that cannot be answered with less-ordered cocrystals. It should be possible to determine: (i) the precise contacts made by the repressor, including contacts made by the flexible NH_2-terminal arm and contacts that might be made via bridging water molecules; (ii) whether the conformation of the protein changes as it binds to the operator; (iii) whether the DNA bends or twists significantly; and (iv) how the local structure of the DNA differs from that of a uniform B-DNA fiber (*23, 24*). Because the protein structure was determined previously and a detailed model for the complex was proposed (*8*), this study should also help evaluate the reliability of model-building and the importance of cocrystals for future studies of protein-DNA interactions.

Although it is tedious, systematic screening of crystallization conditions with helices of different lengths or with different ends may be a useful approach for cocrystallization of other proteins. Our results suggest that crystallization conditions and crystal quality can vary drastically as the length of the duplex and the nature of the ends are changed. Thus the 20-mer gave excellent cocrystals when precipitated with 20 percent PEG 400, but the best conditions with the 21-mer (100 m*M* $MgCl_2$ and 9 percent PEG 400) only gave microcrystals. Cocrystals containing the 22-mer grew quite readily from 15 m*M* $Co(NH_3)_6Cl_3$, but these were poorly ordered.

Is it possible to predict which length of DNA will be optimal for cocrystallization? Anderson, Ptashne, and Harrison predicted that a 21-mer would be particularly useful for cocrystallization (*14*). They assumed (i) that duplexes will stack to form a pseudocontinuous helix, (ii) that DNA fragments, like DNA fragments in solution, will have 10.5 base pairs per helical turn (*25*), and (iii) that a small repeating unit is desirable. Presumably our use of self-complementary ends increases the chances of forming a pseudocontinuous helix in the crystal. If this occurs, there will be exactly two helical turns as the DNA traverses our unit cell, but a 20-mer, rather than a 21-mer, yields the best crystals. Our experience shows that it may be necessary to try several DNA fragments with any particular protein.

References and Notes

1. W. F. Anderson, D. H. Ohlendorf, Y. Takeda, B. W. Matthews, *Nature (London)* **290**, 754 (1981).
2. C. O. Pabo and M. Lewis, *ibid.* **298**, 443 (1982).
3. D. B. McKay and T. A. Steitz, *ibid.* **290**, 744 (1981).
4. R. W. Schevitz, Z. Otwinowski, A. Jouchimiak, C. L. Lawson, P. B. Sigler, *ibid.* **317**, 782 (1985).
5. T. A. Steitz, D. H. Ohlendorf, D. B. McKay, W. F. Anderson, B. W. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3097 (1982).
6. R. T. Sauer *et. al.*, *Nature (London)* **298**, 447 (1982).
7. D. H. Ohlendorf, W. F. Anderson, R. G. Fisher, Y. Takeda, B. W. Matthews, *ibid.*, p. 718.
8. M. Lewis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 435 (1983).
9. I. T. Weber and T. A. Steitz, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3973 (1984).
10. C. O. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984).
11. H. C. M. Nelson, M. H. Hecht, R. T. Sauer, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 441 (1983).
12. M. H. Hecht, H. C. M. Nelson, R. T. Sauer, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2676 (1983).
13. J. E. Anderson, M. Ptashne, S. C. Harrison, *Nature (London)* **316**, 596 (1985).
14. J. Anderson, M. Ptashne, S. C. Harrison, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1307 (1984).
15. C. A. Frederick *et al.*, *Nature (London)* **309**, 327 (1984).
16. C. O. Pabo, R. T. Sauer, J. M. Sturtevant, M. Ptashne, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1608 (1979).
17. C. O. Pabo, thesis, Harvard University (1980).
18. A. J-H. Wang *et al.*, *Nature (London)* **282**, 680 (1979).
19. A. D. Johnson, thesis, Harvard University (1980).
20. A. McPherson, *Preparation and Analysis of Protein Crystals* (Wiley, New York, 1982).
21. M. Fried and D. Crothers, *Nucleic Acids Res.* **9**, 6505 (1981); T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, Cold Spring Harbor, N.Y., 1982).
22. B. W. Matthews, *J. Mol. Biol.* **33**, 491 (1968).
23. S. Arnott and D. W. L. Hukins, *Biochem. Biophys. Res. Commun.* **47**, 1504 (1972).
24. R. E. Dickerson and H. R. Drew, *J. Mol. Biol.* **149**, 761 (1981).
25. J. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 200 (1979).
26. The 17-base pair duplex was synthesized by Satoshi Ikuta and Keiichi Itakura (City of Hope Medical Center, Duarte, Calif.).
27. S. L. Beaucage, and M. H. Caruthers, *Tetrahedron Lett.* **22**, 1859 (1981).

28. Supported by NIH grant GM 31471, a Junior Faculty Research Award from the American Cancer Society, and an Institutional grant from the American Cancer Society. S.R.J. was supported by a training grant from the National Institutes of Health to the Molecular Biology and Genetics Department. J.M.B. was supported by the Jane Coffin Childs Memorial Fund for Medical Research. We thank Professor Donald M. Crothers of Yale University for helpful suggestions about DNA synthesis and purification. We thank the Molecular Biology Department and the Howard Hughes Genetics Institute for the use of an Applied Biosystems DNA Synthesizer and an HPLC.

17 September 1985; accepted 8 November 1985

Brain-Derived Acidic Fibroblast Growth Factor: Complete Amino Acid Sequence and Homologies

Abstract. *Bovine brain-derived acidic fibroblast growth factor (aFGF) is a protein mitogen originally identified in partially purified preparations of whole brain. The protein was purified to homogeneity and shown to be a potent vascular endothelial cell mitogen in culture and angiogenic substance in vivo. The homology of aFGF to human interleukin-1β was inferred from partial sequence data. The complete amino acid sequence of aFGF has now been determined and observed to be similar to both basic FGF and interleukin-1's. A neuropeptide-like sequence, flanked by basic dipeptides, was observed within the aFGF sequence.*

GUILLERMO GIMENEZ-GALLEGO*
Consejo Superior de Investigaciones Cientificas, Centro de Investigaciones Biologicas, Madrid, Spain 28006
JOHN RODKEY
CARL BENNETT
Department of Medicinal Chemistry, Merck Sharp and Dohme Research Laboratories, West Point, Pennsylvania 19486
MARI RIOS-CANDELORE
JERRY DISALVO
KENNETH THOMAS†
Department of Biochemistry and Molecular Biology, Merck Institute for Therapeutic Research, Merck Sharp and Dohme Research Laboratories, Rahway, New Jersey 07065

*On sabbatical leave in the Department of Biochemistry and Molecular Biology, Merck Institute for Therapeutic Research, Rahway, New Jersey 07065.
†To whom correspondence should be addressed.

Brain has been recognized as a plentiful source of mitogenic activity for primary fibroblasts for almost 50 years (*1*). Two unique fibroblast growth factors (FGF's), one with an acidic (*2*) and the other having a basic (*3*) *p*I, have been purified to homogeneity from brain. Both hypothalamus-derived endothelial cell growth factor (ECGF) and eye-derived growth factor II (EDGF II) appear to be similar, or identical, to the acidic mitogen (*4*). Brain-derived acidic fibroblast growth factor (aFGF) is purified as a pair of microheterogeneous forms, aFGF-1 and aFGF-2, differing in the presence or absence of an amino terminal hexapeptide (*5*). Their masses, as estimated from sodium dodecyl sulfate polyacrylamide gel electrophoresis, are 16,000 to 17,000 daltons and their isoelectric points are from *p*H 5 to *p*H 7 (*2*). We reported earlier the identification, purification, amino acid composition, and partial amino acid sequence of bovine aFGF and its apparent homology to human interleukin-1β (IL-1β) (*2, 5, 6*). The acidic and basic FGF's have about the same mass, bind avidly to heparin affinity columns (*3, 7*), and are active on a wide variety of types of cells (*5, 8, 9*), including vascular endothelial cells. The acidic, and probably the basic, FGF's are potent angiogenic substances in vivo

(*5*, *9*). The physiological significance of two unique FGF's with similar spectra of mitogenic activities is unknown.

The complete amino acid sequence of bovine brain-derived aFGF and the microsequencing techniques used for its determination are given in Fig. 1. The 140-residue sequence is well determined, with most regions confirmed by overlapping peptides. The amino acid composition calculated from the complete sequence agrees well with the composition determined by hydrolysis when rescaled to the slightly lower mass of 15,883 daltons. No Asn glycosylation sequences (Asn-X-Ser/Thr) are present; this is in agreement with (i) the lack of amino sugars following acid hydrolysis, and (ii) the failure to bind either to concanavalin A or to wheat germ lectin affinity columns (*2*).

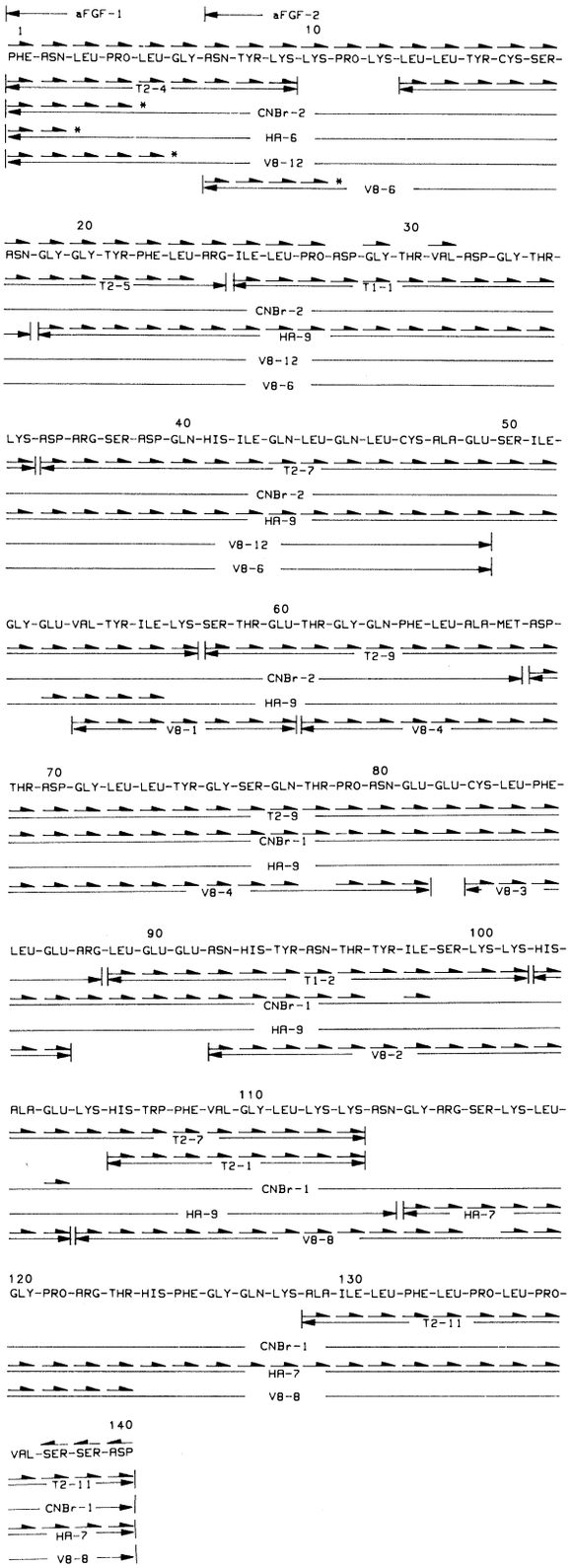
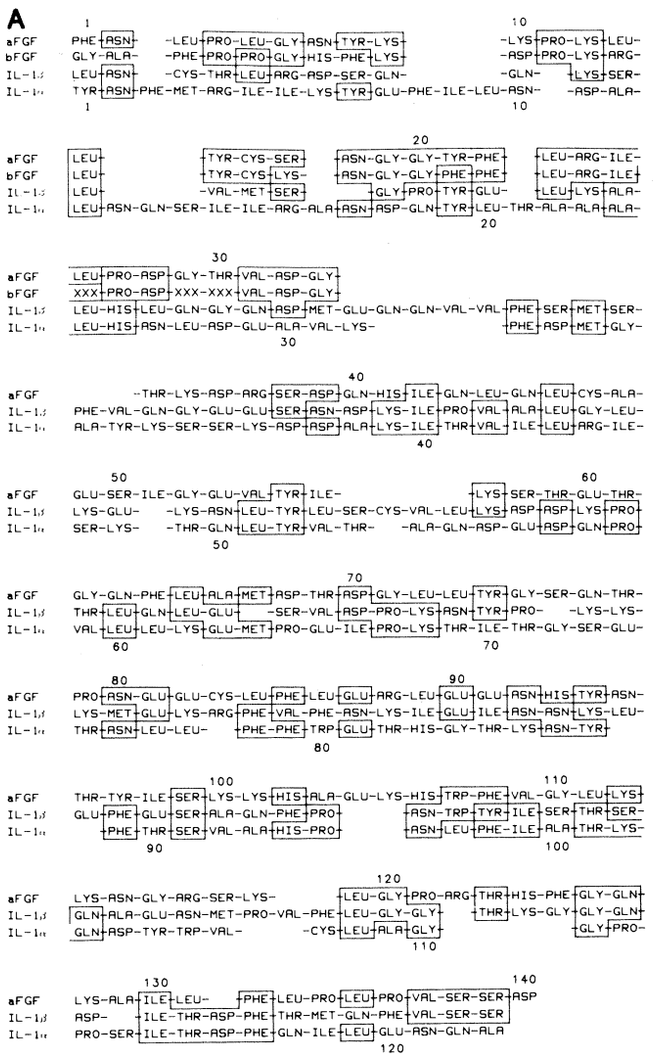


Fig. 1. The complete amino acid sequence of aFGF. Two amino termini, corresponding to two previously identified microheterogeneous forms (*2*), are denoted aFGF-1 and aFGF-2 (*5*). T, CN, HA, and V8 numbers refer to tryptic, cyanogen bromide, hydroxylamine, and *Staphylococcus aureus* V8 protease-generated peptides. All residues identified by amino terminal sequence determinations are denoted by single-headed arrows pointing to the right. Peptide sequences that were prematurely terminated because they were recognized to begin at one of the two previously determined amino termini are marked with asterisks at the end of the last degradation cycle. Residues identified at the carboxyl terminus by timed digestion of the whole protein with carboxypeptidase A are denoted with single-headed arrows pointing to the left. Tryptic peptides denoted by T1 and T2 were obtained after 21- and 6-hour digestions. All peptides were fractionated on 330-Å pore size C_4 and C_{18} Vydac HPLC columns (The Separations Group), with gradients from 10 m*M* trifluoroacetic acid to 33 percent 10 m*M* trifluoroacetic acid and 67 percent acetonitrile. Sequencing was done on a microsequencer (Applied Biosystems 470A).

Brain-derived basic FGF (bFGF) was purified by a modification of the aFGF purification protocol using heparin Sepharose and reversed-phase high-performance liquid chromatography (HPLC) (*10*). The reduced and carboxymethylated protein was sequenced from the amino terminus. Of the amino terminal 42 residues, 39 were unambiguously identified. Bovine brain, human brain, and bovine pituitary bFGF's are clearly equivalent proteins, as judged by comparison of this extended sequence with the amino-terminal 15 residues of human brain and bovine pituitary bFGF's (*3*). Brain bFGF appears to have an amino terminal 9–residue extension not present on aFGF when aligned as shown in Fig. 2A. Beginning at residue 10 of bFGF, 20 of 30 identified amino acid residues are identical with those of aFGF. The significance of the homology is greater than 8 standard deviations from random resemblance (*11*). This shows that the two FGF's descended from a common ancestral protein.

The originally proposed homology (*5*) between the partial sequence of aFGF and equivalent regions in the carboxyl terminal 149 residues of IL-1β precursor (*12*) is now extended to include the entire aFGF molecule and the carboxyl terminal halves of the precursors of both IL-1β and IL-1α (Fig. 2B). The sequences were aligned by a combination of the Needleman and Wunsch algorithm (*11*), hydropathicity (*13*, *14*), residue exposure calculations (*15*), and visual inspection. A family of slightly different alignments generated by small variations in the parameters of the Needleman and Wunsch analysis all gave very similar statistical significance. The best alignments found between aFGF and both IL-1β and IL-1α were 2.2 and 1.8 standard deviations from random, respectively, compared to 5.5 standard deviations between the two IL-1's. The probabilities are less than 3 percent and 7 percent, respectively, that the values 2.2 and 1.8 reflect zero correlation of aFGF with the IL-1β and IL-1α sequences. In the three-way alignment shown in Fig. 2A, 25 percent of the aFGF residues are identical to those of IL-1β, and 19 percent are identical to those of IL-1α; 26 percent of the residues common to the two IL-1's are the same.

Evaluation of the similarities between the patterns of side-chain hydropathicities, a method of detecting common tertiary structural patterns of solvent-ex-



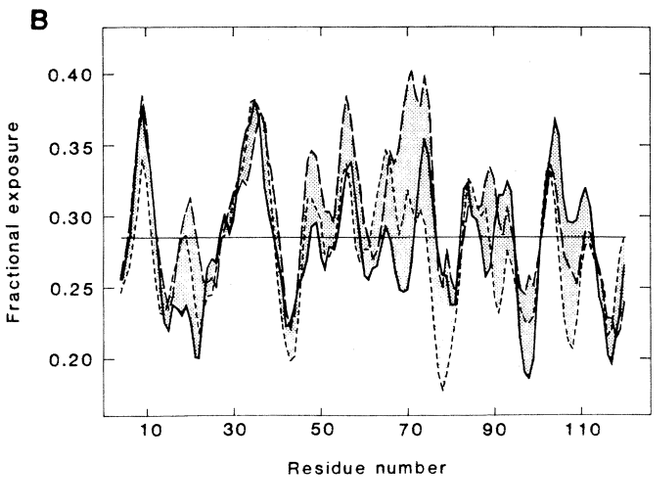


Fig. 2. Homology among FGF's and IL-1's. (A) The complete sequence of bovine aFGF is aligned with the amino terminal end of bovine bFGF and the carboxyl terminal halves of the precursors of both human IL-1β and IL-1α. The bFGF sequence is shown beginning at residue 10 (the first nine residues are Pro-Ala-Leu-Pro-Glu-Asp-Gly-Gly-Ser); unidentified residues are denoted as XXX. Common aligned residues are enclosed in boxes. The numbers above the aFGF sequence denote the aFGF residues, whereas those below the IL-1α sequence refer to the numbering of residues that are common to aFGF and the two IL-1's. (B) The solvent exposure index ["1-f" of Rose *et al.* (*15*)] at each residue, taken as a running average with the preceding and following three residues, is plotted for the common aligned residues of aFGF (—), IL-1β (— —), and IL-1α (- - - -). Maxima and minima correspond to predicted exposed and buried regions. The extent of divergence among the predicted solvent exposures along the three sequences is emphasized by the stippling. The horizontal line is the overall average of all three common sequences.

posed and buried residues in distantly related proteins (*13, 14*), also indicates that aFGF and the IL-1's are related. The similarities between aFGF and the two IL-1's were also evaluated, as shown in Fig. 2B, by a measure of average solvent exposure determined from protein crystallographic coordinates (*15*). Although both hydropathic and exposure diagrams reveal substantial similarities among these proteins, the measure of solvent exposure shows a somewhat greater agreement among the three patterns. When the method and coefficients of Sweet and Eisenberg (*14*) are used to estimate the significance of this similarity, the pairwise correlations among the three proteins are each over six standard deviations from random.

Post-translationally processed mature aFGF, bFGF, IL-1α, and IL-1β are similar by both mass and sequence homology. In addition, IL-1β (*16*) and both FGF's are mitogenic for primary fibroblasts. As aligned, the amino termini of aFGF and bFGF begin five residues after and four residues before that of the processed form of IL-1β (*16*). These data support the proposal that both FGF's and IL-1's are members of a new protein family of homologous growth factors. The amino terminal sequence of an interferon releasing factor (*17*) places it, also, within the same family.

A decapeptide (residues 102 to 111) is contained within aFGF that is flanked on each end by Lys-Lys dipeptides. Since basic dipeptide sequences are signals for proteolytic processing of many active peptides from their longer precursor proteins, we compared the decapeptide sequence to sequences of other known biologically active polypeptides. The brain-derived aFGF decapeptide aligns well with various biologically active neuropeptides, especially in the carboxyl terminal half of the sequence (Fig. 3), a highly conserved region believed to participate in receptor binding and much of the biological activity of the neuropeptides (*18*). Furthermore, at least some of these peptides are mitogenic. The mammalian neuropeptides substance P (from both central and peripheral nervous system) and substance K (from spinal cord) stimulate DNA synthesis in both cultured arterial smooth muscle cells and

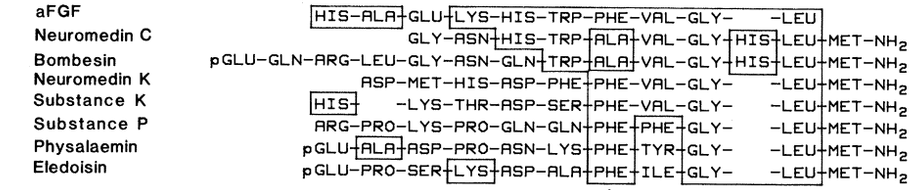


Fig. 3. The sequence similarities of the aFGF neuropeptide-like amino acid sequence with known neuropeptides. The sequence of residues 102 to 111 of bovine aFGF is aligned with neuropeptide sequences. Amino acid residues that are identical to those of aFGF are enclosed in boxes.

skin fibroblasts, whereas the amphibian neuropeptide bombesin stimulates Swiss 3T3 mitogenesis (*19*). If biologically active, the aFGF decapeptide would represent a related but unique neuropeptide since it would not contain the amidated carboxyl-terminal methionine residue required for biological activity and receptor binding of these neuropeptides.

Note added in proof. After we submitted this report, the complete amino acid sequence of the 146-residue pituitary-derived bFGF was published (*20*). The complete aFGF and bFGF sequences are 55 percent identical in the commonly aligned regions.

References and Notes

1. O. A. Trowell, B. Chir, E. N. Willmer, *J. Exp. Biol.* **16**, 60 (1939); R. S. Hoffman, *Growth* **4**, 361 (1940).
2. K. A. Thomas, M. Rios-Candelore, S. Fitzpatrick, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 357 (1984).
3. P. Bohlen, F. Esch, A. Baird, K. L. Jones, D. Gospodarowicz, *FEBS Lett.* **185**, 177 (1985); P. Bohlen, A. Baird, F. Esch, N. Ling, D. Gospodarowicz, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5364 (1984).
4. A. B. Schreiber *et al.*, *J. Cell Biol.*, **101**, 1623 (1985).
5. K. A. Thomas *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6409 (1985).
6. K. A. Thomas, M. C. Riley, S. K. Lemmon, N. C. Baglan, R. A. Bradshaw, *J. Biol. Chem.* **255**, 5517 (1980).
7. R. R. Lobb and J. W. Fett, *Biochemistry* **23**, 6295 (1984).
8. M.-D. Duo, S. S. Huang, J. S. Huang, *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **44**, 695 (1985).
9. D. Gospodarowicz, in *Mediators in Cell Growth and Differentiation*, R. J. Ford and A. L. Maizel, Eds. (Raven, New York, 1985), vol. 37, pp. 109–134.
10. Bovine brain-derived bFGF was purified through the homogenizations, salt precipitations, and CM 50 ion exchange chromatography as described previously (*2*). Material eluted from the CM 50 column was dialyzed against 10 m*M* tris-HCl containing 0.6*M* NaCl (*p*H 7.0), loaded on a heparin-Sepharose column equilibrated in the same buffer, washed extensively with buffer containing 0.6 and 1.0*M* NaCl, and eluted with a 1.5*M* NaCl buffer solution. The eluant was loaded on a C_4 HPLC column and eluted as previously described (*2*). The protein was pure as assessed by high-sensitivity silver-stained 15 percent polyacrylamide electrophoretic gels.
11. M. O. Dayhoff, *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, D.C., 1978), vol. 5, suppl. 3, pp. 345–358.
12. P. T. Lomedico *et al.*, *Nature (London)* **312**, 458 (1984); P. E. Auron *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7907 (1984); C. J. March *et al.*, *Nature (London)* **315**, 641 (1985).
13. T. P. Hopp and K. R. Wood, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3824 (1981); J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
14. R. M. Sweet and D. Eisenberg, *J. Mol. Biol.* **171**, 479 (1983).
15. G. D. Rose, A. R. Geselowitz, G. J. Lesser, R. H. Lee, M. H. Zehfus, *Science*, **229**, 834 (1985).
16. P. Cameron, G. Limjuco, J. Rodkey, C. Bennett, J. A. Schmidt, *J. Exp. Med.*, **162**, 790 (1985); J. Schmidt *et al.*, in *The Physiological, Metabolic, and Immunological Actions of IL-1*, M. J. Kluger, J. J. Oppenheim, M. C. Powanda, Eds. (Liss, New York, in press).
17. J. Van Damme *et al.*, *Nature (London)* **314**, 266 (1985).
18. L. L. Iverson, *Annu. Rev. Pharmacol. Toxicol.* **23**, 1 (1983); N. Minamino *et al.*, *Biochem. Biophys. Res. Commun.* **119**, 14 (1984).
19. J. Nilsson, A. M. von Euler, C. J. Dalsgaard, *Nature (London)* **315**, 61 (1985); E. Rozengurt and J. Sinnett-Smith, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2936 (1983).
20. F. Esch *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6507 (1985).

21. We thank E. Cordes and E. Scolnick for their support of this work.

20 August 1985; accepted 18 October 1985

A Model for Fibrinogen: Domains and Sequence

Abstract. *Electron microscopy of rotary-shadowed fibrinogen demonstrates that the molecules modified for crystallization by limited cleavage with a bacterial protease retain the major features of the native structure. This evidence, together with image processing and x-ray analysis of the crystals and of fibrin, has been used to develop a three-dimensional low resolution model for the molecule. The data indicate that the two large end domains of the molecule would be composed of the carboxyl-terminus of the Bβ chain (proximal) and γ chain (distal), respectively; the carboxyl-terminus of the Aα chain would fold back to form an additional central domain. On this basis, the carboxyl-terminal region of each of the three chains of fibrinogen is folded independently into a globular domain.*

JOHN W. WEISEL
Department of Anatomy, University of Pennsylvania School of Medicine, Philadelphia 19104
CYNTHIA V. STAUFFACHER
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907
ESTHER BULLITT
Graduate Biophysics Program, Brandeis University, Waltham, Massachusetts 02254
CAROLYN COHEN
Rosenstiel Basic Medical Sciences Research Center, Brandeis University

The early trinodular Hall-Slayter model for fibrinogen (*1*), based on electron microscope images of shadowed molecules, has provided a valid, simplified picture of the molecule despite a variety of other conflicting proposals (*2*). Detailed information on the morphology and dimensions of fibrinogen has been derived from analysis of crystalline arrays (*3–9*). Native fibrinogen does not crystallize, but crystals have been produced after limited cleavage of the molecule by a bacterial protease. Image processing of electron micrographs of these ordered arrays had led to an improved structure for the molecule, consisting of a linear arrangement of seven domains, termed the heptad model. This scheme also accounts for the band pattern of fibrin in the electron microscope (*7, 9*). The heptad model differs in detail, however, from some recent electron microscope images of individual molecules (*10–13*). We now show that many of

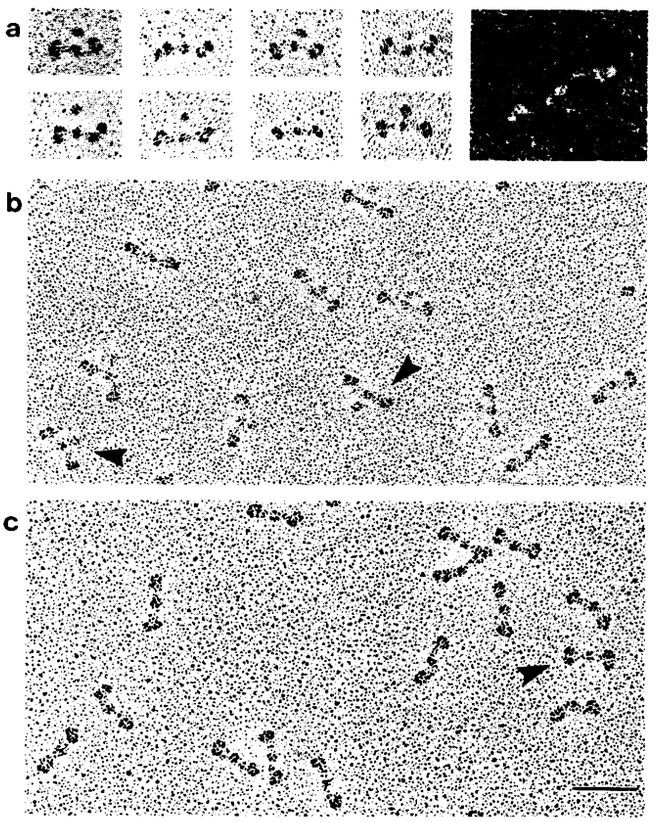


Fig. 1. Electron microscope images of individual, rotary-shadowed bovine fibrinogen molecules. (a and b) Native fibrinogen. (a) A gallery of selected molecules illustrating a range of observed morphologies. The inset (at a higher magnification) shows a negatively stained bovine fibrinogen molecule (*23*) displaying the staggering of the terminal domains. (b) Typical field showing molecules whose overall shape is trinodular. Arrows point to two molecules showing an additional central domain and a subdivision of the end region into two domains. (c) Fibrinogen modified by the *Pseudomonas* protease to a point where crystals are formed (*5*). The molecules are similar to native fibrinogen, but the extra central domain is not seen. The arrow points to a molecule with a distinguishable division of the end domains. The lengths of all molecules average 460 ± 25 Å. Native and modified fibrinogen samples were prepared for electron microscopy and photographed as described in Flicker *et al.* (*24*); scale bar, 500 Å.