Biochem. J. (1988) 256, 1059–1063 (Printed in Great Britain)

BIOCHEMICAL JOURNAL LETTERS

Invertebrate vitellogenin is homologous to human von Willebrand factor

Vitellogenins are ancient proteins, found in egg yolk, and used by oviparious invertebrates and vertebrates to feed the developing embryo [1–3]. Vitellogenin from the nematode [4] and sea urchin [5] is synthesized in the intestine and secreted into the body fluid and transported to oocytes. 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 [6]. Aside from being a food for the developing embryo, vitellogenin has no other known biological function.

Serum proteins involved in blood clotting appeared 450 million years ago in primitive vertebrates [7,8]. Analysis of their amino acid sequences reveals that many of these proteins evolved by a combination of exon shuffling and gene duplication to form mosaic proteins with diverse functions [8,9]. Events in the evolution of serum proteins prior to 450 million years are are still unknown. Here we present evidence that von Willebrand factor, a $M_r \sim 250000$ protein that participates in the blood clotting process in vertebrates by mediating platelet binding to vascular walls [10,11], is homologous to vitellogenin found in the invertebrate *Caenorhabditis elegans*. In addition to providing a link to invertebrates for a protein involved in blood clotting, this finding suggests new functions for vitellogenin other than as a food for the developing embryo.

The sequence of von Willebrand factor contains multiple copies of four distinct domains, A through D, clearly indicating that it formed by exon shuffling and gene duplication [12–16]. It contains three copies of the ~ 200 residue A domain, which is homologous to part of complement factor B; three copies of a ~ 30 residue B domain; two copies of a ~ 120 residue C domain, which is homologous to part of thrombospondin; four complete copies of a ~ 360 residue D domain, and a ~ 100 residue fragment of the D domain. As described below, it is the D domain that we find is similar to a part of vitellogenin found in the nematode *C. elegans*, chicken, and frog.

Our interest in the origin of serum proteins involved in blood clotting and steroid binding [17–19] led us to search the protein database with the Lipman–Pearson FASTP program [20] to see if the amino acid sequence of von Willebrand factor was similar to that of other proteins. Unexpectedly, we found that a ~ 1225 residue segment at the *N*-terminus of pro-von Willebrand factor

```
von Willebrand factor    366  C L V T G Q S H F K S F D N R Y F T F S - G I C Q Y L L A R D
                              :   :     : :       . :   :     .             :     : . : . :
C. elegans vitellogenin  1308 C K V - D Q S E V S T F H N V I Y R A P L T T C Y S L V A K D

von Willebrand factor         C Q D H S - F S I V I E T V Q C A D D R - - D A V C T R S V T
                              :   .       :   . . .     . .       .           :     :
C. elegans vitellogenin       C S E Q P R F A V L A K K I N K N S E E L L V K V V R R E E E

von Willebrand factor         V R L P G L H N S - L V K L K H G A G V A M D G Q D V Q L P L
                              .   .         .     : : : .                 .   . .   . .   .
C. elegans vitellogenin       I V V K K S D D K F L V K V D G K K V N P T E L E Q Y N I E I

von Willebrand factor         L K G D L R I Q H T V T A S V R L S Y G E D L Q M D W D G R G
                              :     . :   .               . : :     : :           . : :
C. elegans vitellogenin       L G D N L I V - - - - - - - I R L P Q G E - - - V R F D G Y T

von Willebrand factor         R L L V K L S P V Y A G K T - C G L C G N Y N G N Q G D D F L
                                .             :             : : : : : :   . : .     . . :
C. elegans vitellogenin       - V K T N M P S V A S Q N Q L C G L C G N N D G E R D N E F M

von Willebrand factor         T P S G L A E P R V E D F G N A W K L - H G D C Q 538
                              :                 : : . :       .   :       . : .
C. elegans vitellogenin       T A D N Y E T E D V E E F H R S Y L L K N E E C E 1475
```

Fig. 1. Comparison of human von Willebrand factor with *C. elegans* vitellogenin

(:) shows identities; (.) shows conservative replacements. Out of 161 possible matches there are 40 identities (25 %) and 34 conservative replacements (21 %). An ALIGN analysis of these segments yielded a score that was 9.9 standard deviations higher than that obtained with 1000 comparisons of randomized sequences of these proteins. The probability of getting such a score by chance is 10^{-23}.

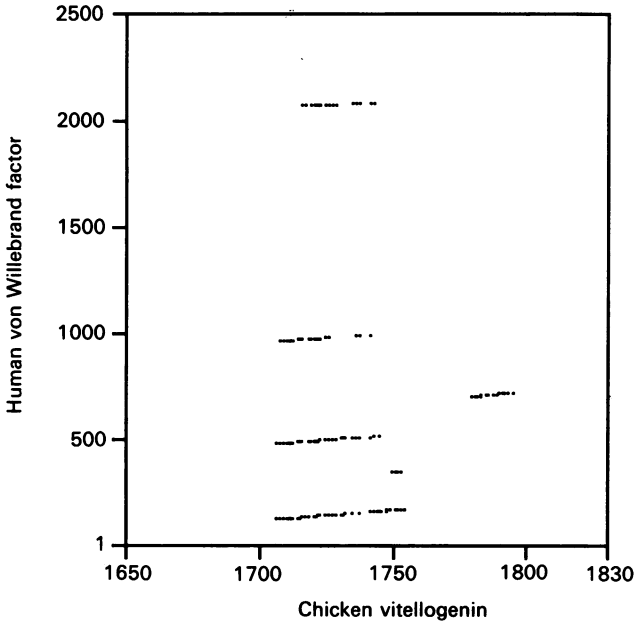

Fig. 2. Comparison of chicken vitellogenin with human von Willebrand factor

The Diagon analysis was used with a segment of 25 residues. Segments with scores above 285 are represented by a dot at the midpoint of the segment. Vitellogenin shows similarity to four homologous D domains and a 100 residue segment of a D domain. Ordinate: residues 1–2500 of human pro-von Willebrand factor; abscissa: residues 1650–1830 of chicken vitellogenin.

was similar to a 250 residue segment at the *C*-terminus of *C. elegans* vitellogenin. This part of human von Willebrand factor consists of domains D1, D2, the 100 residue D fragment, and D3. To quantify the similarity between vitellogenin and von Willebrand factor, we used the RELATE computer program [21], with a segment length of 40 to compare residues 1–1250 of von Willebrand factor with residues 1250–1600 of *C. elegans* vitellogenin. The comparison score was 12 standard deviations higher than that of 100 comparisons of randomized sequences of these segments. The probability of getting this score by chance is 1.7×10^{-33}. A score of this magnitude is usually interpreted as indicating that the segments are homologous, that is they are derived from a common ancestor. The alignment of part of D2, residues 366–538 of von Willebrand factor, with residues 1308–1475 of *C. elegans* vitellogenin is shown in Fig. 1. This invertebrate vitellogenin is homologous to chicken and frog vitellogenin [22], and the part that is similar to von Willebrand factor corresponds to vitellin II in chicken [23] and frog [24] vitellogenin, which comes after the phosvitin segment [25].

We examined both vertebrate vitellogenins for similarity to von Willebrand factor and found that both vitellogenins contain a ~ 250 residue segment at the *C*-terminus that is similar to residues 1–1225 of pro-von Willebrand factor. The RELATE comparison scores for chicken and frog vitellogenin with human von Willebrand factor are 14.5 and 11 standard deviations, respectively. Fig. 2 is a graphical representation of the similarity between chicken vitellogenin and human pre-von Willebrand factor. It shows that the similarity occurs at the parts of pro-von Willebrand factor that correspond to the four copies of the ~ 360 residue D domain and a ~ 100 residue fragment of the D domain [12,15]. D1 and D2 constitute most of the propeptide (also called von Willebrand antigen II), which circulates in plasma [26]. The ~ 100 residue segment and D3 are at the *N*-terminus of processed von Willebrand factor. A fourth D domain is between residues 1925 and 2275. Together the D domains constitute about 1550 residues, more than 50% of pro-von Willebrand factor. Thus a substantial part of von Willebrand factor is homologous to part of invertebrate, amphibian and avian vitellogenin.

The biological properties of von Willebrand factor may provide clues for the evolution of haemostasis, and for other biological functions for vitellogenin [5,27]. von Willebrand factor has adhesive properties, which include binding to platelets, collagen, and other proteins; it acts in haemostasis by promoting the adherence of platelets at sites where there is injury to the epithelium, and it binds factor VIII [11,28]. This binding site has been localized to the D domain [28], which we find is homologous to vitellogenin. The homology between vitellogenin and von Willebrand factor suggests that the *C*-terminal part of vitellogenin was used in the evolution of the intrinsic pathway for coagulation [29].

Amphibians and birds contain the proteins required for blood coagulation, which makes the conservation of the von Willebrand factor-like domain in vertebrate vitellogenin intriguing. An interesting possible function for this part of vitellogenin, based on von Willebrand factor's adhesive properties, is a role in binding the oocyte membrane receptor for vitellogenin [6], which leads to its uptake into the oocyte.

I thank Drs. Darrell Fanestil and Samuel Rappaport for review of an earlier version of this manuscript. The encouragement and support of Marvin Weissberg is gratefully acknowledged.

Michael E. BAKER

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

1. Tata, J. R. & Smith, D. F. (1979) Recent Progr. Hormone Res. **85**, 47–95
2. Shapiro, D J. (1982) CRC Crit. Rev. Biochem. **12**, 187–203
3. Wyatt, G. R. & Pan, M. L. (1978) Annu. Rev. Biochem. **47**, 779–817
4. Speith, J., Denison, K., Zucker, E. & Blumenthal, T. (1985) Nucleic Acids Res. **13**, 7129–7138
5. Shyu, A.-B., Raff, R. A. & Blumenthal, T. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 3865–3869
6. Opresko, L. K. & Wiley, H. S. (1987) J. Biol. Chem. **262**, 4109–4115
7. Doolittle, R. F. (1984) in The Plasma Proteins (Putnam, F. W., ed.), vol. 4, pp. 317–360, Academic Press, New York
8. Doolittle. R. F. (1985) Trends Biochem. Sci. **10**, 233–237
9. Rogers, ' (1985) Nature (London) **315**, 458–459
10. Hoyer, L. W. (1981) Blood **58**, 1–13
11. Lynch, D. C., Zimmerman, T. S. & Riggeri, Z. M. (1986) Br. J. Haematol. **64**, 15–20
12. Bonthron, D. T., Handin, R. I., Kaufman, R. J., Wasley, L. C., Orr, E. C., Mitsock, L. M., Ewenstein, B., Loscalzo, J., Ginsburg, D. & Orkin, S. H. (1986) Nature (London) **324**, 270–273
13. Shelton-Inloes, B. B., Titani, K. & Sadler, J. E. (1986) Biochemistry **25**, 3164–3170

14. Titani, K., Kumar, S., Takio, K., Ericsson, L. H., Wade, R. D., Ashida, K., Walsh, K. A., Chonek, M. W., Sadler, J. E. & Fujikawa, K. (1986) Biochemistry **25**, 3171–3184
15. Shelton-Inloes, B. B., Broze, G. R., Jr., Miletich, J. P. & Sadler, J. E. (1987) Biochem. Biophys. Res. Commun. **144**, 657–665
16. Hunt, L. T. & Barker, W. C. (1987) Biochem. Biophys. Res. Commun. **144**, 876–882
17. Baker, M. E., French, F. S. & Joseph, D. (1987) Biochem. J. **243**, 293–296
18. Baker, M. E. (1985) FEBS Lett. **182**, 47–52
19. Baker, M. E. (1988) Tumor Biol. **9**, 123–136
20. Lipman, D. J. & Pearson, W. R. (1985) Science **227**, 1435–1441
21. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983) Methods Enzymol. **91**, 524–545
22. Spieth, J., Denison, K., Kirtland, S., Cane, J. & Blumenthal, T. (1985) Nucleic Acids Res. **13**, 5283–5295
23. van het Schip, F. D., Samallo, J., Broos, J., Ophuis, J., Mojet, M., Gruber, M. & Geert, A. B. (1987) J. Mol. Biol. **196**, 245–260
24. Gerber-Huber, S., Nardelli, D., Haefliger, J.-A., Cooper, D. N., Givel, F., Germond, J.-E., Engel, J., Green, N. M. & Wahli, W. (1987) Nucleic Acids Res. **15**, 4737–4760
25. Nardelli, D., Gerber-Huber, S., van het Schip, F. D., Gruber, M., Geert, A. B. & Wahli, W. (1987) Biochemistry **26**, 6397–6401
26. Fay, D. J., Kawai, Y., Wagner, D. D., Ginsburg, D., Bonthron, D., Ohlsson-Montgomery, R. R. & Marder, V. J. (1985) Science **232**, 995–998
27. Taborsky, G. (1983) Adv. Inorg. Biochem. **5**, 235–279
28. Foster, P. A., Fulcher, C. A., Marti, T., Titani, K. & Zimmerman, T. S. (1987) J. Biol. Chem. **262**, 8443–8446
29. Ratnoff, O. D. (1987) Perspect. Biol. Med. **31**, 4–33

Received 7 June 1988

Sequence similarity between cyclophilin and elongation factor 2

Cyclophilin is a specific, high-affinity binding protein for the immunosuppressant cyclosporin A (Handschumacher *et al.*, 1984). When the amino acid sequence of bovine cyclophilin was published (Harding *et al.*, 1986), no similarity with sequences of other proteins was reported. Even a very recent paper on the complete nucleotide sequence of a cDNA encoding rat cyclophilin did not mention any similarity with other genes (Danielson *et al.*, 1988).

In 1986, the amino acid sequence of the eukaryotic elongation factor 2 (eEF-2) of protein biosynthesis was published (Kohno *et al.*, 1986). When we compared the amino acid sequences of cyclophilin and eEF-2, we found striking similarities (Fig. 1). A domain of 10 amino acids at the *N*-terminus shows 50% similarity and another domain, from amino acid 115 to 157 in cyclophilin and from 312 to 356 in eEF-2, shows 38% similarity. As indicated in Fig. 1, some of the differences in this region are non-conservative ones. It should be noted that this domain of cyclophilin contains the putative binding site for cyclosporin A (Dalgarno *et al.*, 1986). Furthermore, both proteins contain the sequence Glx-Xaa-Gly-Xaa-Xaa-Gly, which is characteristic for nucleotide binding proteins (Wierenga & Hol, 1983). The elongation factor eEF-2 is known to be a guanine nucleotide binding protein (Henriksen *et al.*, 1975).

These sequence similarities indicate that cyclophilin might be a nucleotide binding protein and that it possibly might play a role in protein biosynthesis. In this context it is intriguing that cyclosporin A is an effective inhibitor of phorbol ester-induced protein synthesis (Gschwendt *et al.*, 1988).

Michael GSCHWENDT, Walter KITTSTEIN and Friedrich MARKS

German Cancer Research Center, Institute of Biochemistry, D-6900 Heidelberg, Federal Republic of Germany

Dalgarno, D. C., Harding, M. W., Lazarides, A., Handschumacher, R. E. & Armitage, I. M. (1986) Biochemistry **25**, 6778–6784
Danielson, P. E., Forss-Petter, S., Brow, M. A., Calavetta, L., Douglass, J., Milner, R. J. & Sutcliffe, J. G. (1988) DNA **7**, 261–267
Dayhoff, M. O. (ed.) (1978) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington, DC
Gschwendt, M., Kittstein, W. & Marks, F. (1988) Skin Pharmacology, in the press
Handschumacher, R. E., Harding, M. W., Rice, J. & Drugge, R. J. (1984) Science **226**, 544–546
Harding, M. W., Handschumacher, R. E. & Speicher, D. W. (1986) J. Biol. Chem. **261**, 8547–8555
Henriksen, O., Robinson, E. A. & Maxwell, E. S. (1975) J. Biol. Chem. **250**, 720–724
Kohno, K., Uchida, T., Ohkubo, H., Nakanishi, S., Nakanishi, T., Fukui, T., Ohtsuka, E., Ikehara, M. & Okada, Y. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 4978–4982
Wierenga, R. K. & Hol, W. G. J. (1983) Nature (London) **302**, 842–844

Received 18 August 1988

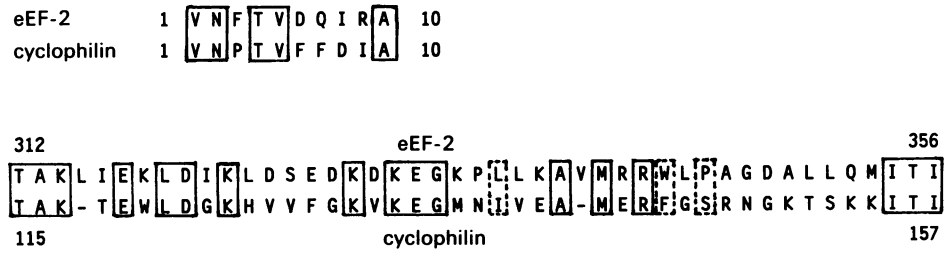

Fig. 1. Similarity between cyclophilin and eEF-2

Gaps in the alignment are indicated as dashes. Non-conservative differences according to the Dayhoff conservative categories (Dayhoff, 1978) are indicated by dashed boxes.