Epidermal growth factor

Is the precursor a receptor?

from Suzanne Pfeffer and Axel Ullrich

THE recent elucidation of the sequences of the cell-surface receptors for low density lipoprotein (LDL)[1,2] and epidermal growth factor (EGF)[3] have revealed an unexpected relationship between their structure and that of the biosynthetic precursor for EGF itself[4,5], and have enhanced speculation that the EGF precursor may play a dual role as both a growth factor precursor and a cell-surface receptor. New clues to this puzzle are provided by Rall and co-workers[6] on page 228 of this issue of *Nature*.

The 53-amino acid, mature EGF molecule is encoded within a 1,217-amino acid EGF precursor (refs 4* and 5). It contains an internal stretch of 20 hydrophobic amino acids, flanked by polar residues, which could anchor the precursor in the membrane and would divide it into a 159-amino acid cytoplasmic domain and a 1,010-amino acid extracellular domain (see figure). Why would a secreted growth factor be encoded within a membrane-bound precursor, and why is the precursor more than 20 times larger than the mature growth factor? In addition to EGF coding sequences, the precursor contains eight repeated units of a sequence that is related to EGF. Each unit encodes approximately 40 amino acids including 6 cysteine residues. These units may represent other biologically active polypeptides but, unlike the polypeptide units in precursors such as pro-opiomelanocortin[7], they are not flanked by obvious proteolytic cleavage sites.

Striking amino acid sequence similarities have recently been detected between bovine[1] and human LDL receptors[2] and the mouse EGF precursor. The most significant similarity (about 40 per cent) is between EGF precursor amino acid residues 565-701 and residues 457-595 in the LDL receptor extracellular domain (see figure). Sequence homology decreases in regions surrounding this core, but remains statistically significantly (approximately 30 per cent) for several hundred more amino acids. In addition to direct sequence homology, the LDL receptor and the EGF precursor share other structural features. Like the EGF precursor, the extracellular domain of the LDL receptor contains eight repeated, cysteine-rich units of 40 amino acids in length. The spatial distribution of the cysteine residues is conserved in each of the repeat units in both molecules but the relative positions of the cysteines differ between the LDL receptor and the EGF precursor repeat units. In both cases, other amino acid residues seem also to have been conserved. The repeat units of both proteins are distributed into two larger, distin-
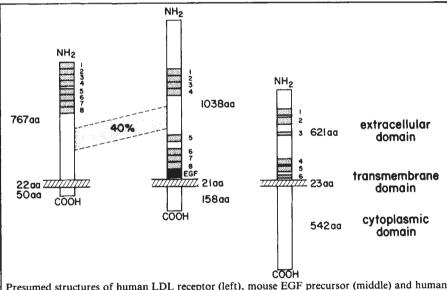
* A clerical error generated a frameshift at position 3707 in the mouse EGF precursor sequence reported in ref. 4.

guishable units, which are of similar size, and are arranged in tandem in the LDL receptor, yet are separated in the EGF precursor (see figure).

The structural theme of repeated, cysteine-rich sequences reappears within the human EGF receptor. Its extracellular domain contains four cysteine-rich repeated sequences of about 40 amino acids in length together with two of about half that size. The EGF receptor repeat units are enclosed within a larger duplicated sequence which involves the entire extracellular domain of the EGF receptor and seems to have led to the generation of six homologous repeat units[8]. Beyond these repeat units, however, the EGF receptor does not display significant sequence homology with either the EGF precursor or the LDL receptor.

The fact that the EGF precursor shares structural features with two cell-surface

Presumed structures of human LDL receptor (left), mouse EGF precursor (middle) and human EGF receptor (right) when inserted in the membrane. Hatched bars represent the membrane bilayer; shaded, numbered boxes depict cysteine-rich repeat units. The hatched band joins domains that are 40% matched in structure. Sizes of structural domains are given in amino acids (aa). The extracellular domain of the mouse EGF precursor includes a putative signal sequence.

receptors, including a hydrophobic domain which could anchor the precursor in the membrane, suggests that it has evolved from an ancestral cell-surface protein. Perhaps exon-shuffling events generated this family of nutrient and growth-factor receptors, with intragenic duplication events producing the repeated structural domains that evolved to serve best the specific needs of each protein. While the significance of the cysteine-rich repeat units for the EGF precursor and EGF receptor can at present only be guessed, they have been postulated to contribute to a multiply-looped, rigid LDL-binding domain in the LDL receptor[2].

Structural homologies between secreted

molecules and membrane receptors have been reported. For example, the receptor for transepithelial transport of two immunoglobulin molecules (IgA and IgM) displays structural homology with its ligands[9]. In this regard, it is interesting to note that the precursor for a polypeptide homologous to EGF, transforming growth factor (TGF)-alpha, also possesses a putative transmembrane sequence in a similar position relative to the mature growth factor sequence[10]. However, the TGF-alpha precursor is considerably smaller than the EGF precursor and lacks cysteine-rich repeats.

Together, these findings suggest that the EGF precursor may be a membrane protein in addition to being the precursor of EGF. That idea is strengthened by the detection by Rall *et al.* of the unprocessed EGF precursor but not mature EGF in regions of the mouse kidney. It will be important next to determine if the precursor is actually a membrane protein, and if so, whether it is localized on the cell surface or remains within the cell. Should this molecule be located on the cell surface, determination of its function will present

a major new challenge. It is possible that by regulation of protein processing, the EGF precursor serves different functions in different tissues.   □

1. Russell, D.W. *et al. Cell* **37**, 577 (1984).
2. Yamamoto, T. *et al. Cell*, **39**, 27 (1984).
3. Ullrich, A. *et al. Nature* **309**, 418 (1984).
4. Gray, A. *et al. Nature* **303**, 722 (1983).
5. Scott, J. *et al. Science* **221**, 236 (1983).
6. Rall, L.B. *et al. Nature* **313**, 228 (1985).
7. Nakanishi, S. *et al. Nature* **287**, 752 (1980).
8. Toh, H. *et al. Nature* (in the press).
9. Mostov, K.E., Friedlander, M. & Blobel, G. *Nature* **308**, 37 (1984).
10. Derynck, R. *et al. Cell* **38**, 287 (1984).

Suzanne Pfeffer is in the Department of Biochemistry, Stanford University, Stanford, California 94305; Axel Ullrich is at Genentech, Inc., 460 Pt. San Bruno Blvd. South San Francisco, California 94080, USA.