Biochem. J. (1988) 255, 1057–1060 (Printed in Great Britain)

Is vitellogenin an ancestor of apolipoprotein B-100 of human low-density lipoprotein and human lipoprotein lipase?

Michael E. BAKER
Department of Medicine, M-023, University of California, San Diego, La Jolla, CA 92093, U.S.A.

Vitellogenin, an ancient animal protein, is the major yolk protein of eggs, where it is used as a food source during embryogenesis. Here it is shown that vitellogenins, including those from the invertebrates *Caenorhabditis elegans* and *Drosophila melanogaster*, contain domains that are homologous with parts of apolipoprotein B-100 (apoB-100) of human low-density lipoprotein and human lipoprotein lipase. As vitellogenins are likely to have been used by invertebrates during embryogenesis well before the circulation of lipids appeared in vertebrates, it is suggested that copies of a precursor gene, serving a function similar to vitellogenin, were modified to code for part of apoB-100 and lipoprotein lipase in vertebrates. In addition to providing a link between invertebrates and vertebrates for proteins involved in lipid transport, these homologies suggest new functions for vitellogenin other than being a yolk food for the developing embryo.

INTRODUCTION

Serum proteins in the early vertebrates are able to function in ligand transport and blood clotting like their descendents in mammals [1–3]. Although it is clear that many of these serum proteins evolved by exon shuffling and gene duplication [2,4,5], the details of events prior to 450 million years ago, when vertebrates first appeared, are still unknown. Here it is shown that two modern vertebrate serum proteins, apolipoprotein B-100 (apoB-100) of human low-density lipoprotein (LDL) and human lipoprotein lipase contain domains that are similar to parts of vitellogenins from the invertebrates *Caenorhabditis elegans* and *Drosophila melanogaster*.

ApoB-100, lipoprotein lipase and vitellogenin have apparently distinct functions. The first two proteins have important roles in the utilization of cholesterol, triacylglycerols, and other lipids by vertebrates. ApoB-100, an $M_r \sim 510000$ protein, is part of very-low-density lipoprotein (VLDL), intermediate-density lipoprotein (IDL) and LDL [4,6–11], and it functions as a transporter of cholesterol and triacylglycerols to cells throughout the body. The first 2150 amino acids of apoB-100 constitute apoB-48, which is found in chylomicrons [12] and chylomicron remnants. Lipoprotein lipase hydrolyses triacylglycerols of circulating VLDL [13–15], converting VLDL to IDL.

Vitellogenin is an egg-yolk protein used by oviparous vertebrates and invertebrates to feed the developing embryo [16–18]. Vertebrates synthesize vitellogenin in the liver, from which it is secreted into the blood and transported to oocytes, where it is selectively taken up by receptor-mediated endocytosis [19]. Once inside the oocyte, frog and chicken vitellogenin are cleaved to form lipovitellin I, phosvitin and lipovitellin II. Nematode vitellogenin [20] contains sequences homologous with vitellin I and II [20a,21], but lacks the phosvitin sequence. In nematodes [20] and sea urchins [22], vitellogenin is synthesized in the intestine. *Drosophila* yolk proteins 1 and 2 [23–25] are synthesized in the female fat body and ovary [26]. Like vertebrate vitellogenins, invertebrate vitellogenins are taken up by the oocyte and then processed for use as food for the developing embryo. Aside from being a food for the developing embryo, vitellogenin has no other known biological function.

Interest in the origin of serum proteins involved in steroid binding and blood clotting [27–29] led to the unexpected finding that a ~ 250 residue segment at the *C*-terminus of vitellogenin is homologous with von Willebrand factor (M. E. Baker, unpublished results). To explore the possibility that other modern vertebrate proteins were derived from vitellogenin, vitellogenins were compared with proteins in the database with the Lipman–Pearson FASTP program [30]. It was found that: (1) segments on vitellogenins from the invertebrate *Caenorhabditis elegans*, as well as frog and chicken vitellogenin, were similar to parts of human apoB-100; and (2) *Drosophila melanogaster* vitellogenin was similar to several lipases, with the strongest similarity to human lipoprotein lipase. It is proposed that vitellogenins are ancestors of apoB-100 and lipoprotein lipase, and that vitellogenin may have biological functions in addition to being a yolk food for the developing embryo.

RESULTS

Drosophila vitellogenin is similar to human lipoprotein lipase

The Lipman–Pearson FASTP search revealed that both yolk proteins from *Drosophila* were similar to three lipases in the database: human lipoprotein lipase [14,15]; porcine pancreatic lipase [31,32]; and dog pancreatic lipase [33]. Yolk proteins 1 and 2 are about 50% similar [25]. Yolk protein 1 showed more similarity to these lipases than did yolk protein 2. Among the lipases, the closest similarity was between human lipoprotein lipase and yolk protein 1. Fig. 1 shows most of the similar

Abbreviations used: ApoB-100, apolipoprotein B-100; LDL, low-density lipoprotein; VLDL, very-low-density lipoprotein; IDL, intermediate-density lipoprotein; DFP, di-isopropyl fluorophosphate.

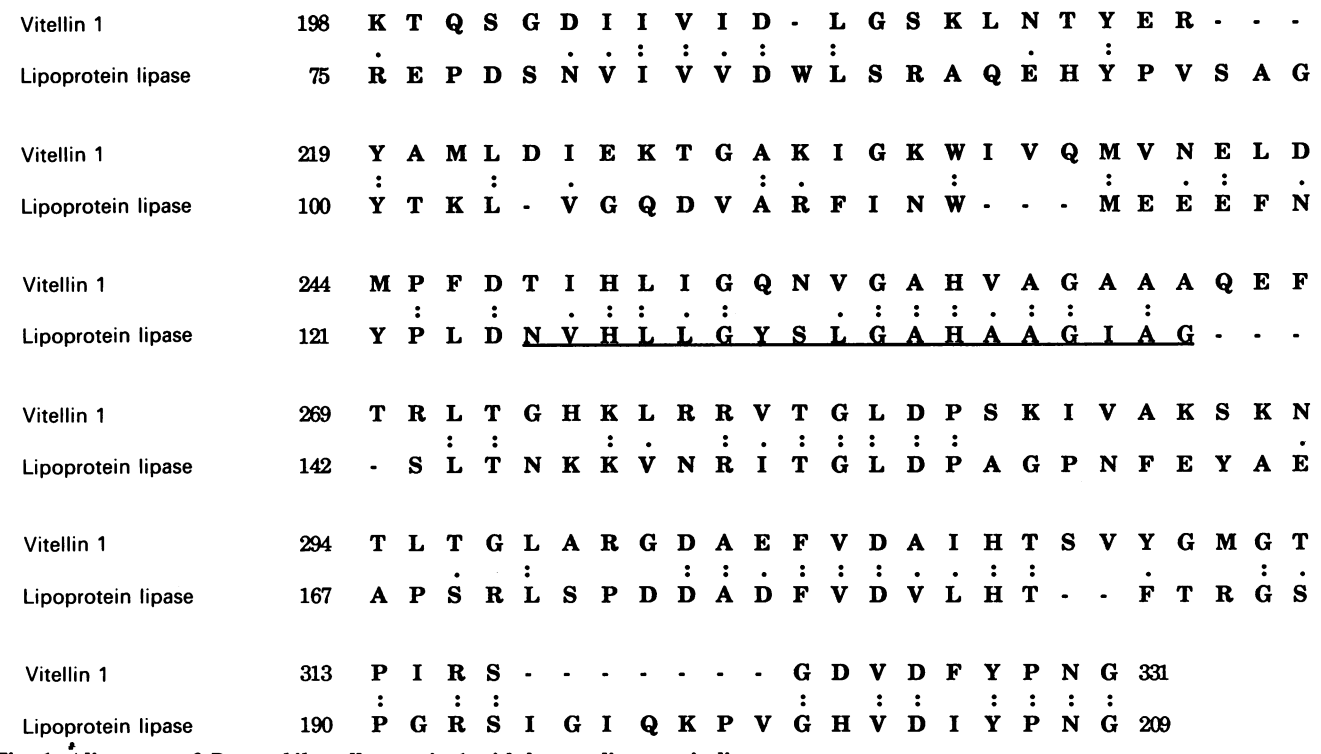

Fig. 1. Alignment of *Drosophila* yolk protein 1 with human lipoprotein lipase

Out of 124 possible matches there are 50 identities (40%) and 34 conservative replacements (27%). ':' Shows identities; '.' shows conservative replacements. The ALIGN analysis of these amino acid sequences, using a gap penalty 8, gives a score that is 11.9 S.D. higher than that obtained with 1000 comparisons of randomized sequences of these segments. The probability of getting such a score by chance is 6×10^{-33}. The segment on lipoprotein lipase with homology to the interfacial lipid-binding site on porcine lipase [31] is underlined. Serine-132 on lipoprotein lipase, which corresponds to the active-site serine on porcine lipase, is replaced with asparagine-225 on vitellin 1.

region in human lipoprotein lipase and *Drosophila* vitellogenin 1, including the region on lipoprotein lipase that is similar to the substrate-binding site on porcine pancreatic lipase [32]. The ALIGN [34] score for these segments is 11.9 S.D. units higher than that obtained with 1000 comparisons of randomized sequences of these segments. The probability of getting this score by chance is 6×10^{-33}.

The sequence similarity does not allow the unequivocal determination of whether vitellogenin and lipoprotein lipase are derived from a common ancestor or whether they are examples of convergent evolution. However, scores about 10 S.D. units usually indicate common ancestry [30], especially when the proteins have other properties in common, such as recognizing similar kinds of hydrophobic ligands.

Caenorhabditis elegans, frog and chicken vitellogenin are similar to human apoB-100

The Lipman–Pearson FASTP search also indicated that apoB-100 was similar to *C. elegans* vitellogenin, which is homologous with frog and chick vitellogenin [20a,21]. A graphical representation of the comparison of the first 1000 residues at the *N*-terminus of apoB-100 with *C. elegans* and *Xenopus laevis* vitellogenin, shown in Fig. 2, shows several similar segments of over 100 residues in these proteins. The ALIGN computer analysis was used to quantify the similarity between apoB-100 and vitellogenins. The ALIGN comparison score for residues 211–684 of apoB-100 with residues 247–725 of *C. elegans* vitellogenin was 10.35 S.D. units ($P < 10^{-25}$) higher than that of 200 comparisons of randomized sequences of these proteins. An ALIGN analysis of residues 19–587 of apoB-100 with residues 24–607 of frog vitellogenin and with residues 24–605 of chick vitellogenin yielded scores of 15.95 and 12.75 S.D. units, respectively.

Comparisons of vitellogenins

Nardelli *et al.* [21] presented an alignment of *C. elegans* vitellogenin with frog and chicken vitellogenin, which indicated that these proteins were homologous. The most similar part of vitellin I was in the *N*-terminal 850 residues. To obtain a quantitative measure of the similarity, the RELATE computer program [34] was used to compare the first 850 residues of *C. elegans* vitellogenin with frog and chicken vitellogenin using a 50 residue segment, which yielded scores of 12.2 and 10.9 S.D. units respectively. Other comparisons showed that *Drosophila* vitellogenins 1 and 2 do not appear to be similar to either frog or chicken vitellogenin. However, *Drosophila* vitellogenin 1 contains a 200 residue segment that is similar to two segments in the vitellin I part of *C. elegans* vitellogenin. A RELATE analysis of residues 1–200 of *Drosophila* vitellogenin 1 with residues 300–800 of *C. elegans* vitellogenin, using a 30 residue segment, yielded a comparison score of 6.7 S.D. units ($P = 10^{-11}$). The part of *Drosophila* vitellogenin that is similar to *C. elegans* vitellogenin does not include the segment that is similar to lipases. As sequences of other vitellogenins become available, it will be possible to establish more definitively the evolutionary relationships among these proteins, as well as their relationship to other lipophilic proteins.

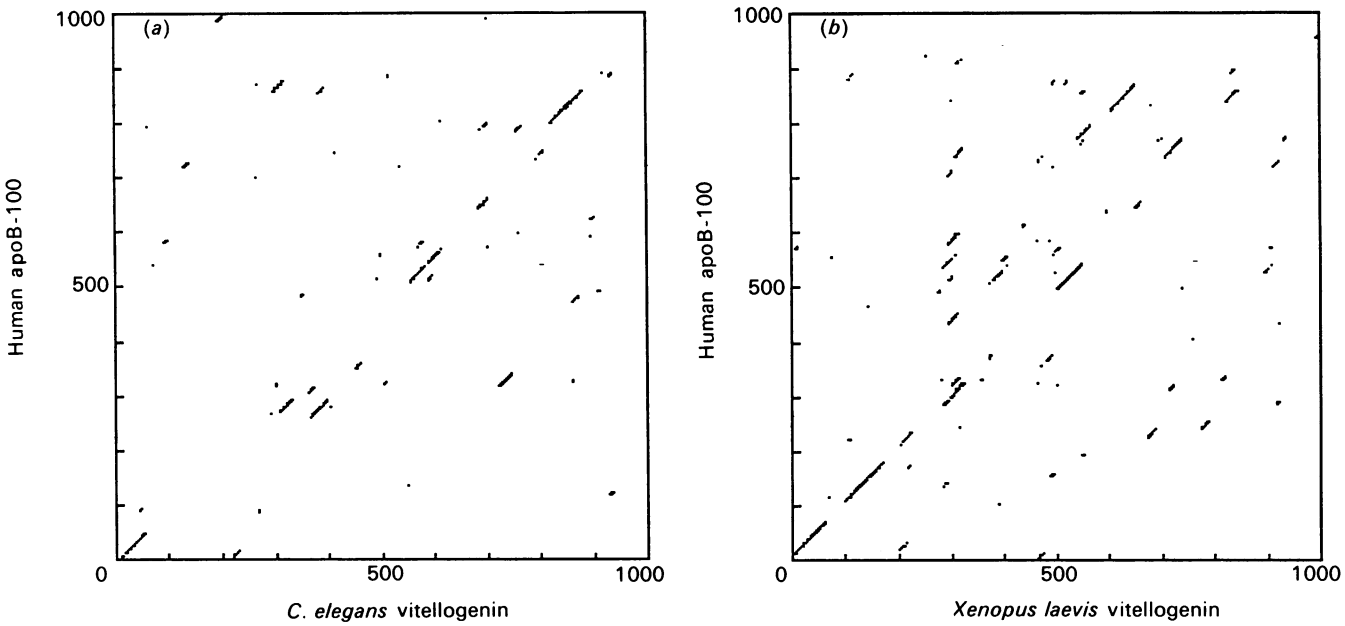

Fig. 2. Graphical comparison of human apoB-100 with *C. elegans* vitellogenin and *Xenopus laevis* vitellogenin

The first 1000 residues of human apoB-100 were compared with the first 1000 residues of vitellogenin from *C. elegans* (*a*) and *Xenopus laevis* (*b*) using segments of 45 and 49 residues respectively. Each segment was scored with the Dayhoff matrix [34]. Scores above a minimum are represented by a dot at the midpoint of the segment. Thus there is additional similarity of 23 residues for the *C. elegans* and 25 residues for *Xenopus laevis* extending beyond each segment shown in the Figure. An ALIGN analysis of residues 247–725 of *C. elegans* vitellogenin with residues 211–684 of human apoB-100, using a gap penalty of 8, yielded a score of 10.35 S.D. units. An ALIGN analysis of residues 24–607 of frog vitellogenin with residues 19–587 of apoB-100 yielded a score of 15.95 S.D. units.

DISCUSSION

Similarity of apoB-100 to vitellogenin

There are biological similarities to support the notion of a common ancestor for vitellogenin and apoB-100. Like apoB-100, vitellogenin binds hydrophobic molecules such as phospholipids, triacylglycerols and cholesterol and it transports these molecules to a target cell. Also, like apoB-100, vitellogenin binds to a membrane receptor and enters the cell by endocytosis [19]. Moreover, chicken apoB-100 and vitellogenin synthesis is induced by oestrogen [35,36]. Interestingly, there is evidence for cross-reactivity of antibodies to mammalian LDL receptor with chicken-oocyte vitellogenin receptor [37]. The *N*-terminal 1000 residues of apoB-100, which were found to be similar to vitellogenin, comprise about 40% of apoB-48, which is found in chylomicrons [10,12]. Like *C. elegans* vitellogenin, apoB-48 is synthesized in the intestine. Based on the sequence similarity between vitellogenin and apoB-100 (and apoB-48) and their biological similarities, I suggest that the transfer of cholesterol and other lipophilic compounds to cells by human apoB-100 descended from the actions of vitellogenin secreted from *C. elegans* intestine.

Similarity of human lipoprotein lipase to vitellogenin

The similarity of *Drosophila* vitellogenin to human lipoprotein lipase is intriguing. Human lipoprotein lipase and dog and pig pancreatic lipase belong to a family of lipases [14,38], which include rat hepatic lipase, rat lingual lipase and human lecithin:cholesterol acyltransferase [39]. Although some of these lipases show substantial differences in parts of their amino acid sequences, they have conserved a sequence that has been identified as the interfacial lipid-binding region of porcine pancreatic lipase [32], where it is thought that substrate binding and catalysis by a serine residue, which also reacts with di-isopropyl fluorophosphate (DFP) [40,41], occurs.

The high comparison score for yolk protein 1 and lipoprotein lipase and the high conservation of the presumed lipase active site, underlined in Fig. 1, suggests that they have some common biological properties. However, examination of this segment on yolk protein 1 and 2 reveals that the serine residue that reacts with DFP is replaced with an asparagine and a glycine respectively, which may mean that yolk proteins 1 and 2 are not catalytically active. Of course, it is possible that a serine or another amino acid from another part of the primary structure is properly positioned in the tertiary structure to act as a nucleophile for an enzymic reaction. In seeking an enzymic activity in vitellogenin, it should be kept in mind that some lipases use different cofactors to modulate their enzymic activity. For example, the enzymic activity of human lipoprotein lipase is dependent on the cofactor apolipoprotein C-II, while pancreatic lipase requires the cofactor colipase for maximal activity. Thus it may require special conditions to observe enzymic activity or reaction with DFP in vitellogenin. A relevant precedent for lipid-transporting proteins to recognize DFP and have esterase activity comes from albumin and α-foetoprotein, the major transporters of fatty acids in adult and foetal serum, respectively. Both albumin and α-foetoprotein have sites that recognize DFP [28, 42–44]. Albumin uses a tyrosine residue to catalyse the hydrolysis of esters [42,43].

It is also interesting that lipoprotein lipase has a functional role in the conversion of VLDL to LDL,

which contain apoB-100, which is similar to a vitellogenin.

Vitellogenins are ancient animal proteins that were used by protostomes and deuterostomes to feed the developing embryo. However, it should be realized that the sequences that are available for analysis are different from the ancestral proteins, even if the organism, be it the nematode or the fly, retains many features of the ancestor that originated hundreds of millions of years ago. In this context, the good conservation of parts of invertebrate and vertebrate vitellogenin sequences with two mammalian proteins involved in lipid transport and metabolism over such a long time suggests that vitellogenins may have a function other than as a food source, for example, as a carrier for apolar molecules and perhaps as an enzyme. One good candidate in which to examine for such novel functions is sea-urchin vitellogenin, as suggested by Shyu *et al.* [22] who found large quantities of vitellogenin in male sea urchins. The sequence of this vitellogenin may contain more clues to the function(s) of this interesting protein.

I thank Darrell Fanestil and Bill Loomis for reviewing an earlier version of this manuscript and Bill Loomis for vigorous and thought-provoking discussions of principles of evolutionary biology. The encouragement and support of Andrew Ansel and Marvin Weissberg for this research is gratefully acknowledged. After this manuscript was completed I learned that the similarity between *Drosophila* vitellogenin and lipases was also discovered by Dr. Russell Doolittle, University of California, San Diego (unpublished results), and by Coulson *et al.* [45].

REFERENCES

1. Doolittle, R. F. (1984) in The Plasma Proteins (Putnam, F. W., ed.), vol **4**, pp. 317–360, Academic Press, New York
2. Doolittle, R. F. (1985) Trends Biochem. Sci. **10**, 233–237
3. Chapman, M. J. (1980) J. Lipid Res. **21**, 789–853
4. Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L. & Russell, D. W. (1984) Cell (Cambridge, Mass.) **39**, 27–38
5. Rogers, J. (1985) Nature (London) **315**, 458–459
6. Brown, M. S. & Goldstein, J. L. (1987) Nature (London) **330**, 113–114
7. Law, S. W., Grant, S. M., Higuchi, K., Hospattankar, A., Lackner, K., Lee, N. & Brewer, H. B., Jr. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 8142–8146
8. Knott, T. J., Pease, R. J., Powell, L. M., Wallis, S. C., Rall, S. C., Jr., Innerarity, T. L., Blackhart, B., Taylor, W. H., Marcel, Y., Milne, R., Johnson, D., Fuller, M., Lusis, A. J., McCarthy, B. J., Mahley, R. W., Levy-Wilson, B. & Scott, J. (1986) Nature (London) **323**, 734–738
9. Yang, C.-Y., Chen, S.-H., Gianturco, S. H., Bradley, W. A., Sparrow, J. T., Tanimura, M., Li, W.-H., Sparrow, D. A., DeLoof, H., Rosseneu, M., Lee, F.-S., Gu, Z.-W., Gotto, A. M., Jr. & Chan, L. (1986) Nature (London) **323**, 738–742
10. Cladaras, C., Hadzopoulou-Cladaras, M., Nolte, R. T., Atkinson, D. & Zannis, V. I. (1986) EMBO J. **5**, 3495–3507
11. Brown, M. S. & Goldstein, J. L. (1986) Science **232**, 34–47
12. Chen, S.-H., Habib, G., Yasng, C.-Y., Gu, Z.-W., Lee, B. R., Weng, S.-A., Silberman, S. R., Cai, S.-J., Deslypere, J. P., Rosseneu, M., Gotto, A. M., Jr., Li, W.-H. & Chan, L. (1987) Science **238**, 363–366
13. Nilsson-Ehl, P., Garfinkel, A. S. & Schotz, M. C. (1980) Annu. Rev. Biochem. **49**, 667–693
14. Wion, K. L., Kirchgessner, T. G., Lusis, A. J., Schotz, M. C. & Lawn, R. M. (1987) Science **235**, 1638–1641
15. Senda, M., Oka, K., Brown, W. V., Qasba, P. K. & Furuichi, Y. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 4369
16. Tata, J. R. & Smith, D. F. (1979) Recent Prog. Horm. Res. **35**, 47–95
17. Shapiro, D. J. (1982) CRC Crit. Rev. Biochem. **12**, 187–203
18. Wyatt, G. R. & Pan, M. L. (1978) Annu. Rev. Biochem. **47**, 779–817
19. Opresko, L. K. & Wiley, H. S. (1987) J. Biol. Chem. **262**, 4109–4115
20. Speith, J., Denison, K., Zucker, E. & Blumenthal, T. (1985) Nucleic Acids Res. **13**, 7129–7138

20a. Spieth, J., Denison, K., Kirtland, S., Cane, J. & Blumenthal, T. (1985) Nucleic Acids Res. **13**, 5283–5295

21. Nardelli, D., Gerber-Huber, S., van het Schip, F. D., Gruber, M., Geert, A. B. & Wahli, W. (1987) Biochemistry **26**, 6397–6402
22. Shyu, A.-B., Raff, R. A. & Blumenthal, T. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 3865–3869
23. Hovemann, B., Galler, R., Walldorf, U., Kupper, H. & Bautz, K. F. (1981) Nucleic Acids Res. **9**, 4721–4734
24. Hung, M.-C. & Wensink, P. C. (1981) Nucleic Acids Res. **9**, 6407–6419
25. Hung, M.-C. & Wensink, P. C. (1983) J. Mol. Biol. **164**, 481–492
26. Bownes, M. (1986) Annu. Rev. Entomol. **31**, 507–531
27. Baker, M. E., French, F. S. & Joseph, D. (1987) Biochem. J. **243**, 293–296
28. Baker, M. E. (1985) FEBS Lett. **182**, 47–52
29. Baker, M. E. (1988) Tumor Biol. **9**, 123–136
30. Lipman, D. J. & Pearson, W. R. (1985) Science **227**, 1435–1441
31. DeCaro, J., Boudourd, M., Guidoni, A., Desnuelle, P. & Rovery, M. (1981) Biochim. Biophys. Acta **671**, 129–138
32. Guidoni, A., Benkouka, F., DeCaro, J. & Rovery, M. (1981) Biochim. Biophys. Acta **660**, 148–150
33. Kerfelec, B., LaForge, K. S., Puigserver, A. & Scheele, G. (1986) Pancreas **1**, 430–437
34. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983) Methods Enzymol. **91**, 524–545
35. Evans, M. I., O'Malley, P. J., Krust, A. & Burch, J. B. E. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 8493–8497
36. Evans, M. I., Silva, R. & Burch, J. B. E. (1988) Genes Dev. **2**, 116–124
37. Stifani, S., George, R. & Schneider, W. J. (1988) Biochem. J. **250**, 467–475
38. Komaromy, M. C. & Schotz, M. C. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 1526–1530
39. McLean, J., Fielding, C., Drayna, D., Dieplinger, H., Baer, B., Kohr, W., Henzel, W. & Lawn, R. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 2335–2339
40. Reddy, M. N., Maraganore, J. M., Meredith, S. C., Heinrikson, R. L. & Kezdy, F. J. (1986) J. Biol. Chem. **261**, 9678–9683
41. Maraganore, J. M. & Heinrikson, R. L. (1986) Trends Biochem. Sci. **11**, 497–498
42. Sanger, F. (1963) Proc. Chem. Soc. 76–83
43. Means, G. E. & Wu, H. L. (1979) Arch. Biochem. Biophys. **194**, 526–530
44. Baker, M. E. (1986) Biochem. Int. **12**, 331–340
45. Coulson, A. F. W., Collins, J. F. & Lyall, A. (1987) Comput. J. **30**, 420–424

Received 14 July 1988/15 August 1988; accepted 25 August 1988