PRIMARY STRUCTURE OF HUMAN INSULIN-LIKE GROWTH FACTOR II

Ernst RINDERKNECHT and René E. HUMBEL
Biochemisches Institut, Universität Zürich, CH-8028 Zürich, Switzerland

Received 20 March 1978

1. Introduction

Vertebrate blood contains an insulin-like activity distinguishable from insulin by the lack of cross-reactivity with antibodies to insulin [1,2]. The low molecular weight form of this non-suppressible insulin-like activity (NSILA-s) [2] could be identified with 2 polypeptides isolated from human plasma [3]. Besides having insulin-like activity, both polypeptides display growth-promoting activity on chick embryo fibroblasts if added to a chemically defined medium at a concentration of 10^{-9} M [3]. The structural similarity of these 2 substances with the amino acid sequence of insulin has prompted us to name these polypeptides insulin-like growth factors (IGF I and II) [4]. The total primary structure of IGF I has been reported [5]. We now have determined the complete amino acid sequence of IGF II.

2. Methods

IGF II was isolated from a human plasma fraction as in [3]. Modification by *S*-[^{14}C]carboxymethylation or *S*-aminoethylation, cleavage by trypsin, chymotrypsin and 0.25 M acetic acid and separation of the resulting peptides followed the methods used for the sequence determination of IGF I [5]. Peptides A-1 and A-7 (see fig.1) had to be separated from each other on SP-Sephadex C-25 in 30% aqueous acetic acid by elution with a NaCl gradient. The 28 residue peptide A-5 (see fig.1) was further digested with thermolysin and the peptides separated by paper electrophoresis, at pH 6.5. Sequence determination was carried out on an automatic sequencer (Beckman Model 890 B, updated) as in [5]. Repetitive yields were from 90–95%.

3. Results

The complete amino acid sequence of IGF II is shown in fig.1, together with notations indicating the peptides and the methods used in establishing the structure. The sequence of residues 1–31 has been determined by Edman degradation of S-pyridylethylated IGF II [4]. The proposed sequence was confirmed by sequence analysis of fragments obtained with trypsin, chymotrypsin and acetic acid. The lack of N-terminal Ala in about 25% of all molecules of IGF II could be confirmed by Edman degradation of peptides T-1 and A-1. Residues 32–49 were determined on peptides T-3, T-4, T-6 and T-7 and on thermolytic subpeptides of A-5, residues 50–67 on peptides T-8, C-6 and C-8. Further confirmation of the sequence was obtained from supplementary analyses of overlapping peptides as indicated in fig.1. The amino acid composition derived from the sequence is shown in table 1, the calculated minimal mol. wt is 7471. Earlier determinations had shown that the 6 half-cystine residues form 3 intrachain disulfide bridges [3].

4. Discussion

The primary structure of IGF II is similar to that of IGF I. If aligned as in fig.2, 45 of 73 amino acid positions, or 62%, are identical. Their most characteristic common feature is the homology to proinsulin.

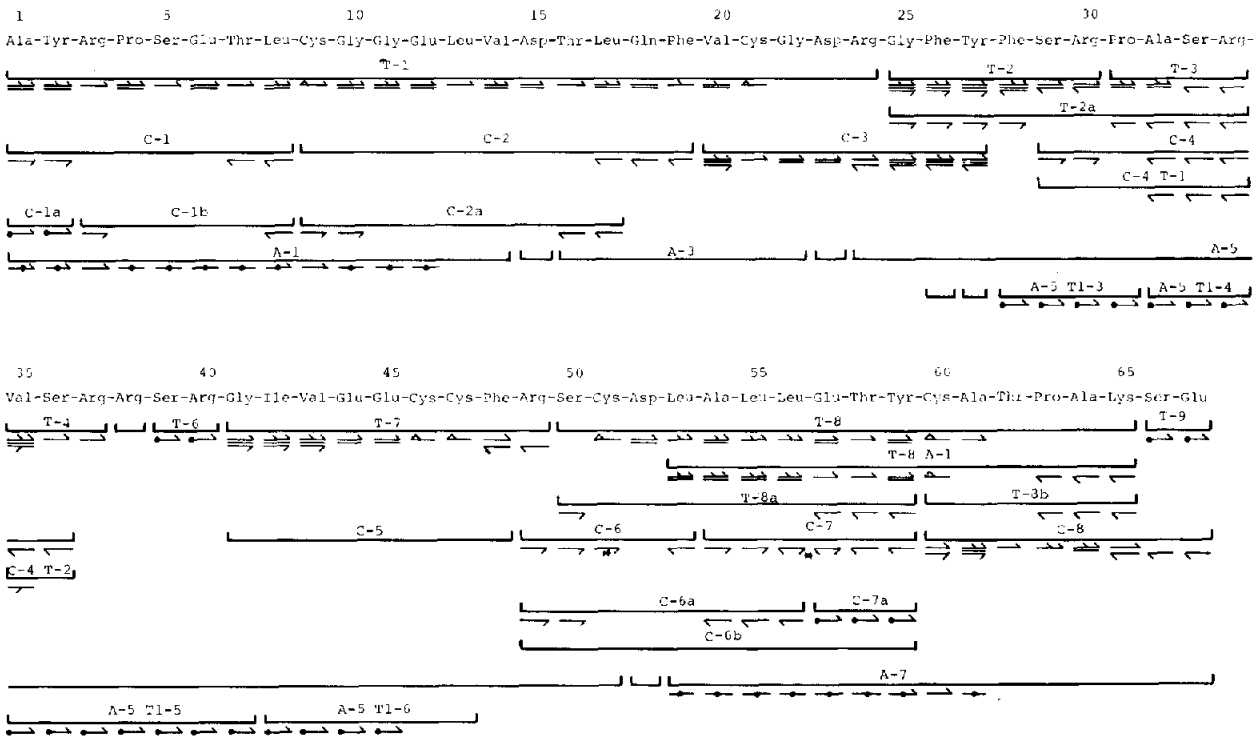

Fig.1. Amino acid sequence of human IGF-II and a schematic outline of the data supporting the sequence. The letter code indicates the cleavage method used to produce the peptides: T, trypsin; C, chymotrypsin; A, acetic acid; Tl, thermolysin. Sequence data obtained by automatic Edman degradation of the intact protein (31 residues, [4]) are not included in fig.1. To indicate the method of sequential degradation and of identification of residues, the following symbols are used: (⇀) automatic Edman degradation and amino acid analysis of the hydrolyzed PTH-derivatives; (⇁) automatic Edman degradation and gas chromatography of PTH-derivatives; (=) automatic Edman degradation and thin-layer chromatography of the PTH-derivatives; (⇀●) automatic Edman degradation and high-pressure liquid chromatography of PTH-derivatives; (⇀△) automatic Edman degradation and liquid scintillation counting; (⇀) digestion with aminopeptidase M or aminopeptidase from *Aeromonas* and amino acid analysis; (⇀) dansylation and thin layer chromatography; (⇀) digestion with carboxypeptidase C and amino acid analysis; (●⇀) manual Edman degradation (direct and/or subtractive). Peptides marked with an asterisk were further digested with trypsin or staphylococcal protease, the fragments separated and analyzed for their amino acid composition.

As proinsulin, IGF consists of the domains B, C and A. The sequence similarity of IGF to proinsulin is confined to the B and A domain (fig.2). The C domain (connecting peptide) of IGF II is only 8 residues long compared to 12 in IGF I and to 35 in proinsulin. It has been pointed out that a connecting peptide as short as 3 residues suffices to span residues B30 and A1 and to allow the proper 3-dimensional folding of insulin (cf. [6]).

Out of 19 invariant residues present in all insulins so far sequenced, the same 17 residues are present in IGF II as in IGF I. In fact, a 3-dimensional model of IGF I could be constructed very similar to the one of insulin [6], a further corollary to the postulated evolutionary relationship of IGF and insulin.

Another feature common to IGF I and II, but not found in proinsulin, is a C-terminal extension (D domain) which is 6 residues long in IGF II, 2 residues shorter than in IGF I. An analogous extension in the nucleotide sequence of proinsulin messenger RNA has been described in [7]. Although this nucleotide sequence shows no homology to the D domain of IGF, the former might be an evolutionary relic of a common ancestor of proinsulin and IGF.

Table 1
Amino acid composition of IGF I and II (no./molecule from amino acid sequence)

Residue	IGF-I	IGF-II
Lys	3	1
Arg	6	8
Asp	4	3
Asn	1	0
Thr	3	4
Ser	5	7
Glu	4	6
Gln	2	1
Pro	5	3
Gly	7	5
Ala	6	5
Cys	6	6
Val	3	4
Met	1	0
Ile	1	1
Leu	6	6
Tyr	3	3
Phe	4	4
Total residues	70	67
Calc. mol. wt	7649	7471

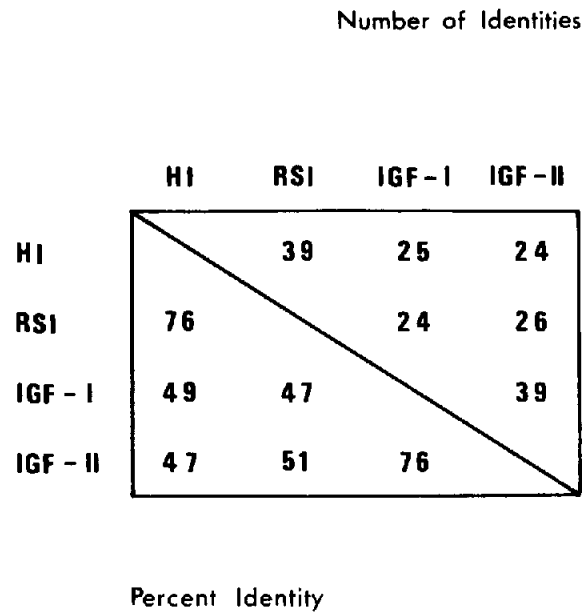

Fig.3. Matrix comparing sequence identities between human insulin (HI), rattlesnake insulin (RSI) and A and B domains of IGF I and II. The data for insulins are taken from [10].

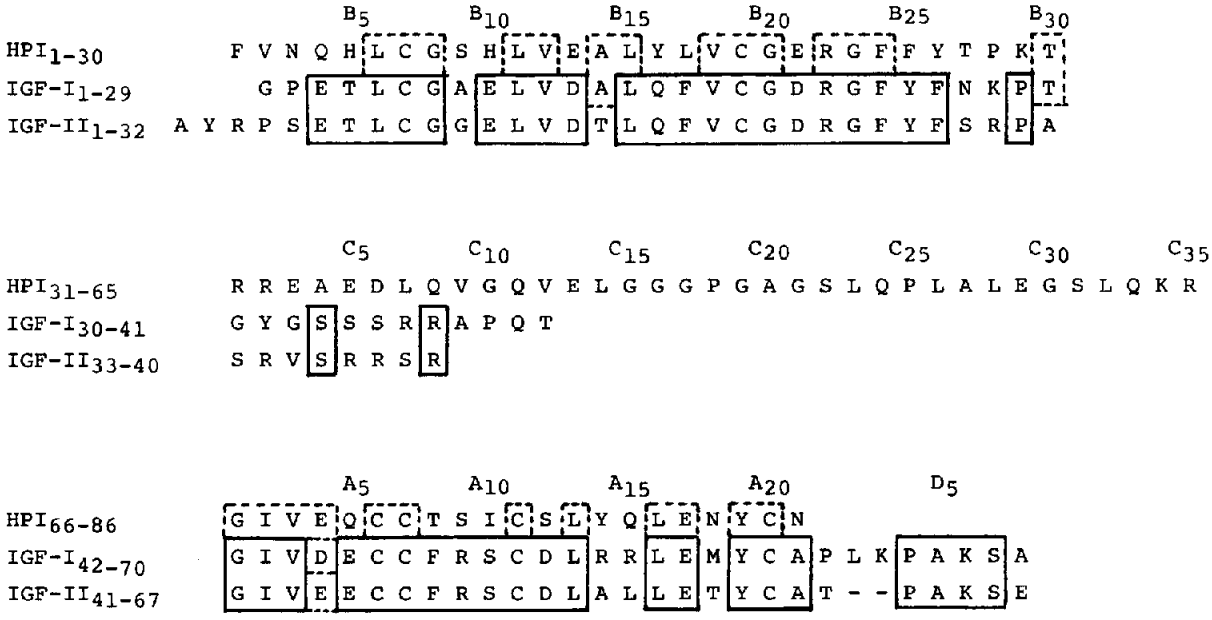

Fig.2. Primary structure of human proinsulin (HPI), IGF I and IGF II. Alignment has been so chosen as to give maximal homology. Boxes in solid lines indicate residues identical in IGF I and II; boxes in broken lines indicate residues identical in HPI and in IGF I and/or II.

A somewhat different extension beyond residue A21 has been found in nerve growth factor, which may be another, though disputed, member of the insulin family [8]. The extension in nerve growth factor has been interpreted as the result of a gene duplication [8]. In relaxin, the 4th member of this family, neither a C nor a D domain has been found as yet [9].

Based on the differences in amino acid sequence between insulins from different classes of vertebrates and IGF I, we have suggested that the divergence of IGF and proinsulin from a common ancestor took place before the appearance of the vertebrates [5]. Figure 3 lists the % and no. of identities in amino acid positions of human and rattlesnake insulin and of the corresponding A and B domains of IGF I and II. The % identity between human and rattlesnake insulin is 76 [10], the same as between IGF I and II (A and B domains). Under the assumption that the A and B domains of IGF had the same rate of evolution as A and B chains of insulin, the time of divergence of IGF I and II can thus be placed at about the time of the appearance of the first mammals.

As to the biological function of IGF, evidence is accumulating that IGF is a growth hormone-dependent hormone [11,12], thus belonging to a class of growth factors called somatomedins (cf. [13]). It remains to be seen whether other somatomedins such as somatomedin A, somatomedin C and multiplication stimulating activity (MSA) (cf. [13]) are as close homologues of proinsulin as is IGF.

Acknowledgements

We thank Lotte Braun, Christina Kasper and Hanspeter Lutz for technical assistance. This work was supported by the Swiss National Science Foundation, grant 3.708.76.

References

[1] Froesch, E. R., Bürgi, H., Ramseier, E. B., Bally, P. and Labhart, A. (1963) J. Clin. Invest. 42, 1816–1834.
[2] Oelz, O., Froesch, E. R., Bünzli, H. F., Humbel, R. E. and Ritschard, W. J. (1972) in: Handbook of Physiology; endocrinology I (Steiner, D. F. and Freinkel, N. eds) pp. 685–702, Am. Physiol. Soc., Washington, DC.
[3] Rinderknecht, E. and Humbel, R. E. (1976) Proc. Natl. Acad. Sci. USA 73, 2365–2369.
[4] Rinderknecht, E. and Humbel, R. E. (1976) Proc. Natl. Acad. Sci. USA 73, 4379–4381.
[5] Rinderknecht, E. and Humbel, R. E. (1978) J. Biol. Chem. in press.
[6] Blundell, T. L., Bedarkar, S., Rinderknecht, E. and Humbel, R. E. (1978) Proc. Natl. Acad. Sci. USA 75, 180–184.
[7] Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W. J. and Goodman, H. M. (1977) Science 196, 1313–1319.
[8] Frazier, W. A., Angeletti, R. H. and Bradshaw, R. A. (1972) Science 176, 482–488.
[9] James, R., Niall, H., Kwok, S. and Bryant-Greenwood, G. (1977) Nature 267, 544–546.
[10] Dayhoff, M. O., ed (1976) Atlas of Protein Sequence and Structure, suppl. 2, p. 129, Natl. Biomed. Res. Found., Washington, DC.
[11] Schlumpf, U., Heimann, R., Zapf, J. and Froesch, E. R. (1976) Acta Endocrinol. 81, 28–42.
[12] Zapf, J., Schoenle, E. and Froesch, E. R. (1978) Eur. J. Biochem. in press.
[13] Rechler, M. M. and Nissley, S. P. (1977) Nature 270, 665–666.