Pliskin, M. E., Ginsberg, S. M., & Carp, N. (1980) *Transplantation 29*, 255–258.
Ruegg, U. T., & Rudinger, J. (1974) *Int. J. Pept. Protein Res. 6*, 447–456.
Sanger, F., & Thompson, E. O. P. (1963) *Biochim. Biophys. Acta 71*, 468–471.
Shing, Y., Folkman, J., Sullivan, R., Butterfield, C., Murray, J., & Klagsbrun, M. (1984) *Science (Washington, D.C.) 223*, 1296–1299.
Strydom, D. J., Fett, J. W., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry* (second of three papers in this issue).
Tsukamoto, K., & Sugino, Y. (1979) *Cancer Res. 39*, 1305–1309.
Tuan, D., Smith, S., Folkman, J., & Merler, E. (1973) *Biochemistry 12*, 3159–3165.
Vallee, B. L., Riordan, J. F., Lobb, R. R., Higashi, N., Fett, J. W., Crossley, G., Buhler, R., Budzik, G., Breddam, K., Bethune, J. L., & Alderman, E. M. (1985) *Experientia 41*, 1–15.
Weiss, J. B., Brown, R. A., Kumar, S., & Phillips, P. (1979) *Br. J. Cancer 40*, 493–496.

Amino Acid Sequence of Human Tumor Derived Angiogenin†

Daniel J. Strydom, James W. Fett, Roy R. Lobb, Edward M. Alderman, J. Lemuel Bethune, James F. Riordan, and Bert L. Vallee*

Center for Biochemical and Biophysical Sciences and Medicine and Departments of Biological Chemistry, Pathology, and Radiology, Harvard Medical School, and Brigham and Women's Hospital, Boston, Massachusetts 02115

Received May 9, 1985

ABSTRACT: The amino acid sequence and disulfide bond pairing of human tumor derived angiogenin, the first tumor angiogenesis factor to be isolated in pure form from human sources, have been determined by conventional sequencing techniques adapted and applied to nanomole and subnanomole levels of material. Angiogenin, obtained from conditioned media of a human colonic adenocarcinoma cell line, is a single-chain protein consisting of 123 amino acids with the following sequence: <Glu^{1}-Asp-Asn-Ser-Arg-Tyr-Thr-His-Phe-Leu-Thr-Gln-His-Tyr-Asp^{15}-Ala-Lys-Pro-Gln-Gly-Arg-Asp-Asp-Arg-Tyr-Cys-Glu-Ser-Ile-Met^{30}-Arg-Arg-Arg-Gly-Leu-Thr-Ser-Pro-Cys-Lys-Asp-Ile-Asn-Thr-Phe^{45}-Ile-His-Gly-Asn-Lys-Arg-Ser-Ile-Lys-Ala-Ile-Cys-Glu-Asn-Lys^{60}-Asn-Gly-Asn-Pro-His-Arg-Glu-Asn-Leu-Arg-Ile-Ser-Lys-Ser-Ser^{75}-Phe-Gln-Val-Thr-Thr-Cys-Lys-Leu-His-Gly-Gly-Ser-Pro-Trp-Pro^{90}-Pro-Cys-Gln-Tyr-Arg-Ala-Thr-Ala-Gly-Phe-Arg-Asn-Val-Val-Val^{105}-Ala-Cys-Glu-Asn-Gly-Leu-Pro-Val-His-Leu-Asp-Gln-Ser-Ile-Phe^{120}-Arg-Arg-Pro^{123}-OH. Three disulfide bonds link the half-cystinyl residues 26–81, 39–92, and 57–107. The sequence is homologous to that of the pancreatic ribonucleases with 35% identity and many of the remaining residues conservatively replaced. Similarities are especially apparent around the major active-site residues His-12, Lys-41, and His-119 of ribonuclease which are conserved as are three of the four disulfide bonds. The complete chemical characterization of a unique human organogenic messenger molecule, i.e., one that can induce organ formation, accomplishes the first major objective of this long-term investigation of organogenesis in general and angiogenesis in particular. The unexpected homology to ribonuclease suggests novel approaches to the investigation of the biological process of angiogenesis.

The preceding paper describes the isolation and purification of an angiogenic protein from human tumor cells (Fett et al., 1985b). We have designated this protein human tumor *angiogenin* to differentiate it from other angiogenic factors encountered in normal or pathological sources. It is capable of inducing blood vessel growth in both the chick embryo chorioallantoic membrane and the rabbit cornea. It is secreted by HT-29 human adenocarcinoma cells into the extracellular medium from which it can be isolated in yields of about 0.5 μg/L. Detailed physical and chemical characterization has been hampered owing to this limited availability. However, recent experimental advances have facilitated both isolation of angiogenin and examination of its structural properties. It is a single polypeptide of M_r 14400 with a blocked amino terminus and a proline carboxyl terminus (Fett et al., 1985b).

The present paper describes the results of amino acid sequence studies carried out by Edman degradation chemistry and other procedures adapted to nanomolar and subnanomolar amounts of material. The sequence of four polypeptide segments accounts for all 123 amino acid residues. The amino- and carboxyl-terminal segments are identified on the basis of the unique nature of their respective terminal residues. The alignment of the remaining two segments to give the overall primary structure is based on a surprising homology to the family of pancreatic ribonucleases, and on the corresponding cDNA sequence determined concurrently.

Angiogenin is the first human tumor derived angiogenic factor to be isolated in pure form and sequenced. Other mediators of blood vessel growth have also been described but are as yet uncharacterized, and hence, their relationship to angiogenin is still unknown. We expect, however, that angiogenin will be a member of what will prove to be a new and important class of proteins that can induce organ formation, i.e., organogenic messenger molecules.

Experimental Procedures

Materials. Angiogenin was isolated from the serum-free conditioned medium of an established human colonic adeno-

† This work was supported by funds from the Monsanto Co. under agreements with Harvard University.

* Address correspondence to this author at the Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School.

0006-2960/85/0424-5486$01.50/0

carcinoma line (HT-29) as previously described (Fett et al., 1985b), but employing 3:2:2 2-propanol/acetonitrile/water containing 0.08% trifluoroacetic acid (TFA)[1] in the final HPLC purification step.

Trifluoroacetic acid was sequenal grade from Pierce Chemical Co.; water was either from J. T. Baker Chemical Co. (HPLC grade) or from an in-house supply (Milli-Q, Millipore). Both acetonitrile (J. T. Baker Chemical Co.) and 2-propanol (Millipore, Waters Associates) were of HPLC grade.

Sequencer reagents were obtained from Beckman Instruments, Inc., and sequencer solvents from Burdick and Jackson Laboratories, Inc., Muskegon, MI.

Glassware used for handling protein solutions was always siliconized by treatment with dichlorodimethylsilane (Sigma Chemical Co.).

Amino Acid Analysis. Peptide samples were dried in Pyrex–Corning culture tubes (6 × 50 mm) which had been pyrolyzed in a muffle furnace at 500 °C. Hydrolyses were performed by a slight modification of the method of Dreyer & Bynum (1967). The tubes containing the dry samples were placed in a 15-cm-diameter dry-sealing vacuum desiccator (Wheaton) over 10 mL of 6 N HCl containing 0.5% phenol. A Viton O ring provided the seal, and a stainless-steel coupling (Fisher Scientific Co.) clamped the lid to the body of the desiccator. The desiccator was alternately evacuated and purged with N_2 (3×) and finally evacuated to 0.01 torr and then heated in an oven at 110 °C for 18–20 h.

After hydrolysis, the tubes were kept briefly in a dry desiccator under vacuum. Hydrolysates were analyzed for their amino acid content by either of two procedures. The first employed ion-exchange chromatography followed by reaction with ninhydrin, utilizing a Durrum D500 amino acid analyzer. The second involved derivatization of hydrolysates by the Picotag procedure (Bidlingmeyer et al., 1984) prior to analysis by reversed-phase chromatography on Novapak C18 columns (Waters Associates) using a Waters Associates HPLC system.

Microsequencing. A Beckman 890C spinning-cup sequencer equipped with a cold trap attachment was used. Polybrene (5 mg) (Klapper et al., 1978; Tarr et al., 1978) and Gly-Tyr (50 nmol) were loaded into the cup and then degraded through five cycles before the peptide or protein samples were introduced. To preclean the sample, one cycle of degradation was performed without the addition of phenyl isothiocyanate. The degradation was carried out by using Beckman program 111978 for 1 M quadrol buffer but actually employing 0.1 M quadrol. Reagents and solutions were pretreated as described (Fett et al., 1985a).

Anilinothiazolinones were converted to phenylthiohydantoins (Strydom & Bazzone, 1985; Fett et al., 1985a), and the PTH-amino acids were identified and quantitated by HPLC using one or both of two separation systems. System 1 used an IBM octadecylsilane column (25 × 0.4 cm, 5-μm particles) at room temperature. The solvents and gradient systems were the following: solvent A, 20% methanol, 0.3% triethylamine, and 0.6% acetic acid, pH 4; solvent B, 50% 2-propanol, 0.3% triethylamine, and 0.6% acetic acid. Consecutive gradients from 96% solvent A to 80% solvent A in 5 min (curve 7, concave) and from 80% solvent A to 47% solvent A for an additional 8 min (curve 6, linear) allow a separation in 25 min at 1 mL/min flow rate. System 2 employed a Waters Novapak (10 × 0.4 cm) column at 40 °C (S. A. Cohen, personal communication).

Reduction and S-Sulfopropylation. Angiogenin was reduced and sulfopropylated as described by Fett et al. (1985b). The reaction mixture containing microgram quantities of derivatized protein was desalted by molecular-sieve HPLC on a Waters I-125 column in 17% 2-propanol, 17% acetonitrile, and 0.09% TFA.

Trypsin Digestion. Angiogenin in lots of 12–66 μg was digested with HPLC-purified trypsin (Titani et al., 1982), 2–3% by weight, in 100 μL of 0.1 M *N*-ethylmorpholine buffer, pH 8.5, 35 °C, for 18–20 h under N_2. The digest was applied directly to an HPLC column for fractionation as described under Results.

Cyanogen Bromide Digest. Cyanogen bromide (Pierce Chemical Co.), 15 mg, was dissolved in 100 μL of 70% TFA and after 10 min at room temperature was added to 15 μg of reduced and sulfopropylated angiogenin. The tube was flushed with N_2, stoppered, and kept at room temperature for 24 h in the dark. A few drops of water were added, and the mixture was dried, redissolved in 500 μL of TFA, and applied to the sequencer cup.

Pyroglutaminase Digest. Angiogenin (140 μg) was digested with 15 μg of bovine liver pyroglutaminase (Sigma Chemical Co.) (Podell & Abraham, 1978; Zalut et al., 1980) in 100 μL of 0.05 M phosphate buffer, pH 8, containing 0.01 M β-mercaptoethanol and 0.001 M EDTA for 18–20 h at 30 °C. The reaction mixture was desalted by reversed-phase chromatography on a SynChropak RP-P column (SynChrom Inc., Linden, IN) employing a 30-min gradient from 10% solvent B to 40% solvent B. Solvent A was 0.1% TFA, and solvent B was 3:2:2 2-propanol/acetonitrile/water containing 0.08% TFA.

Cleavage with Hydroxylamine. Angiogenin was cleaved at the two Asn-Gly sequences by hydroxylamine according to a modification of the method of Bornstein (1970). The protein (35 μg) was suspended in 200 μL of 2 M hydroxylamine hydrochloride (Aldrich Chemical Co.) (pH adjusted to 9.2 with KOH) containing 0.2 M potassium carbonate (final pH 9.6) and dissolved by the addition of 200 mg of guanidine hydrochloride. After 18 h at 23 °C, the precipitated salts were dissolved by the addition of 220 μL of water, and the mixture was desalted by molecular-sieve HPLC on a Waters I-125 column in 17% 2-propanol, 17% acetonitrile, and 0.09% TFA. The fraction containing salt was taken to dryness under a stream of N_2 and rechromatographed as for the pyroglutaminase digest, except that the gradient was from 5% solvent B to 50% solvent B in 60 min.

Thermolysin Digest and Isolation of Peptides. Thermolysin (Sigma Chemical Co.) was recrystallized according to Holmquist & Vallee (1974); it was dissolved in a minimal volume of 5 M NaBr, 0.01 M $CaCl_2$, and 0.01 M Tris buffer, pH 8, and diluted into 0.05 M Tris and 0.02 M $CaCl_2$, pH 8.0. Reduced and sulfopropylated angiogenin (10 μg) was dissolved in 50 μL of the latter buffer, and 0.2 μg of thermolysin in solution was added. After 18 h at 40 °C, the digest was loaded directly onto an Ultrapore C3 (Beckman Instruments) column in 0.1% TFA. A 60-min gradient in 2-propanol/acetonitrile did not yield any peptide fractions. The breakthrough fraction was pooled and loaded onto an IBM C18 (5-μm) column. A 60-min gradient from 3.75% to 18% acetonitrile in 0.1 M perchlorate/phosphate buffer, pH 2.5, at a flow rate of 0.8

[1] Abbreviations: HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; PTH, phenylthiohydantoin; C18, octadecylsilane; C3, propylsilane; EDTA, ethylenediaminetetraacetic acid; HFBA, heptafluorobutyric acid; PITC, phenyl isothiocyanate; DTE, dithioerythritol; Cys(Sp), *S*-(sulfopropyl)cysteine; Tris, tris(hydroxymethyl)-aminomethane; FAB, fast atom bombardment; Cys(CM), *S*-(carboxymethyl)cysteine.

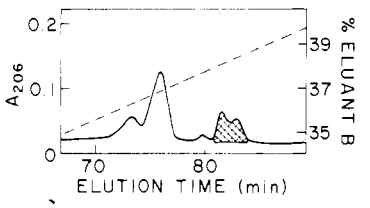

FIGURE 1: Separation and purification of angiogenin from 6 mg of CM II fraction (Fett et al., 1985b) by HPLC on a SynChropak RP-P column (0.41 × 25 cm). Solvent A was 0.1% TFA, and solvent B was 3:2:2 2-propanol/acetonitrile/water in 0.08% TFA. The gradient was linear from 20% to 40% solvent B in 90 min at a flow rate of 1 mL/min. The cross-hatched double peak is angiogenin.

mL/min was used to elute the peptide fractions.

Reduction and S-Carboxamidomethylation of NT 10. Peptide NT 10 (1 nmol) was dissolved in 100 μL of a 1:1 mixture of 1-propanol and 0.5 M sodium bicarbonate, and 10 μL of a 5% solution of tributylphosphine in 1-propanol was added to reduce the disulfide bond (Ruegg & Rudinger, 1977). After 1 h under N_2, 9 μg of iodoacetamide dissolved in 1-propanol was added. After a further 2 h, the reaction mixture was fractionated by HPLC on an Altex Ultrasphere-IP column using a 1-h linear gradient of 5–60% solvent B where solvent A was 0.1% TFA and solvent B was 3:2:2 2-propanol/acetonitrile/water containing 0.08% TFA.

RESULTS

Isolation of Angiogenin. Biologically active angiogenin (Fett et al., 1985b) was isolated as two partially resolved peaks, well separated from the preceding ones of the chromatogram (Figure 1). In preparations on columns that had seen considerable use, only one angiogenin peak was found. Since there were no significant differences between the amino acid analyses of the chromatographic fractions ranging across both peaks, their pool was used as "angiogenin".

Isolation of Tryptic Peptides. The larger peptides from tryptic digests of native angiogenin were isolated by chromatography on an Altex Ultrapore C3 column (Beckman Instruments Inc.) (Figure 2a). The breakthrough fraction was rechromatographed on an Altex Ultrasphere-IP (Beckman Instruments Inc.) column employing either a volatile (0.1% TFA) or a nonvolatile (0.1 M perchlorate/0.1% phosphate, pH 2.5) buffer (Meek, 1980). These separations are shown in Figure 2b,c. In the latter case, the peptides had to be desalted prior to sequencing by chromatography on an IBM 5-μm C18 column with 0.1% HFBA and acetonitrile as solvents. Peptide NT 1 was desalted by reversed-phase chromatography on a C18 column, using 0.1% TFA and acetonitrile as solvents.

Amino Acid Compositions of Peptides. Amino acid compositions of the tryptic peptides of angiogenin, determined on 10–100-pmol aliquots of the HPLC fractions, are listed in Table I. The analyses are not corrected for background quantities of Ser, Gly, or Glu which were found at this level in most samples isolated by HPLC.

Sequences of Tryptic Peptides. Table II presents the sequences derived from Edman degradation of each of the peptides isolated as outlined above. Unique sequences were found for peptides NT 3a, NT 4a, NT 4b, NT 5, NT 6, NT 7, NT 8, and NT 12; consequently, these were considered pure. Analysis of peptide NT 1 + 13 gives the sequence Arg-Arg. This, together with its composition, indicates that NT 1 + 13 is a mixture of Arg-Arg and the N-blocked amino-terminal peptide NT 1. Desalted peptide NT 1 was analyzed by FAB mass spectrometry and found to have an MH^+ of 602, indicating the presence of <Glu, Asp, Asn, Ser, and Arg. Peptide

Table I: Amino Acid Compositions of Tryptic Peptides of Angiogenin[a]

peptide:	1	2	3a	3b	4a	4b	5	6	7	8	9	10	11	12	1 + 13	angiogenin from sequence	angiogenin amino acid analysis[b]
$^1/_2$-Cys						nd[c] (1)					1.22 (2)	1.23 (2)	nd (2)	0.18 (1)		6	6.08
Asp	1.63 (2)		0.38		1.58 (2)	1.10 (1)	0.84 (1)		1.30 (1)	2.71 (3)	1.69 (2)		3.86 (4)	2.39 (3)	2.20 (2)	15	15.96
Thr								0.99 (1)	1.73 (2)	1.30 (1)	1.79 (2)	1.18 (1)	0.28			7	7.35
Ser	0.83 (1)		0.94 (1)	1.54 (1)		1.00	0.50	0.28	0.18	0.61	3.00 (3)	2.28 (2)	1.37 (1)	0.83 (1)	1.01 (1)	9	9.80
Glu	1.02 (1)		0.10		0.16	1.12 (1)	1.03 (1)		1.91 (2)	0.67	2.11 (2)	1.41 (1)	3.01 (3)	2.01 (2)	1.16 (1)	10	11.47
Pro		0.94 (1)			1.13 (1)				1.06 (1)			3.70 (4)	1.20 (1)	1.02 (1)		8	7.93
Gly		0.20	0.36		0.94 (1)	0.63	0.54	1.29 (1)	0.95 (1)	1.30 (1)	0.52	4.33 (3)	1.66 (1)	1.14 (1)	0.32	8	10.41
Ala		0.18	0.13			1.06 (1)		1.89 (2)	0.89 (1)	0.40		0.50	1.91 (2)	0.89 (1)	0.10	5	6.05
Val									0.10		1.08 (1)	0.28	2.70 (4)	2.84 (4)		5	4.86
Met									0.11		0.69 (1)					1	0.97
Ile			0.86 (1)	0.96 (1)	0.14	0.90 (1)			0.22	1.50 (2)	0.80 (1)		1.60 (2)	0.74 (1)		7	6.31
Leu					0.29		0.84 (1)		0.98 (1)			1.83 (2)	1.89 (2)	1.78 (2)		6	6.10
Tyr									1.85 (2)		0.77 (1)	0.94 (1)				4	3.99
Phe								1.00 (1)	1.11 (1)	0.93 (1)	1.04 (1)	0.29	0.93 (1)	0.88 (1)		5	4.90
His			0.21		0.86 (1)				1.68 (2)	1.00 (1)		0.79 (1)	0.98 (1)	0.66 (1)		6	5.72
Lys			1.02 (1)	1.02 (1)		0.92 (1)			1.06 (1)	1.11 (1)	0.96 (1)	0.97 (1)	1.17 (1)	0.12		7	7.49
Arg	1.18 (1)	1.06 (1)	0.30	4.21[d]	1.01 (1)		1.00 (1)	0.99 (1)	1.13 (1)	1.53 (1)	1.92 (2)	1.20 (1)	1.00 (1)	1.17 (1)	2.03 (1 + 2)[d]	13	13.00
Trp												+[e] (1)				1	(1)
total	5	2	3	3 + 1	6	6	4	6	16	11	19	20	26	20	5 + 2	123	

[a] Relative molar amounts of amino acids are given, based on the numbers of amino acids found upon sequencing (in parentheses). [b] Fett et al. (1985b). [c] nd, not determined. [d] These peptides also contain free Arg (3b) and Arg-Arg (13). [e] From the absorbance ratio for this peptide in the peptide maps (254 nm vs. 206 nm).

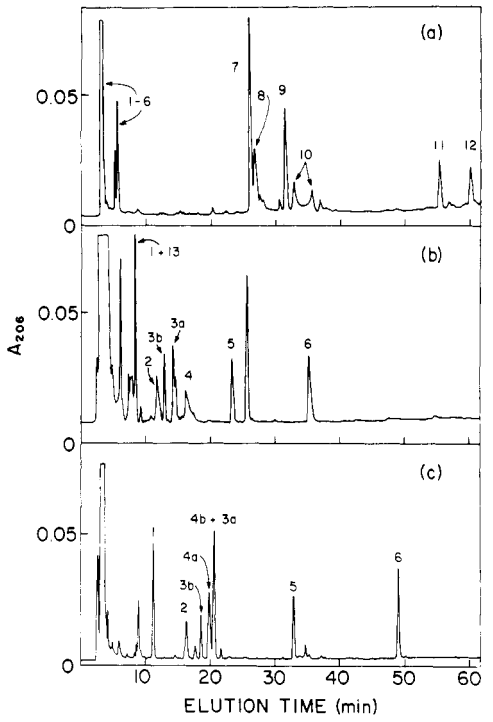

FIGURE 2: (a) Fractionation of a tryptic digest of native angiogenin by HPLC on a Beckman Ultrapore (C3) column. Solvents A and B were as in Figure 1. The gradient was linear from 5% to 95% solvent B in 3 h at a flow rate of 1 mL/min. Peptides NT 7 through NT 12 are indicated by 7–12. The breakthrough fractions are labeled 1–6. (b) Fractionation of breakthrough fractions 1–6 on a Beckman Ultrasphere (C18, IP) column. Solvents A and B were as in Figure 1. The gradient was linear from 5% to 50% solvent B in 2 h, at a flow rate of 1 mL/min. The fractions that contained significant quantities of amino acids after hydrolysis are labeled. The unlabeled peak between peptides NT 5 and NT 6 has an abnormally high absorbance at 254 nm and did not contain peptide material. It is probably due to a buffer or solvent component from prior steps. Likewise, the earliest eluting peaks only showed the presence of the usual "background" amino acids, Ser, Glu, and Gly. (c) Alternative fractionation of breakthrough fractions 1–6 on a Beckman Ultrasphere (C18) column. Solvent A was 0.1 M sodium perchlorate and 0.1% orthophosphoric acid, pH 2.5, and solvent B was 75% acetonitrile in 25% solvent A. The gradient was linear from 3% to 50% solvent B in 2 h at 1 mL/min. The fractions containing peptides are labeled.

NT 2 gives Arg on the first degradation cycle, but nothing after that. Its composition, Arg plus Pro, and the fact that the carboxyl-terminal residue of angiogenin is Pro identify the sequence of NT 2 as Arg-Pro and locate it at the carboxyl terminus of angiogenin. Peptide NT 3b gives both Arg and Ile on the first cycle but only Ser and Lys thereafter. Hence, NT 3b is probably the tripeptide Ile-Ser-Lys plus free arginine. Both NT 9 and NT 10 are mixtures of two peptides (designated NT 9′ + NT 9″ and NT 10′ + NT 10″, respectively) probably linked by a disulfide in each instance. NT 11 is also a mixture. In this case, comparison of its sequence information with that of NT 4b and NT 12 demonstrates that NT 11 consists of these two peptides linked by a disulfide.

Jointly, these peptides account for nearly the entire amino acid composition of angiogenin.

Thermolysin Peptides. A thermolysin digest of 700 pmol of reduced and S-sulfopropylated angiogenin enabled the isolation of six pure peptides (Figure 3) whose amino acid compositions were determined (Table III). Three of these were helpful in providing overlaps for the tryptic peptides.

Alignment of Tryptic Peptides. Since the amino terminus of angiogenin is blocked (Fett et al., 1985b) and the mass spectrometric evidence indicates that NT 1 contains <Glu, it appeared that this must be the N-terminal residue.

Table II: Amino Acid Sequences of Tryptic Peptides of Angiogenin Obtained by Micro-Edman Degradation

peptide	sequencer results[a]
NT 1 + 13	Arg-Arg 163 108
NT 2	Arg-Pro 20
NT 3a	Ser-Ile-Lys 890 795 163
NT 3b	(Arg+Ile)-Ser-Lys 264 273 144 43
NT 4a	Asn-Gly-Asn-Pro-His-Arg 103 168 80 139 ~50 ~20
NT 4b	Ala-Ile-Cys-Glu-Asn-Lys 570 505 100 83 100 25
NT 5	Glu-Asn-Leu-Arg 500 270 1000 70
NT 6	Ala-Thr-Ala-Gly-Phe-Arg 640 258 471 252 136 32
NT 7	Tyr-Thr-His-Phe-Leu-Thr-Gln-His-Tyr-Asp- 314 340 100 235 268 167 170 17 156 30 Ala-Lys-Pro-Gln-Gly-Arg 228 86 44 42 54 trace
NT 8	Asp-Ile-Asn-Thr-Phe-Ile-His-Gly-Asn-Lys 147 283 120 150 235 231 10 150 20 56
NT 9 (NT 9′ + 9″)	(Asp+Ser)-(Asp+Ser)-(Arg+Phe)-(Gln+Tyr)- 403 465 458 571 151 290 232 285 Val-(Thr+Glu)-(Thr+Ser)-Ile-(Met+Lys) 660 205 256 120 172 133 236 13
NT 10 (NT 10′ + NT 10″)	(Gly+Leu)-(His+Leu)-(Thr+Gly)-(Ser+Gly)- 862 909 171 745 323 861 405 833 (Ser+Pro)-Pro-(Trp+Lys)-Pro-Pro-Nothing- 455 255 460 206 12 437 382 Gln-Tyr 130 133
NT 11 (NT 4b + NT 12)	(Asn+Ala)-(Val+Ile)-Val-(Val+Glu)-(Asn+Ala)- 25 250 253 140 212 200 70 10 202 nothing-Glu-Asn-Gly-Leu 7 trace 58 56
NT 12	Asn-Val-Val-Val-Ala-Cys-Glu-Asn-Gly-Leu- 467 756 756 562 683 150 105 70 118 204 Pro-Val-His-Leu-Asp-Gln-Ser-Ile-Phe-Arg 119 212 50 188 24 20 108 36 72 33

[a] The established sequence is given, with the yield of amino acid residues at each cycle given *below* the residue, in picomoles. The parentheses indicate that two residues were found after a given degradation cycle.

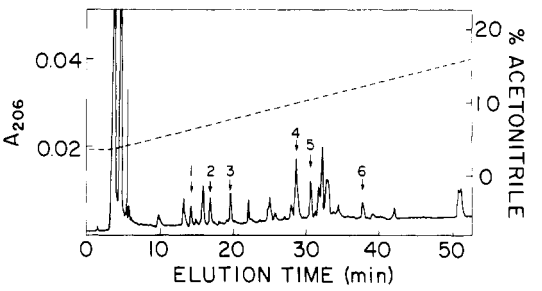

FIGURE 3: Fractionation of a thermolysin digest of 700 pmol of reduced and S-sulfopropylated angiogenin by HPLC on an IBM C18 (5-μm, 25 × 0.4 cm) column. Solvents A and B were as in Figure 2c. The gradient was from 5% to 80% solvent B in 3 h at 0.8 mL/min. No other peptides were eluted after 51 min. The elution positions of peptides L1 through L6 are indicated.

Therefore, 140 μg of the intact protein was digested with pyroglutaminase in order to render the amino terminus accessible to Edman degradation. The product, isolated by reversed-phase HPLC, was then processed through 40 cycles of degradation. As expected, the amino terminus of the digested protein was no longer blocked. Of the first 36 cycles, 35 were identified (Figure 4a), providing the sequence of NT 1, aligning NT 7 as the next peptide, and supplying the sequence of the NT 9′ component of NT 9 and part of the sequence of the NT 10′ component of NT 10 while aligning them into the amino-terminal sequence. The rest of the sequence of NT 10′ was determined from the analysis of NT 10 (Table II). In addition, NT 10′ was isolated by HPLC after

Table III: Amino Acid Composition of Some Thermolysin Peptides of Angiogenin

peptide:	L1	L2	L3	L4	L5	L6
Cys(Sp)		0.43 (1)				
Asp[a]	0.55 (1)			0.81 (1)	0.50 (1)	
Thr		1.97 (2)	0.68 (1)			
Ser	0.94 (1)	0.46	0.56	0.73 (1)	0.46	0.15
Glu[a]	0.80 (1)	0.58	0.92 (1)	0.37	0.30	
Pro						
Gly	0.38	0.48	0.71	1.39 (1)	0.63	0.31
Ala						0.71 (1)
Val		0.95 (1)				
Met						
Ile				0.68 (1)		1.00 (1)
Leu	1.17 (1)		1.03 (1)			0.16
Tyr						
Phe				1.04 (1)	0.94 (1)	0.13
His			0.92 (1)	0.78 (1)	0.14	
Lys		1.19 (1)		0.56 (1)		1.00 (1)
Arg	0.17			1.32 (1)	1.03 (1)	
total	4	5	4	8	3	3
sequence position	115–118	78–82	10–13	45–52	100–102	53–55

[a] Asp and Glu are low (by 30–50% and 5–20%, respectively) due to the presence of salt which interfered with the PITC reaction.

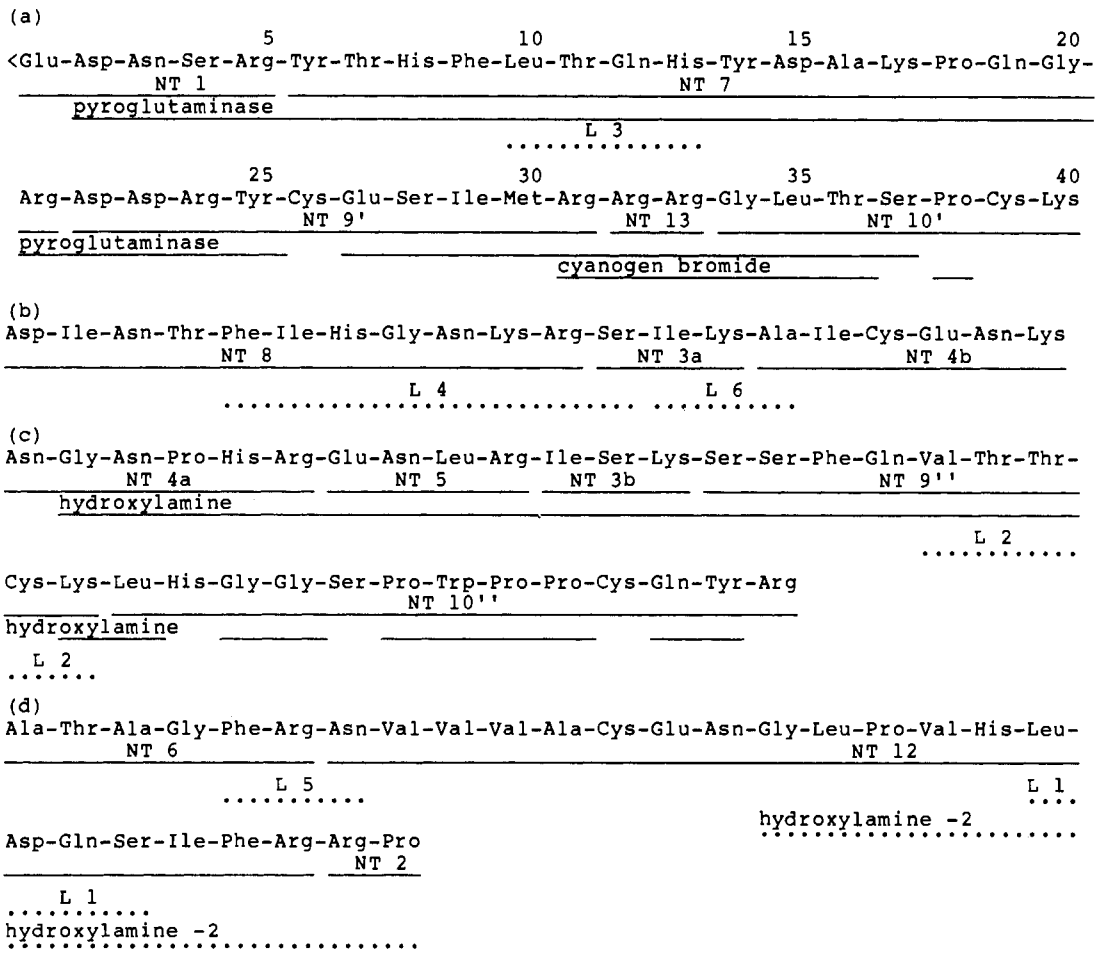

FIGURE 4: Alignment of the tryptic peptides of angiogenin. The four segments determined by protein chemistry are shown under (a), (b), (c), and (d). The extent of sequencing of the pyroglutaminase, cyanogen bromide, and hydroxylamine digested products is indicated by the solid lines, with gaps where no identification could be made. Dotted lines denote proposed areas ascribed to thermolysin (L) and hydroxylamine peptides by amino acid composition.

reduction and alkylation of NT 10 (data not shown). It had the following composition: Cys(CM), 0.46 (1); Glu, 0.4; Ser, 1.08 (1); Gly, 1.39 (1); Arg 0.23; Thr, 1.13 (1); Ala, 0.23; Pro, 1.27 (1); Leu, 0.95 (1); Lys, 0.80 (1), which is consistent with the assigned sequence (Figure 4a). Peptide NT 13 and perhaps also the free Arg of NT 3b would fit into positions 32 and 33 of the overall sequence of angiogenin.

The amino-terminal sequence demonstrates that the single methionine occupies position 30. Cyanogen bromide cleavage of the chain at that point allowed extension of its sequence for seven of the first eight residues following Met-30 as indicated in Figure 4a.

The overlap for peptides NT 3a and NT 4b in the order 3a-4b is provided by thermolysin peptide L6 (Ile, Lys, and

Ala), which by thermolysin specificity (Feder & Schuck, 1970) must have the sequence Ile-Lys-Ala. Thermolysin specificity also dictates that the Ala of L6 cannot be part of NT 6. This location of L6 is also consistent with the sequence of peptide L4 whose composition identifies it as extending from the Phe in peptide NT 8 through the Lys-Arg carboxyl terminus to the Ser of NT 3a. This, therefore, aligns peptides NT 8-3a-4b in that order (Figure 4b).

The existence of two Asn-Gly sequences in angiogenin, in peptides NT 4a and NT 12, respectively, provided an opportunity for chemical cleavage of the polypeptide chain at these two positions by the hydroxylamine method of Bornstein (1970). The hydroxylamine digest was desalted by molecular-sieve HPLC and the protein peak sequenced directly. A single clear sequence of 27–33 residues was obtained, enabling the alignment of peptides NT 4a-5-3b-9″-10″, in that order (Figure 4c), and thereby placing an additional block of 35 residues in the sequence. The salt-containing fraction of the hydroxylamine digest was refractionated by reversed-phase chromatography, and one pure peptide, hydroxylamine 2, was isolated. Its amino acid composition was as follows: Asp, 1.17 (1); Glu, 1.12 (1); Ser, 0.99 (1); Gly, 1.10 (1); His, 0.84 (1); Arg, 2.05 (2); Pro, 2.11 (2); Tyr, 0.22; Val, 0.97 (1); Ile, 0.81 (1); Leu, 1.77 (2); and Phe, 1.02 (1). These 14 amino acids provide the overlap allowing alignment of NT 12 and NT 2, in that order.

Peptide L5 (Phe, Arg, Asx) could arise from two different combinations of tryptic peptides—NT 8 or NT 12 providing its amino-terminal Asx and NT 6 providing its carboxyl-terminal Phe-Arg sequence. The actual combination NT 6-12 (Figure 4d) was deduced on the basis of homology considerations (vide infra). The other three thermolysin peptides—L1, L2, and L3—correspond to sequences identified within peptides NT 12, NT 9″, and NT 7 (Figure 4).

Finally, the assignment of peptide NT 2 as the carboxyl-terminal peptide is in agreement with the fact that proline is the carboxyl-terminal amino acid residue of angiogenin.

Discussion

The strategy and tactics for sequencing angiogenin were determined largely by the scarcity of material. Since a conventional sequence study was not feasible, it was necessary to choose a wider ranging approach. As has been discussed in detail in the preceding paper (Fett et al., 1985b), the isolation of angiogenin was made possible by the development of conditions for growing HT-29 cells and maintaining them in the absence of exogenous serum proteins, by HPLC methods for the purification of extremely small amounts of protein, and by statistical evaluation of biological assays. Even under optimal conditions, the yield of angiogenin was such that it took up to several months to acquire enough material for a single trypsin digestion experiment. To prevent unnecessary losses, all glassware had to be siliconized, and special care was required throughout.

Peptide separation at the nanomole and subnanomole level was greatly facilitated by reversed-phase HPLC using volatile solvents. However, in some instances it was necessary to employ perchlorate/phosphate buffer which meant that the resulting peptides had to be desalted by rechromatography on C18 columns. Despite these repeated fractionations, we were able to recover sufficient product to carry out amino acid analyses at or below the 100-pmol level and sequencing at or below the 1-nmol level.

Hydrolysis of samples for amino acid analysis presented a serious contamination problem which was eventually overcome by glassware pyrolysis and the "dry" HCl procedure described under Experimental Procedures. Conventional amino acid analysis gives useful results at the 100–200-pmol level and with fluorescence detection at the 50-pmol level. Precolumn derivatization by PITC and analysis by HPLC reduced the limits of detection to the 1–5-pmol level.

Sequencing very small amounts of protein or peptide on the Beckman 890C instrument was another challenge that was addressed by using clean reagents, implementing suggestions for greater efficiency (Frank, 1979), and detecting manually converted PTH-amino acids by high-sensitivity HPLC (Strydom & Bazzone, 1985). Many of the peptides examined were ultimately sequenced by using nanomole amounts of material.

Once it became clear that it would be possible to isolate angiogenin in amounts suitable for microsequencing, two principal objectives of the work were defined: first, to obtain as much information as possible on the protein sequence by direct studies on the protein, and second, to use this information to design a probe for isolation of a cDNA clone from an appropriate cDNA library. The sequence of the cDNA would then confirm the accuracy of the chemically derived sequence which in turn would be needed to identify the presence or absence of any posttranslational modifications.

The peptides isolated from the tryptic digest of the native protein almost completely accounted for the amino acid composition of the whole protein. Minor discrepancies were found for Asp, Ser, Glu, Gly, and Ala which are likely due either to background problems or, in the case of Ala, to slight contamination by lysozyme which is the major component (95%) of the fraction from which angiogenin is isolated (Fett et al., 1985a). All were isolated with good yields and, except for three, were sequenced easily at the nanomole level. Only one was found to be blocked; its amino acid composition and molecular weight by mass spectrometry (on <1 nmol) were consistent with a single residue each of <Glu, Asn, Asp, Ser, and Arg and established <Glu as the amino-terminal residue of the protein. Two other peptides, NT 9 and NT 10, each yielded nearly equimolar quantities of two sequences, and each was shown to consist of two peptides linked by a disulfide bond.

Peptide NT 11 was also found to be a mixture of two disulfide-linked peptides which from the Edman degradation results could readily be identified as NT 4b and NT 12. The fact that these two peptides were present in a trypsin digest of native angiogenin would suggest either that they are not joined by a disulfide bond in the intact protein or that this bond is labile. Whether or not this has any biological relevance must await further investigation.

Peptide NT 7 was especially important for the second objective of this work. It was the first of the longer peptides to be isolated in pure form and to be completely sequenced. As will be described in a separate paper (Kurachi et al., 1985), the sequence data not only allowed this peptide to be synthesized in large amounts, e.g., for immunological studies, but also served as the basis for the design and synthesis of an oligonucleotide probe. This, in turn, was used to isolate and clone the desired cDNA. It was still necessary, and indeed particularly critical, to determine the primary structure of tumor-derived angiogenin by direct amino acid sequencing methods. Eventually all of the tryptic peptides were sequenced (Table II). Enough chemical sequencing data were acquired to describe the structure of angiogenin in four parts. They are the amino-terminal 40 residues, a 20-residue section composed of the sequence NT8-3a-4b, another 35-residue segment from NT 4a through NT 10″, and the carboxyl-terminal 28 residues of NT 6-12-2. The amino and carboxyl

```
                1
Horse RNase     KESPAMKFERQHMDSGSTSSNPTNYCNQMMKRRNMTQGCK
Angiogenin      QDNSRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLTSPCK
Probability                          344566668
```

FIGURE 5: Comparison of the amino-terminal 40 residues of angiogenin with the sequence of horse pancreatic ribonuclease (Scheffer & Beintema, 1974). Program SCOR (T. French and D. J. Strydom, unpublished results) based on the method of MacLachlan (1971) was used by employing the matrix of similarities of Schwartz & Dayhoff (1978) and a span length of 17. Probabilities for finding alignments of the 17-residue segments with that much similarity are plotted under the sequences, centered under the respective spans, as n, where n is the exponent of 10^{-n}.

segments were identified by their unique pyroglutamic acid and proline terminal residues, respectively, while the remaining two parts were then aligned into the final structure on the basis of homology considerations and by comparison with the sequence derived from the cDNA.

The longest continuous sequence to be established was the amino-terminal 40 residues of angiogenin. A comparison of this region with the protein sequences in the National Biomedical Research Foundation protein sequence data bank revealed several sets of possible homologies, but, in particular, that to the family of pancreatic ribonucleases was remarkable. The same homology was identified when the computer search was carried out with NT 12 and also by direct visual comparison of the sequence defined by cDNA (Kurachi et al., 1985) and that of human ribonuclease.

Sequence similarities to proteins other than ribonuclease also became apparent from the protein sequence data bank search using the complete sequence of angiogenin, but none of these were of such a magnitude that they could be ascribed to descent from a common ancestor, i.e., homology. Thus, bovine prochymosin has a 6-residue stretch of identity to residues 103–108 of angiogenin and 9 out of 14 for residues 103–116. Since the rest of these two proteins have little or no correspondence, this apparent similarity is likely due to chance rather than homology. The same is true for yeast triosephosphate isomerase, cAMP-dependent protein kinase (type III regulatory chain), human influenza virus neuraminidase, mouse transforming factor (fos), bovine carboxypeptidase B, mouse epidermal growth factor precursor, and a number of others. A particularly interesting relationship has been noted with DNA-dependent RNA polymerase where residues 1136–1149 align with 7 of the last 14 residues of angiogenin. A detailed discussion of sequence homologies will be the subject of a separate communication.

Figure 5 shows how the amino-terminal sequence of angiogenin aligns with the corresponding region of horse ribonuclease. The probabilities that these similarities are due to chance are in the 10^{-6}–10^{-8} range, and thus, the homology is highly significant. On the basis of this relationship, the tryptic peptides have been aligned into a presumptive final sequence (Figure 6). Not only do peptides NT 1, 7, 9′, 13, and 10′ align as per the pyroglutaminase experiment but also they are followed by NT 8 which is homologous to horse/human ribonuclease residues 42–52. NT 12 and its disulfide-linked partner NT 4b then align with residues 105–126 and 56–61, respectively, corresponding to the known disulfide bond between Cys-58 and Cys-110 of ribonuclease (Smyth et al., 1963). An insertion of a gap corresponding to two residues in peptide NT 12 is needed for this alignment. The disulfide-linked partners of peptides NT 9′ and NT 10′ are likewise aligned with the ribonuclease region 77–98. All four

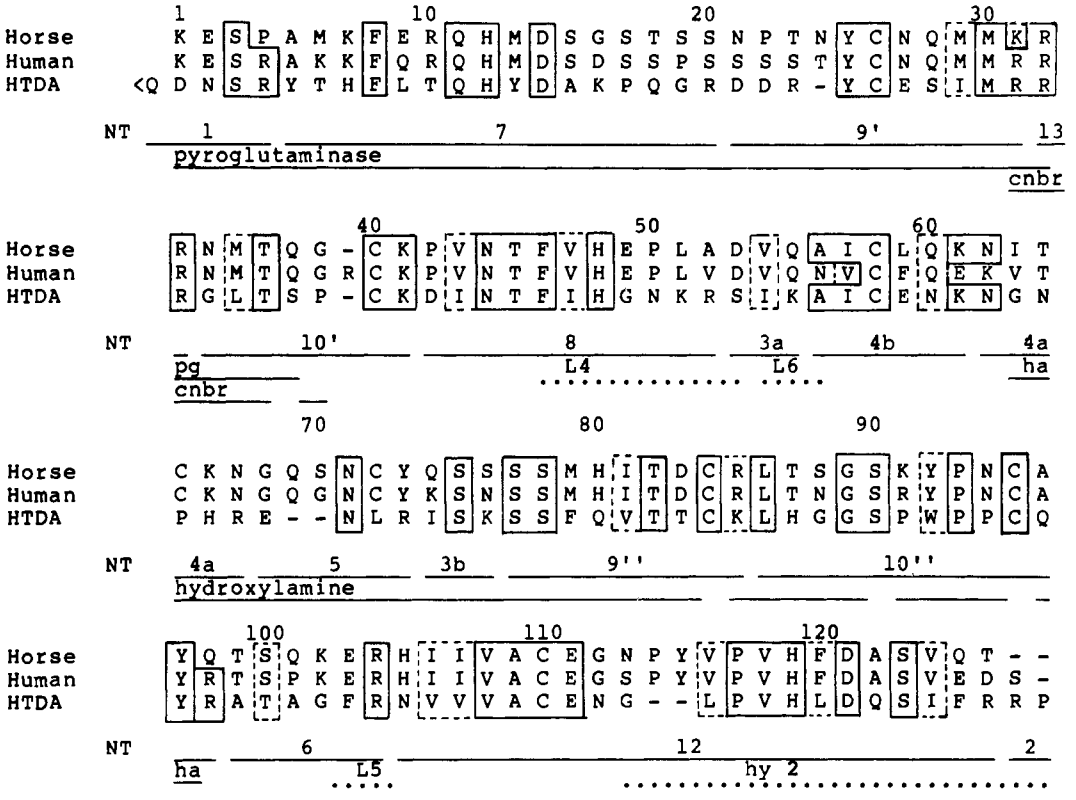

FIGURE 6: Alignment of angiogenin and its peptides with the sequences of horse and human pancreatic ribonucleases (Scheffer & Beintema, 1974; Beintema et al., 1984). Gaps were introduced in four places to maximize homology. The gap at 24 is placed arbitrarily and could be anywhere from 15 through 24. The one-letter amino acid notation is used. The alignment is numbered according to the human enzyme. Abbreviations: <Q, pyroglutamic acid; pg, pyroglutaminase (sequence determined after pyroglutaminase digest); ha, hydroxylamine (sequence of major fraction of the hydroxylamine digest); cnbr, sequence of CNBr digests; hy 2, area represented by the amino acid composition of peptide hydroxylamine 2. The Ln peptides are placed by amino acid composition. Identities between angiogenin and either human or horse pancreatic ribonucleases are boxed, while similarities are enclosed by dashed lines.

of these latter peptides show homology to the pancreatic enzymes, in addition to their being disulfide bonded in identical fashion. Only the disulfide between residues 65 and 72 of ribonuclease is missing in angiogenin. The sequence of the nonblocked partner of the major hydroxylamine cleavage product provides the alignment for tryptic peptides NT 4a-5-3b-9″-10″. The amino acid composition of the thermolysin peptides L4 and L6 indicates that they can only arise from the alignment NT 8-3a-4b, assuming thermolysin exerts its usual specificity (Feder & Schuck, 1970) by cleaving Thr-Phe in NT 8, Ser-Ile in NT 3a, and Ala-Ile in NT 4b. Similarly, peptide L5 (Phe, Arg, Asx) can only be accounted for by the alignment NT 6-12. Although NT 8 starts with Asp-Ile, the ribonuclease homology calls for the NT 10′-8 sequence rather than the NT 6-8 sequence. It would be quite unreasonable to insert a hexapeptide after NT 10′ while leaving a six-residue gap after NT 10″ in the alignment with the ribonuclease.

All 123 amino acids have thus been placed in this predicted sequence. The gene sequence for the protein corresponds precisely to this chemical alignment (Kurachi et al., 1985). Moreover, there is no indication from the sequencer data that any of the side chains of angiogenin other than the amino-terminal pyroglutamic acid have been subjected to post-translational modification.

We not only have noted the high degree (35%) of sequence identity to the ribonucleases but also have noted that the important active-site residues of the pancreatic enzyme, His-12, His-119, and Lys-41, are conserved in angiogenin. This would seem to be very important and could be pertinent to the function of angiogenin. Such homology with maintenance of the active-site residues, i.e., the actual residues involved in catalytic activity, suggests strongly that angiogenin might exhibit a ribonuclease-like activity. Although preliminary studies have not detected any ribonuclease A like activity, whether or not angiogenin is catalytically active and what relationship, if any, this may have to angiogenic activity await further enzymatic and biological studies now in progress. In this regard, it should be noted that sequence analysis of notexin, an Australian tiger snake venom toxin, revealed an unexpected high degree of homology to pancreatic phospholipase A_2 (Halpert & Eaker, 1975). Subsequent enzymatic assays demonstrated that notexin was, indeed, a potent phospholipase (Halpert & Eaker, 1976). However, in two other cases where hydrolytic enzyme homologues with new functions have been found, namely, the serine protease–haptoglobin A homology (Kurosky et al., 1980) and the classic lysozyme–lactalbumin case (Findlay & Brew, 1972; Browne et al., 1969; Klee & Klee, 1970), the "new" function does not include a hydrolytic step, and one or more of the catalytic residues are lost. Conversely, within the heterogeneous group of serine proteases, for example, the catalytically essential active-site residues are always conserved, even though profound differences in the rest of the structures may be found. The only exception is rat submaxillary tonin, a serine protease, in which the active-site Asp-102 of chymotrypsin is replaced by leucine (Lazure et al., 1984). Conservation of the ribonuclease active-site residues in angiogenin could therefore be thought to signal a possible hydrolytic function for angiogenin. The specificity of such an enzymatic action need not be that of ribonuclease A and—by analogy to the serine proteases—might very well be unique.

As indicated above, angiogenin does not exhibit ribonuclease A like activity under standard assay conditions. Angiogenin also lacks the Cys-65–Cys-72 disulfide bond of ribonuclease A. However, this bond can be selectively reduced and mercurated with no effect on the activity of ribonuclease (Sperling et al., 1969) and is therefore not critical for activity.

The identification and chemical characterization of molecular entities involved in organogenesis in general, and angiogenesis in particular, are a basic quest of biology that is now becoming amenable to study. As a rule, such biological messenger molecules are extremely potent; hence, they are usually present in biological material in miniscule quantities that are often insufficient for even the most advanced experimental approaches. Pathological conditions have proved advantageous at times by enhancing their production, thereby improving yields. This appears to be the case for angiogenesis, the process of developing a hemovascular network. For over 100 years, solid tumors have been known to elicit the proliferation of blood vessels and have indeed shown themselves to be exceptionally good sources of angiogenesis factors [Folkman & Cotran, 1976; see Vallee et al. (1985) and references cited therein].

The complete chemical characterization of a unique human organogenic messenger molecule described here constitutes the first major goal of this long-term investigation of organogenesis in general and angiogenesis in particular. The homology to ribonuclease was entirely unexpected and in itself requires study. It opens up novel directions for the investigation of the biological process of angiogenesis.

ACKNOWLEDGMENTS

We are indebted to Drs. K. Biemann and S. Martin for the analysis of peptide NT 1 by FAB mass spectrometry. The expert technical assistance of Nancy Hanley, Nazik Sarkissian, and Ramdas Pai at various stages of the work is gratefully acknowledged.

REFERENCES

Beintema, J. J., Wietzes, P., Weickmann, J. L., & Glitz, D. G. (1984) *Anal. Biochem. 136*, 48–64.

Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. L. (1984) *J. Chromatogr. 336*, 93–104.

Bornstein, P. (1970) *Biochemistry 9*, 2408–2421.

Browne, W. J., North, A. C. T., Phillips, D. C., Brew, K., Vanaman, T. C., & Hill, R. L. (1969) *J. Mol. Biol. 42*, 65–86.

Dreyer, W., & Bynum, E. (1967) *Methods Enzymol. 11*, 32–39.

Feder, J., & Schuck, J. M. (1970) *Biochemistry 9*, 2784–2791.

Fett, J. W., Strydom, D. J., Lobb, R. R., Alderman, E. M., & Vallee, B. L. (1985a) *Biochemistry 24*, 965–975.

Fett, J. W., Strydom, D. J., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985b) *Biochemistry* (first of three papers in this issue).

Findlay, J. B. C., & Brew, K. (1972) *Eur. J. Biochem. 27*, 65–86.

Folkman, J., & Cotran, R. S. (1976) *Int. Rev. Exp. Pathol. 16*, 207–248.

Frank, G. (1979) *Hoppe-Seyler's Z. Physiol. Chem. 360*, 997–999.

Halpert, J., & Eaker, D. (1975) *J. Biol. Chem. 250*, 6990–6997.

Halpert, J., & Eaker, D. (1976) *J. Biol. Chem. 251*, 7343–7347.

Holmquist, B., & Vallee, B. L. (1974) *J. Biol. Chem. 249*, 4601–4607.

Klapper, D. G., Wilde, C. E., & Capra, J. D. (1978) *Anal. Biochem. 85*, 126–131.

Klee, W. A., & Klee, C. B. (1970) *Biochem. Biophys. Res. Commun. 39*, 833–841.

Kurachi, K., Davie, E. W., Strydom, D. J., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry* (third of three papers in this issue).
Kurosky, A., Barnett, D. R., Lee, T. H., Touchstone, B., Hay, R. E., Arnott, M. S., Bowman, B. H., & Fitch, W. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3388-3392.
Lazure, C., Leduc, R., Seidah, N. G., Thibault, G., Genest, J., & Chretien, M. (1984) *Nature (London) 307*, 555-558.
McLachlan, A. D. (1971) *J. Mol. Biol. 61*, 409-424.
Meek, J. L. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1632-1636.
Podell, D. N., & Abraham, G. N. (1978) *Biochem. Biophys. Res. Commun. 81*, 176-185.
Ruegg, U. T., & Rudinger, J. (1977) *Methods Enzymol. 47*, 111-116.
Scheffer, A. J., & Beintema, J. J. (1974) *Eur. J. Biochem. 46*, 221-233.
Schwarz, R. M., & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, pp 353-358, National Biomedical Research Foundation, Washington, DC.
Smyth, D. G., Stein, W. H., & Moore, S. (1963) *J. Biol. Chem. 238*, 227-234.
Sperling, R., Burstein, Y., & Steinberg, I. Z. (1969) *Biochemistry 8*, 3810-3820.
Strydom, D. J., & Bazzone, T. J. (1985) *J. Protein Chem.* (in press).
Tarr, G. E., Beecher, J. F., Bell, M., & McKean, D. J. (1978) *Anal. Biochem. 84*, 622-627.
Titani, K., Sasagawa, T., Resing, K., & Walsh, K. A. (1982) *Anal. Biochem. 123*, 408-412.
Vallee, B. L., Riordan, J. F., Lobb, R. R., Higachi, N., Fett, J. W., Crossley, G., Bühler, R., Budzik, G., Breddam, K., Bethune, J. L., & Alderman, E. M. (1985) *Experientia 41*, 1-15.
Zalut, C., Henzel, W. J., & Harris, H. W., Jr. (1980) *J. Biochem. Biophys. Methods 3*, 11-30.

Sequence of the cDNA and Gene for Angiogenin, a Human Angiogenesis Factor†

Kotoku Kurachi,‡,§ Earl W. Davie,‡ Daniel J. Strydom,§ James F. Riordan,§ and Bert L. Vallee*,§

Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School, Boston, Massachusetts 02115, and Department of Biochemistry, University of Washington, Seattle, Washington 98195

Received June 14, 1985

ABSTRACT: Human cDNAs coding for angiogenin, a human tumor derived angiogenesis factor, were isolated from a cDNA library prepared from human liver poly(A) mRNA employing a synthetic oligonucleotide as a hybridization probe. The largest cDNA insert (697 base pairs) contained a short 5′-noncoding sequence followed by a sequence coding for a signal peptide of 24 (or 22) amino acids, 369 nucleotides coding for the mature protein of 123 amino acids, a stop codon, a 3′-noncoding sequence of 175 nucleotides, and a poly(A) tail. The gene coding for human angiogenin was then isolated from a genomic λ Charon 4A bacteriophage library employing the cDNA as a probe. The nucleotide sequence of the gene and the adjacent 5′- and 3′-flanking regions (4688 base pairs) was then determined. The coding and 3′-noncoding regions of the gene for human angiogenin were found to be free of introns, and the DNA sequence for the gene agreed well with that of the cDNA. The gene contained a potential TATA box in the 5′ end in addition to two Alu repetitive sequences immediately flanking the 5′ and 3′ ends of the gene. The third Alu sequence was also found about 500 nucleotides downstream from the Alu sequence at the 3′ end of the gene. The amino acid sequence of human angiogenin as predicted from the gene sequence was in complete agreement with that determined by amino acid sequence analysis. It is about 35% homologous with human pancreatic ribonuclease, and the amino acid residues that are essential for the activity of ribonuclease are also conserved in angiogenin. This provocative finding is thought to have important physiological implications.

Angiogenesis, the process leading to the development of a vascular network in normal as well as in malignant tissues, occurs under various physiological and pathological conditions including wound healing, embryonic development, rheumatoid diseases, diabetic retinopathies, and progressive tumor growth (Folkman & Cotran, 1976; Gullino, 1981; Schor & Schor, 1983; Vallee et al., 1985). The proteins that mediate angiogenesis have been difficult to characterize since they occur in trace amounts. Recently, the first human tumor derived angiogenesis factor, designated as angiogenin, has been isolated and characterized (Fett et al., 1985). It is a protein of molecular weight ~14000 and induces neovascularization on the chick chorioallantoic membrane at femtomolar levels and in the rabbit cornea at picomolar concentrations. It is purified from the serum-free conditioned medium of a human colon adenocarcinoma cell line (HT-29) (Alderman et al., 1985). The amino acid sequence of angiogenin isolated from this cell line is described in an accompanying paper (Strydom et al., 1985).

In the present report, the cloning and sequence of cDNAs and the gene for human angiogenin are described. The amino acid sequence predicted from the gene agrees completely with that obtained by protein sequencing studies. There is a high

† This work was supported in part by funds from the Monsanto Co. under agreements with Harvard University.
‡ University of Washington.
§ Harvard Medical School.

0006-2960/85/0424-5494$01.50/0