural bovine chain.

S-Sulfonated Derivatives of the A and B Chains of Bovine Insulin. These compounds were prepared by oxidative sulfitolysis of bovine insulin followed by electrophoretic separation of the resulting S-sulfonated chains as described previously.[13] The B chain of bovine insulin is identical with the corresponding chain of sheep insulin.[11,12]

Synthesis and Isolation of [Sar^1-A] Sheep Insulin. The synthesis of this analog by the interaction of the sulfhydryl form of the [Sar^1]A chain with the S-sulfonated form of the B chain was carried out as described previously.[19,20] In a typical experiment 20 mg of the S-sulfonated [Sar^1]A chain was converted to the sulfhydryl form upon exposure to 2-mercaptoethanol (6–8 min, pH 5.0, 100°) and allowed to react with 5 mg of the S-sulfonated B chain. The combination mixture was processed as we have reported in detail previously.[19,21] Isolation of the insulin analog from the combination mixture was accomplished by chromatography on a 0.9 × 23 cm CM-cellulose column with an acetate buffer (0.024 *M*, pH 3.3) and an exponential NaCl gradient, according to the procedure we have reported previously.[19,21] Chromatography of two combination mixtures, each corresponding to the amounts of materials indicated above, gave the pattern shown in Figure 5. The [Sar^1-A] sheep insulin is eluted with the application of the gradient and was isolated from the effluent (175–250 ml) via picrate as the hydrochloride (1.23 mg) following the procedure described previously.[21]

Amino acid analysis of this analog after acid hydrolysis gave a composition expressed in molar ratios shown in Table II, in good agreement with the theoretically expected values. On thin-layer electrophoresis in 0.5 *N* acetic acid and 3500 V the insulin analog moved as a single component (Pauly reaction) and had a slightly different mobility than bovine insulin (Figure 6). The [Sar^1-A] sheep insulin by the mouse convulsion assay method was found to possess a potency of 20 IU/mg and by the radioimmunoassay method had a potency of 9 IU/mg.

Acknowledgments. The authors express their appreciation to Dr. Agnes Horvat and Mrs. Edith Li for the radioimmunoassays, to Mrs. Mei-Ying Liu for the biological assays, and to Mrs. Kay Tai for the amino acid analyses and the enzymatic analysis.

References and Notes

(1) This work was supported by the National Institute for Arthritis and Metabolic Diseases, U.S. Public Health Service (AM 12925).
(2) For the previous paper of this series, see A. Cosmatos and P. G. Katsoyannis, *J. Biol. Chem.*, in press.
(3) For a review see T. Blundell, G. G. Dodson, D. C. Hodgkin, and D. Mercola, *Adv. Protein Chem.*, **26**, 280 (1972).
(4) P. G. Katsoyannis and C. Zalut, *Biochemistry*, **11**, 3065 (1972).
(5) P. G. Katsoyannis and C. Zalut, *Biochemistry*, **11**, 1128 (1972).
(6) D. Brandenburg, H.-G. Gattner, M. Weinert, L. Herbertz, H. Zahn, and A. Wollmer, *Excerpta Med. Found. Int. Congr. Ser.*, **231**, 363 (1971).
(7) D. Brandenburg, *Hoppe-Seyler's Z. Physiol. Chem.*, **350**, 741 (1969).
(8) D. G. Lindsay and S. Shall, *Biochem. J.*, **121**, 737 (1971); **115**, 587 (1969).
(9) D. G. Lindsay, O. Loge, W. Losert, and S. Shall, *Biochim. Biophys. Acta*, **263**, 658 (1972).
(10) M. Weinert, K. Kircher, D. Brandenburg, and H. Zahn, *Hoppe-Seyler's Z. Physiol. Chem.*, **352**, 719 (1971).
(11) F. Sanger and H. Tuppy, *Biochem. J.*, **49**, 463, 481 (1951).
(12) H. Brown, F. Sanger, and R. Kitai, *Biochem. J.*, **60**, 556 (1955).
(13) P. G. Katsoyannis, A. Tometsko, C. Zalut, S. Johnson, and A. Trakatellis, *Biochemistry*, **6**, 2635 (1967).
(14) P. G. Katsoyannis, A. Tometsko, C. Zalut, and K. Fukuda, *J. Am. Chem.*

Soc., 88, 5625 (1966).
(15) P. G. Katsoyannis, A. Tometsko, and C. Zalut, *J. Am. Chem. Soc.*, **89**, 4505 (1967).
(16) P. G. Katsoyannis, Y. Okada, and C. Zalut, *Biochemistry*, **12**, 2516 (1973).
(17) R. H. Sifferd and V. duVigneaud, *J. Biol. Chem.*, **108**, 753 (1935).
(18) J. L. Bailey and R. D. Cole, *J. Biol. Chem.*, **234**, 1733 (1959).
(19) P. G. Katsoyannis, A. Trakatellis, S. Johnson, C. Zalut, and G. Schwartz, *Biochemistry*, **6**, 2642 (1967).
(20) P. G. Katsoyannis and A. Tometsko, *Proc. Natl. Acad. Sci. U.S.A.*, **55**, 1554 (1966).
(21) P. G. Katsoyannis, A. Trakatellis, C. Zalut, S. Johnson, A. Tometsko, G. Schwartz, and J. Ginos, *Biochemistry*, **6**, 2656 (1967).
(22) A. Tometsko and N. Delihas, *Anal. Biochem.*, **18**, 72 (1967).
(23) (a) D. H. Spackman, W. H. Stein, and S. Moore, *Anal. Chem.*, **30**, 1190 (1958); (b) R. L. Hill and E. L. Smith, *J. Biol. Chem.*, **228**, 577 (1957).
(24) A. Gilman, F. S. Philips, and E. S. Koelle, *Am. J. Physiol.*, **148**, 348 (1946).
(25) O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, *J. Biol. Chem.*, **193**, 265 (1951).
(26) C. N. Hales and P. J. Randle, *Biochem. J.*, **88**, 137 (1963).
(27) L. Zervas, I. Photaki, and N. Ghelis, *J. Am. Chem. Soc.*, **85**, 1337 (1963).
(28) E. Sandrin and R. A. Boissonas, *Helv. Chim. Acta*, **46**, 1637 (1963).
(29) S. Hörnle, *Hoppe-Seyler' Z. Physiol. Chem.*, **48**, 1355 (1967).
(30) H. C. Beyerman, C. A. M. Boers-BooneKamp, and H. Maasen van den Brink-Zimmermannova, *Recl. Trav. Chim. Pays-Bas*, **87**, 257 (1968).
(31) H. Zahn, H. Bremer, W. Sroka, and J. Melenhofer, *Z. Naturforsch.*, **20B**, 646 (1965).
(32) R. S. Timpson, *J. Org. Chem.*, **26**, 4698 (1961).

Biosynthesis of Anthramycin. Determination of the Labeling Pattern by the Use of Radioactive and Stable Isotope Techniques

Laurence H. Hurley,*[1a] Milton Zmijewski,[1a] and Ching-Jer Chang[1b]

Contribution from the Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Kentucky, Lexington, Kentucky 40506, and the Department of Medicinal Chemistry and Pharmacognosy, School of Pharmacy and Pharmacal Sciences, Purdue University, West Lafayette, Indiana 47907. *Received January 2, 1975*

Abstract: The building blocks for anthramycin, an antitumor antibiotic produced by a strain of *Streptomyces refuineus*, have been shown to be L-tryptophan, probably via 3-hydroxyanthranilic acid, L-tyrosine which loses two of its aromatic carbons, and L-methionine which contributes two methyl groups. While one of the two methyl groups is transferred intact, the other loses all of its hydrogens and becomes the carbonyl of an amide group. A mechanism involving extradiol cleavage of Dopa is proposed on the basis of double labeling and stable isotope experiments. A general scheme for the biosynthetic origin of the C_3-proline moieties of anthramycin, lincomycin A, and sibiromycin and the C_2-proline moieties of tomaymycin and lincomycin B is proposed.

Anthramycin (Ia) is an antitumor antibiotic produced by *Streptomyces refuineus* var *thermotolerans*.[2] The structure and stereochemistry of anthramycin have been elucidated[3,4] and the total synthesis of anthramycin has been reported.[5] Anthramycin has been shown to possess antitumor,[6] antibiotic,[6] amebicidal,[6] and chemosterilant properties.[7] All the biological properties of the antitumor agent can be attributed to the effect of anthramycin on nucleic acid biosynthesis. Anthramycin appears to act by virtue of its ability to bind to DNA and therefore interfere with the function of DNA.[8-10] Anthramycin causes inhibition of DNA and RNA synthesis in both bacterial and mammalian cells. However, the synthesis of RNA is most powerfully affected.[11]

The biosynthetic origin of the non-4-methyl-3-hydroxyanthranilic acid part of anthramycin was of prime interest in this investigation. A similarly structured element can be found in tomaymycin (II),[12] sibiromycin (III),[13] and the lincomycin group of antibiotics (IVa,b).[14] (See Chart I.) This structural similarity might suggest a common biogenetic origin for all of these antibiotics. The biogenetic origin of the propylproline group of lincomycin A has been shown to be tyrosine and two one-carbon units via methionine.[15] In this case the *N*-methyl and terminal methyl groups were those derived from methionine. In the case of tomaymycin, we have established that this antibiotic is derived in an analogous way to anthramycin.[16]

In this paper we wish to report on the determination of the biogenetic origin of anthramycin. Some of these results have been communicated in preliminary form.[17]

Results

Biosynthetic experiments were carried out with shake cultures of the anthramycin-producing *S. refuineus* strain NRRL 3143 grown in a complex medium. The amount of anthramycin produced at the time of harvest was about 80 μg/ml, as determined by a spectroscopic method.[18] The antibiotic was isolated from the culture medium by extraction with butanol and further purified if necessary by countercurrent distribution.

The time course of anthramycin production was followed by measuring spectroscopically the amount of anthramycin produced and the incorporation of radioactive methionine into the antibiotic at different times during the culture period. As shown in Figure 1, maximum production of anthramycin occurred at 15 hr and maximum incorporation of methionine into anthramycin occurred when the precursor was added at 12 hr.

The efficiency of incorporation of potential radioactive precursors into anthramycin was evaluated as follows. The thin-layer chromatograms of the methanol solubilized extracts were scanned for radioactivity and the areas under the peaks were integrated to obtain the percentage of the radioactivity of the extract residing in anthramycin. From this value and from the total radioactivity of the methanol extract, the total radioactivity incorporated into anthramycin was calculated, which was related to the total radioactivity of the precursor fed to give the percentage incorporation. This latter figure was also obtained independently by recrystallizing to constant specific activity those samples of