Waite, J. H. (1984) *Ann. Chem. 56*, 1935-1939.
Waite, J. H., & Tanzer, M. L. (1981) *Anal. Biochem. 111*, 131-136.
Waite, J. H., & Benedict, C. V. (1984) *Methods Enzymol. 107*, 391-413.
Waite, J. H., & Rice-Ficht, A. C. (1987) *Biochemistry 26*, 7819-7825.
Wharton, D. A. (1983) *Parasitology 86*, 85-97.
Zurita, M., Bieber, D., Ringold, G., & Mansour, T. E. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2340-2344.

Amino Acid Sequence of Bovine Angiogenin†

Michael D. Bond and Daniel J. Strydom*

Center for Biochemical and Biophysical Sciences and Medicine and Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115

Received December 29, 1988; Revised Manuscript Received March 27, 1989

ABSTRACT: The amino acid sequence and disulfide bridges of bovine plasma derived angiogenin were determined by sequencer analysis of the intact protein and fragments derived by enzymatic and chemical digestion. Bovine angiogenin is a single-chain protein of 125 amino acids; it contains six cysteines and has a calculated molecular weight of 14 595. In contrast to the human protein its amino terminus is unblocked. It has the following sequence: H_2N-Ala[1]-Gln-Asp-Asp-Tyr-Arg-Tyr-Ile-His-Phe[10]-Leu-Thr-Gln-His-Tyr-Asp-Ala-Lys-Pro-Lys[20]-Gly-Arg-Asn-Asp-Glu-Tyr-Cys-Phe-Asn-Met[30]-Met-Lys-Asn-Arg-Arg-Leu-Thr-Arg-Pro-Cys[40]-Lys-Asp-Arg-Asn-Thr-Phe-Ile-His-Gly-Asn[50]-Lys-Asn-Asp-Ile-Lys-Ala-Ile-Cys-Glu-Asp[60]-Arg-Asn-Gly-Gln-Pro-Tyr-Arg-Gly-Asp-Leu[70]-Arg-Ile-Ser-Lys-Ser-Glu-Phe-Gln-Ile-Thr[80]-Ile-Cys-Lys-His-Lys-Gly-Gly-Ser-Ser-Arg[90]-Pro-Pro-Cys-Arg-Tyr-Gly-Ala-Thr-Glu-Asp[100]-Ser-Arg-Val-Ile-Val-Val-Gly-Cys-Glu-Asn[110]-Gly-Leu-Pro-Val-His-Phe-Asp-Glu-Ser-Phe[120]-Ile-Thr-Pro-Arg-His-OH. Disulfide bonds link Cys(27)-Cys(82), Cys(40)-Cys(93), and Cys(58)-Cys(108). Bovine angiogenin is 64% identical with human angiogenin; like the human protein, it is homologous to the pancreatic ribonucleases, with conservation of active site residues. Two regions, 6-22 and 65-75, are highly conserved between the angiogenins but are significantly different from those of the ribonucleases, suggesting a possible role in the molecules' biological activity.

Angiogenin is a nonmitogenic protein that stimulates angiogenesis. It was first isolated from human tumor cell conditioned media (Fett et al., 1985) and later from normal human serum (Shapiro et al., 1987). Its amino acid sequence is homologous to that of RNase A[1] (Strydom et al., 1985), and it contains many of the important structural and active site residues critical to RNase activity, including the catalytically essential His-12, Lys-41, and His-119 (RNase A numbering). Indeed, angiogenin displays ribonucleolytic activity (Shapiro et al., 1986) and also binds tightly to a protein RNase inhibitor from placenta (Shapiro & Vallee, 1987).

Despite the similarities between the two molecules, only angiogenin is able to induce blood vessel growth. Therefore, this activity derives at least in part from some structural feature(s) unique to angiogenin. One means of identifying such features is to examine the expected sequence variation in angiogenins from different species. Individual residues and/or regions involved in angiogenesis should be conserved and should vary, perhaps dramatically, from the corresponding residues in RNase A.

To investigate this possibility, we have isolated (Bond & Vallee, 1988) and sequenced angiogenin from bovine plasma. The isolation was made possible by the recent development of a convenient assay system (Bond, 1988). This paper describes the first sequence determination of an angiogenin other than the human protein, namely, the bovine one, examines correspondences and differences between the two angiogenins, and points to specific regions of the molecule that might be important for biological activity.

Materials and Methods

Angiogenin was isolated from bovine plasma as described (Bond & Vallee, 1988). The methodology of sequence determination is given in the supplementary material.

Nomenclature

Unless otherwise indicated, the numbering of amino acid residues is based on the sequence of bovine angiogenin. Tryptic peptides are numbered according to their location in the sequence, from T1, the amino-terminal peptide, to T18, the carboxyl-terminal peptide.

Results

Overlaps and Complete Sequence. Figure 1 summarizes the sequence of bovine angiogenin and the major pieces of

† This work was supported by funds from Hoechst, A.G., under an agreement with Harvard University.

* Address correspondence to this author at the Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School, 250 Longwood Ave., Boston, MA 02115.

[1] Abbreviations: RNase(s), ribonuclease(s); RNase A, pancreatic ribonuclease A; RNase S-protein and S-peptide, subtilisin-generated fragments of RNase A containing residues 21-124 and 1-20, respectively; C18, octadecylsilane; HPLC, high-performance liquid chromatography; Tris, tris(hydroxymethyl)aminomethane; TFA, trifluoroacetic acid; DTT, dithiothreitol; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; PTH, phenylthiohydantoin.

0006-2960/89/0428-6110$01.50/0

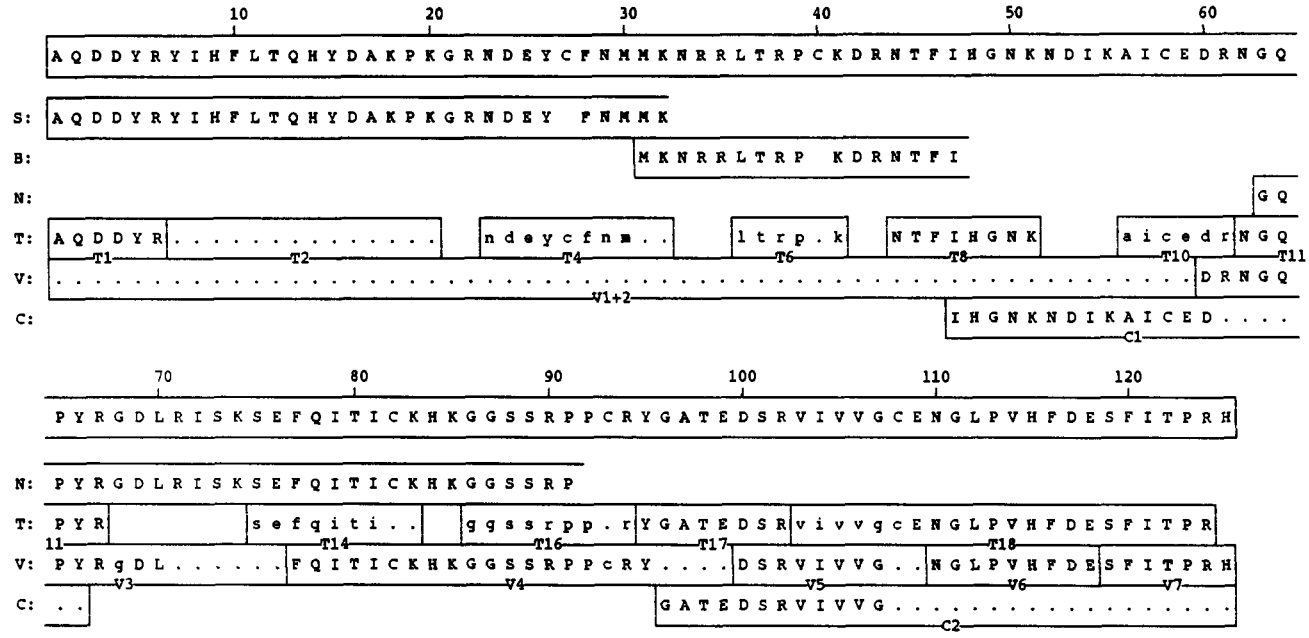

FIGURE 1: Amino acid sequence of bovine plasma derived angiogenin. The extents of sequencing of the amino terminus (S) of the intact protein and of fragments derived from treatment with cyanogen bromide (B), hydroxylamine (N), trypsin (T), *Staphylococcus aureus* V8 protease (V), and chymotrypsin (C) are indicated by capital letters. Tentative sequencer assignments are indicated by lower-case letters, whereas dots represent regions of peptides that were not sequenced but are consistent with the proposed sequence on the basis of amino acid compositions. The lower-case designation of some tryptic peptides indicates regions where disulfide-bonded pairs of peptides were sequenced simultaneously.

evidence for assigning it. The amino-terminal sequence up to residue 32 was obtained by Edman degradation of the intact protein. It was extended to residue 47 by analysis of the cyanogen bromide digest. Tryptic peptide T8 and chymotryptic peptide C1 further extended this to residue 60. Tryptic peptides T10 and T11, overlapped by peptide V3, stretch from residue 56 to residue 67. Edman degradation of the hydroxylamine digest overlapped this sequence starting at residue 63 and extended it to residue 91. The sequence and composition of peptide V4 then continued through to peptide T17, and peptide C2 extended the overlap on into peptide T18, bringing the sequence to residue 124. Finally, the sequence of peptide V7 completed that of angiogenin up to the carboxyl-terminal His-125. Separate hydrazinolysis of bovine angiogenin afforded mainly histidine as the carboxyl-terminal amino acid in 68% yield.

Tryptic peptide pairs T4 and T14, T6 and T16, and T10 and T18 were isolated together and yielded sequence information in equivalent yields for both members of the pair. These results in conjunction with their amino acid compositions established the presence of disulfide bonds linking these specific peptides.

Quality of Sequence Data. Most parts of the sequence were determined at least twice, either by repeated sequencer analyses on one peptide or by the use of different peptides. In the proposed structure the Lys^{32}-Asn-Arg-Arg-Leu region was sequenced only once, and in the Ser-Ser-Arg^{90}-Pro-Pro region low yields made identification and quantitation difficult. In the latter case, no sequencing products were inconsistent with the proposed structure, and in both cases the amino acid compositions of the appropriate peptides are in agreement with the sequences shown.

Discussion

Bovine Angiogenin Sequence. Bovine angiogenin is a single-chain polypeptide of 125 residues. No posttranslational modifications were found. Comparison with human angiogenin reveals that the two proteins are 64% identical (compared to 35% with bovine RNase A) and no insertions/deletions are needed for their alignment (Figure 2). The six Cys residues, forming the three disulfide bonds, are conserved. A cladistic classification of the two angiogenins and the ribonucleases employing program PHYMOD (Strydom, 1973) definitively segregrated the angiogenins from all the ribonucleases, further confirming the identification of the bovine protein as angiogenin. In contrast to human angiogenin, the amino terminus is not blocked and extends one residue beyond that of the human. The amino-terminal alanine residue matches the alanine of the human presequence (Kurachi et al., 1985), signifying a possible difference in the proteolytic processing of the human and bovine pre-angiogenins. Among the pancreatic ribonucleases, the only example of such an N-terminal extension is that of rat (Beintema & Gruber, 1967; Beintema et al., 1986) with a three-residue extension.

There is also a one-residue extension of the carboxyl terminus of the bovine protein relative to the human. The extension, a histidine, is unusual in the angiogenin/RNase family of proteins. All pancreatic RNase terminate either at residue 120 (bovine angiogenin numbering) or at residue 124 (Beintema et al., 1986); the genes of rat RNase and human angiogenin have a stop codon at residue 125. The histidine in position 125 therefore signals the first example of a mutation in the stop codon of the angiogenin/ribonuclease gene family. The consequences for activity appear to be negligible (Bond & Vallee, 1988).

Comparison of Bovine and Human Angiogenin Sequences. The relatively large number of changes that have been incorporated in the angiogenin proteins since divergence of the human and bovine species suggest a rate of accepted mutations similar to that of ribonucleases, one of the faster rates seen among protein families. These changes have, however, been irregularly spaced over the linear sequence, leaving a few regions relatively unaltered. In particular, conserved regions occur around the three catalytically essential residues His-14, Lys-41, and His-115. The five residues to either side of His-14 are identical in both proteins, constituting the longest unvaried segment. The two second-longest conserved regions extend eight residues each and occur immediately preceding His-115

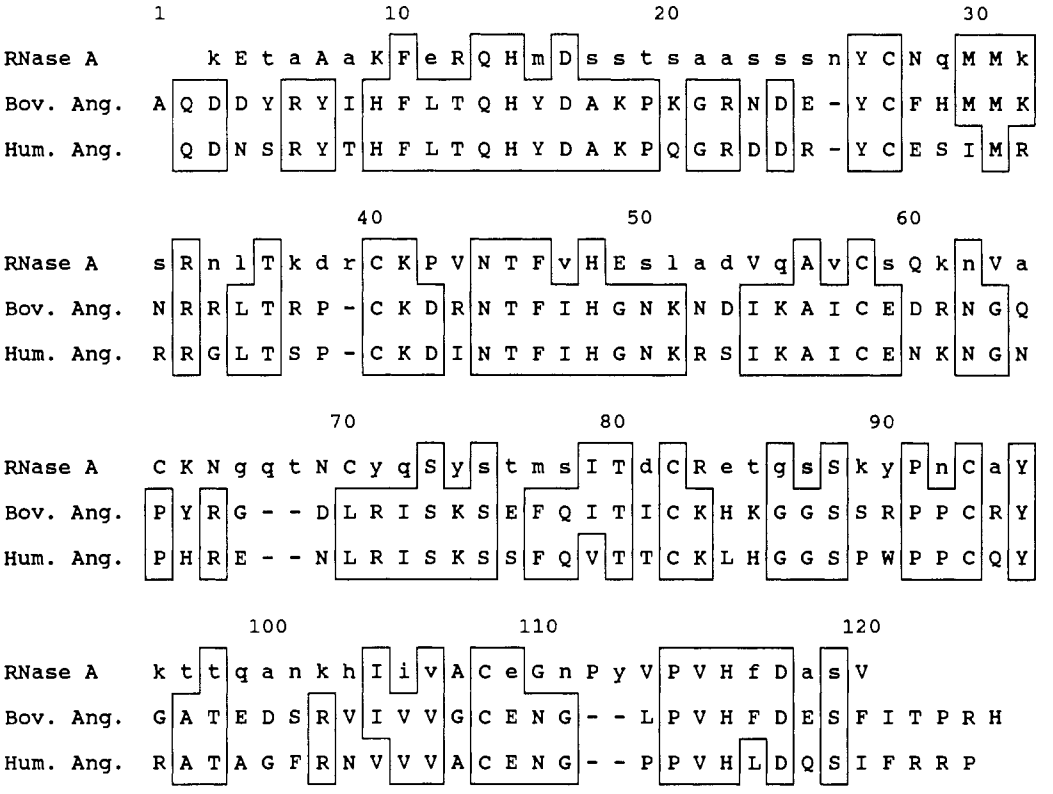

FIGURE 2: Alignment of bovine angiogenin with the sequences of human angiogenin and bovine RNase A (Smyth et al., 1963). The alignment is numbered according to bovine angiogenin. The amino-terminal Q of human angiogenin is pyroglutamic acid. Identities between bovine angiogenin and either human angiogenin or RNase A are boxed. Capital letters in the RNase A sequence indicate residues that are conserved in at least 38 of 39 mammalian pancreatic RNases (Beintema, 1986).

and beginning with the third residue after Lys-41. Three residues surrounding Lys-41 are also conserved. Thus, many of the residues likely to be involved in substrate binding or in positioning of catalytically essential residues are conserved in the angiogenins. This is consistent with the hypothesis that ribonucleolytic activity is important to the mode of action of the molecule.

Some potentially important differences between the two angiogenins are noticeable. Most informative, perhaps, is the replacement of Leu-116 in the human protein by Phe in the bovine. The corresponding residue in RNases, Phe-120, has been thought to be important in maintaining the tertiary structure of the protein and in ensuring the proper alignment of the catalytic His residue preceding it (Lin et al., 1972; Puett, 1972). On this basis it seemed plausible that the occurrence of Leu in this position in human angiogenin might be largely responsible for the low conventional RNase activity of the molecule. This now appears unlikely. The bovine protein contains Phe in this position, and even so, no significant activity toward RNase A type substrates has been detected (Bond & Vallee, 1988); its activity toward tRNA and CpA is equal to or less than that of human angiogenin (R. Shapiro, personal communication). Other differences occur at Gly-107 (Ala in all pancreatic ribonucleases and in human angiogenin), at Arg-43 (Val in the RNases and conservatively replaced with Ile in human angiogenin), at Arg-90 (Trp in human angiogenin and Tyr or Phe in RNases), and at Tyr-5 (Ser or Thr in RNases and Ser in human angiogenin).

Structure–Activity Relationships. The delineation of regions of the angiogenin molecule that are required for specific angiogenic activity, as contrasted with RNase A type activity, was a central aim of this study. One would expect that these areas would be conserved in the angiogenins and would differ from the corresponding sequences in RNase A. Several general regions meet these criteria, but two that stand out occur in the N-terminal region and in the sequence corresponding to the "loop" region of RNases, from Pro-65 to Cys-70.

The N-terminal, or "S-peptide", region is highly conserved in the angiogenins. Fifteen of 17 residues, from Arg-6 to Arg-22, are unvaried, and all but four differ from those present in the pancreatic RNases. Several amino acids which occur in this region in all mammalian RNases (Beintema et al., 1986) are replaced by quite different residues in the angiogenins. Thus, Tyr-7 replaces the smaller, hydrophic Ala in the RNases; His-9 replaces a Lys; and Thr-12 replaces the bulky, cationic Arg residue. Perhaps equally as important, five of the last six amino acids in this region are conserved in the angiogenins, as compared with none in the RNases. Whereas the RNase A residues from Ser-15 to Ala-20 (positions 17–22 in bovine angiogenin) can be entirely removed from the molecule without detrimental effect on enzymatic activity (Hofmann et al., 1966; Richards & Wyckoff, 1971), this short segment is highly conserved in the angiogenins.

A similar situation exists for the sequence from Pro-65 to Ser-75. Part of this region forms a short loop in RNase A, linked covalently by a disulfide bond between the two cysteines in positions 65 and 70. Two residues in the loop, the Gln and Asn-69, are thought to be involved in formation of hydrogen bonds to the purine base of the substrate (Richards & Wyckoff, 1971; Blackburn & Moore, 1982; Brunger et al., 1985; Wodak et al., 1977), but neither of these residues nor the two cysteines are found in the angiogenins. Instead, there are conserved Pro, Arg, and Leu residues in positions 65, 67, and 70, respectively, and a double deletion, relative to RNase A, between residues 68 and 69. Thus, the local three-dimensional structure is likely to vary significantly from that in the RNases. The latter half of the region in angiogenin also

contains a highly conserved segment that is variable in RNase A. Six consecutive residues are unvaried, from Leu-70 to Ser-75, whereas only two are unchanged in the RNases.

In addition to the above regions, other conserved sections in angiogenin differ from those in the RNases, but less dramatically so. The region from Cys-108 to His-115, for example, is conserved completely in the angiogenins. However, five of eight residues are identical with those in RNase A, and one is replaced conservatively (Leu-112). Likewise, in the region surrounding Lys-41, from Pro-39 to Lys-51, six of the residues that are conserved in the angiogenins are the same as those in RNase A. Interestingly, both of these regions and the two above contain substitutions involving a Pro residue conserved in either the ribonucleases or the angiogenins. Such changes are potentially significant due to the unique effects of this amino acid, with fixed torsional angle (ϕ), on local protein structure.

In summary, comparison of similarities and differences between the sequences of the angiogenins and RNase A has pointed out potentially significant changes in the primary, secondary, and, perhaps, tertiary structure of angiogenin as contrasted to RNase A. The pattern of conserved regions, residue deletions, and proline substitutions, as well as the differences in the location and concentration of charged residues, argues most strongly for such alterations in structure. Two regions of potential importance for angiogenic activity, as contrasted with ribonuclease activity, were delineated. These and other regions have been targeted for further investigations of angiogenin structure–function relationships.

Added in Proof

Since submission of the manuscript Maes et al. (1988) have reported the amino acid sequence of an angiogenin isolated from bovine milk; the sequence is identical with that presented here.

Acknowledgments

We thank Dr. B. L. Vallee for his advice and support and gratefully acknowledge the expert technical assistance of Rebecca Ettling, Niem Nguyen, and Nazik Sarkissian.

Supplementary Material Available

Materials and Methods section, results of sequence studies on fragments of angiogenin, and eight tables and one figure containing amino acid composition and sequence data (17 pages). Ordering information is given on any current masthead page.

Registry No. Angiogenic factor, 121055-26-3; bovine angiogenin, 121055-25-2; ribonuclease, 9001-99-4.

References

Beintema, J. J., & Gruber, M. (1967) *Biochim. Biophys. Acta 147*, 612–614.

Beintema, J. J., Fitch, W. M., & Carsana, A. (1986) *Mol. Biol. Evol. 3*, 262–275.

Blackburn, P., & Moore, S. (1982) *Enzymes (3rd Ed.) 15*, 317–433.

Bond, M. D. (1988) *Anal. Biochem. 173*, 166–173.

Bond, M. D., & Vallee, B. L. (1988) *Biochemistry 27*, 6282–6287.

Brünger, A. T., Brooks, C. L., III, & Karplus, M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 8458–8462.

Fett, J. W., Strydom, D. J., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry 24*, 5480–5486.

Hofmann, K., Finn, F. M., Limetti, M., Montibeller, J., & Zanetti, G. (1966) *J. Am. Chem. Soc. 88*, 3633–3639.

Kurachi, K., Davie, E. W., Strydom, D. J., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry 24*, 5494–5499.

Lin, M. C., Gutte, B., Caldi, D. G., Moore, S., & Merrifield, R. B. (1972) *J. Biol. Chem. 247*, 4768–4774.

Maes et al. (1988) *FEBS Lett. 241*, 41–45.

Puett, D. (1972) *Biochemistry 11*, 4304–4307.

Richards, F. M., & Wyckoff, H. W. (1971) *Enzymes (3rd Ed.) 4*, 647–806.

Shapiro, R., & Vallee, B. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2238–2241.

Shapiro, R., Riordan, J. F., & Vallee, B. L. (1986) *Biochemistry 25*, 3527–3532.

Shapiro, R., Strydom, D. J., Olson, K. A., & Vallee, B. L. (1987) *Biochemistry 26*, 5141–5146.

Smyth, D. G., Stein, W. H., & Moore, S. (1963) *J. Biol. Chem. 238*, 227–234.

Strydom, D. J. (1973) *Comp. Biochem. Physiol. 44B*, 269–281.

Strydom, D. J., Fett, J. W., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry 24*, 5486–5494.

Wodak, S. Y., Liu, M. Y., & Wyckoff, H. W. (1977) *J. Mol. Biol. 116*, 855–875.