The Complexity of the IGF1 Gene Splicing, Posttranslational Modification and Bioactivity

Anastassios Philippou,1 Maria Maridaki,2 Spiros Pneumaticos,3 and Michael Koutsilieris1

1Department of Experimental Physiology, Medical School, 2Department of Sports Medicine and Biology of Physical Activity, Faculty of Physical Education and Sport Science, and 3Third Department of Orthopaedic Surgery, Medical School, National and Kapodistrian University of Athens, Athens, Greece

The insulinlike growth factor-I (IGF-I), also called somatomedin C, is a cellular and secreted growth factor which is critical for normal body growth, development and maintenance, and has important roles in multiple biological systems (1–3). A variety of cellular responses are induced by IGF-I, including cell proliferation, differentiation, migration and survival (4–8). These cellular responses have implicated IGF-I in several conditions such as the pathophysiology of several cancers (9–11), or the mitogenic and myogenic processes during muscle development, regeneration or hypertrophy, since, unlike other growth factors, IGF-I acts as both a mitogen and a differentiation factor (12,13).

IGF-I is produced by many tissues, indicating that a significant component of IGF-I action is due to its autocrine and paracrine mode of function and, thus, both common and unique or complementary pathways exist for the IGF-I isoforms to promote biological effects. The multiple peptides derived from IGF-I and the differential expression of its various transcripts in different conditions and pathologies appear to be compatible with the distinct cellular responses observed to the different IGF-I peptides and with the concept of a complex and possibly isoform-specific IGF-I bioactivity. This concept is discussed in the present review, in the context of the broad range of modifications that this growth factor undergoes which might regulate its mechanism(s) of action.

Online address: http://www.molmed.org
doi: 10.2119/molmed.2014.00011

INTRODUCTION

The insulinlike growth factor-I (IGF-I), also called somatomedin C, is a cellular and secreted growth factor which is critical for normal body growth, development and maintenance, and has important roles in multiple biological systems (1–3). A variety of cellular responses are induced by IGF-I, including cell proliferation, differentiation, migration and survival (4–8). These cellular responses have implicated IGF-I in several conditions such as the pathophysiology of several cancers (9–11), or the mitogenic and myogenic processes during muscle development, regeneration or hypertrophy, since, unlike other growth factors, IGF-I acts as both a mitogen and a differentiation factor (12,13).

IGF-I is produced by many tissues, indicating that a significant component of IGF-I action is due to its autocrine and paracrine mode of function, although it also acts as a classical circulating hormone. In the endocrine mode of action, IGF-I acts as a mediator (somatomedin C) of the growth-promoting effects of pituitary growth hormone (GH, somatotropin), which induces the synthesis and release of IGF-I by the liver (2,14,15). Circulating IGF-I is mainly derived from the liver, but also from skeletal muscle (3,16–18), and is mostly bound to high affinity IGF-binding proteins, which protect it from proteolytic degradation and modulate its bioavailability to the IGF-I receptors (2,19).

Different IGF-I mRNA transcripts are produced as a result of the alternative splicing of the IGF1 gene, encoding for several IGF-I precursor isoforms. These IGF-I protein isoforms differ by the structure of their extension peptides, or E-peptides, on the carboxy-terminal end and by the length of their amino-terminal signal peptides. However, they share the same mature peptide, which is the common part of all the IGF-I precursors (20–23). IGF-I mediates its actions through the binding and activation of several receptors and the IGF-I domain...
which is responsible for the receptor binding is the biologically active mature peptide. It is derived after the posttranslational cleavage of the pro-IGF-I (iso)forms and the removal of the E-peptides (23–26). Interestingly, it has been proposed that the E-peptides also possess bioactivity that is distinct from that of mature IGF-I (20,27).

Thus, during the last decade, many in vitro and in vivo studies have investigated the aspect of the differential IGF-I isoforms or their E-peptides actions in various conditions and pathologies (28–38). This concept was further supported by recent findings which revealed differential, E-peptide– or IGF-I isoform–specific signaling (31,33,34,38–40).

In the present review, focus has been on the propounded concept of the differential roles and bioactivity of the IGF-I isoforms or peptides, in the context of the complexity that characterizes the alternative splicing, posttranscriptional regulation and posttranslational modifications of this growth factor and which might modulate its mechanism(s) of action.

HUMAN IGF1 GENE STRUCTURE AND ALTERNATIVE SPlicing

The IGF1 gene spans a region of over 80 kb of genomic DNA located on the long arm of chromosome 12 in humans, it is a highly conserved sequence in mammals and primates (23), and contains six exons, which give rise to heterogeneous mRNA transcripts by a combination of multiple transcription initiation sites (that is, alternative leader sequences), alternative splicing and different polyadenylation signals (41,42). These multiple IGF-I transcripts code different precursor polypeptides, which also undergo posttranslational modifications (26,43,44), (Figures 1A–C).

More specifically, the different leader sequences result in two different classes of IGF-I mRNA variants: class 1 transcripts have their initiation sites on exon 1 (promoter 1), whereas class 2 transcripts use exon 2 as leader exon (promoter 2), and class 1 (exon 1 to exon 3) or class 2 (exon 2 to the exon 3) mRNA transcripts are produced by differential splicing of exons 1 and 2 to the common exon 3. Alternative splicing of exon 5 also results in different mRNA variants containing exon 5, generally defined as class B (IGF-IEb), or containing exon 6 (and excluding exon 5) defined as class A (IGF-IEa) (26,45,46), (see Figures 1A, B). A third variant, the IGF-IEc, which corresponds to IGF-IEb in rodents, also is generated by alternative splicing in the human IGF1 gene and contains both exon 5 and 6 (22), (see Figure 1C). Similar to the human IGF1 gene, multiple forms of pro-IGF-I mRNA have been described in other species in which IGF-I genomic sequences have been determined, such as the designated pro-IGF-I Ea-1, Ea-2, Ea-3 and Ea-4 in teleosts (23,47–50).

All possible combinations between promoter usage and terminal exon (5 or 6) can occur in different IGF-I transcripts (45,51,52). It has been proposed that the use of promoter 1 could be associated with the synthesis of paracrine IGF-I and may influence interactions with insulin-like growth factor binding proteins (IGF-BPs), or promote the formation of the truncated IGF-I peptide (46), (see below: IGF-I Processing, Secretion and Glycosylation). Transcripts initiating at promoter 1 are widely expressed in many tissues, whereas transcripts initiating at promoter 2 are expressed mainly in the liver (circulating forms) and kidney (53) and are thought to be more GH-dependent (23,54–58), or equally GH-responsive (59,60). However, the two promoters are probably not mutually exclusive, and GH can also stimulate the expression of tissue-specific (local) transcripts, although the existing evidence is still equivocal (48,61–65).

IGF-IEa transcript derives from the splicing pattern exon 1 or 2–3–4–6 of the IGF1 gene, which represents the main pro-IGF-I mRNA produced in liver (systemic IGF-IEa) but also in other tissues with similar exon sequence (22,66), (see Figure 1A). IGF-IEb transcript is a splice variant of exon 1 or 2–3–4–5–6. Its expression was firstly detected in the human liver (67), while it was also found to be expressed in lung carcinoma cells (20), in skeletal muscle (31,62,68) and more recently in various tissues and cells such as prostate, endometrium and lens epithelial cells (34,35,65,69), (see Figure 1B). Whether there is equivalent to human IGF-IEa splice pattern in nonhuman primates is not known (23,43,70). IGF-IEc mRNA transcript is an exon 1 or 2–3–4–5–6 splice variant (see Figure 1C), which was initially identified also in human liver, where, however, it is expressed approximately at 10% relative to the main IGF-IEa transcript (22). Structurally, its cDNA differs from the IGF-IEa variant by the presence of the first 49 base pairs from exon 5 (52 bp in rodents), It results from a splice acceptor site in the intron preceding exon 6 and, due to a reading frame shift, it gives rise to a different carboxy-terminal peptide sequence and a premature stop codon in exon 6. This transcript was named mecha-no-growth factor (MGF) since it was found to be upregulated in response to muscle stretch and/or damage (71); for review see (72). However, its expression also has been identified in various tissues such as endometrium (35), normal and cancerous prostatic cells (34), as well as in osteoblast-like osteosarcoma cells (73).

The biological significance of IGF-I splice variants is currently unknown and the physiological and molecular mechanisms that regulate their expression are unclear; however, the presence of distinct transcripts is indicative of diverse responses of cells to different stimuli (74) and they probably reflect the complexity of IGF-I actions mediated via its various isoforms (75,76).

Recent studies in humans have shown that the IGF-I splice variants are differentially transcribed in response to varying conditions and pathologies, such as exercise-induced muscle damage (31,68), endometriosis (35), and prostate (34), or cervical cancer (69), as well as in some human cell lines after hormonal treatment (65,73), while their differential transcription is possibly a function of age.
The differential expression of the IGF-I splice variants observed in various pathologies is of particular interest, as it could indicate distinct regulatory mechanisms and biological roles of the different IGF-I isoforms; however, their particular functions remain as yet unclear.

The different IGF-I mRNA transcripts encode the corresponding precursor proteins IGF-IeA, IGF-IeB and IGF-IeC (23,44). The 5′ end, by alternative splicing of exons 1, 2 and 3, encodes for the signal peptide of the IGF-I prohormone. Four different transcription start sites are present in exon 1 and their positions relative to the translational initiation codons (that is, Met-48 located in exon 1, and Met-25 and Met-22 located in exon 3) can give rise to three distinct IGF-I signal peptides from class 1 mRNAs (exon 1 to exon 3), (23,43). However, translation of mRNAs initiated at the four transcription start sites mentioned above is expected to produce signal sequences of 48 (transcription start sites 1, 2 and 3) and 25 (transcription start site 4) amino acids, since between two translation start sites contained in an mRNA, (for example, Met-25 and Met-22, located downstream of the transcription start site 4), the preference is given to the upstream site (42,44,55,66), (see Figures 1A–C). It was suggested that when translation initiates at Met-25, the nucleotide sequence corresponding to the first 21 amino acids (encoded by exon 1) of the sequence of the

(64,72). The differential expression of the IGF-I splice variants observed in various pathologies is of particular interest, as it could indicate distinct regulatory mechanisms and biological roles of the different IGF-I isoforms; however, their particular functions remain as yet unclear.

The different IGF-I mRNA transcripts encode the corresponding precursor proteins IGF-IeA, IGF-IeB and IGF-IeC (23,44). The 5′ end, by alternative splicing of exons 1, 2 and 3, encodes for the signal peptide of the IGF-I prohormone. Four different transcription start sites are present in exon 1 and their positions relative to the translational initiation codons (that is, Met-48 located in exon 1, and Met-25 and Met-22 located in exon 3) can give rise to three distinct IGF-I signal peptides from class 1 mRNAs (exon 1 to exon 3), (23,43). However, translation of mRNAs initiated at the four transcription start sites mentioned above is expected to produce signal sequences of 48 (transcription start sites 1, 2 and 3) and 25 (transcription start site 4) amino acids, since between two translation start sites contained in an mRNA, (for example, Met-25 and Met-22, located downstream of the transcription start site 4), the preference is given to the upstream site (42,44,55,66), (see Figures 1A–C). It was suggested that when translation initiates at Met-25, the nucleotide sequence corresponding to the first 21 amino acids (encoded by exon 1) of the sequence of the
48 amino acids (aa) signal peptide may have a specific function as part of the promoter 1, or may play a role in post-transcriptional regulation of IGF-I mRNA by GH (23,55,58).

From class 2 (exon 2 to exon 3) transcripts, three transcription start sites and their upstream position relative to the translational initiation codon Met-32 (located in exon 2) give rise to an IGF-I precursor polypeptide with 32 aa long signal sequence (23,42,44), (see Figures 1A–C).

The mature IGF-I peptide is coded by exons 3 (25 aa) and 4 (45 aa). The first 16 amino acids of the amino-terminal portion of the IGF-I E-peptide are coded by exon 4. Exons 5 and 6 encode, by alternative splicing, distinct portions of the E-peptide with alternative carboxy-terminal sequences that contain also distinct termination codons (22,44). Three different E-peptides have been identified in humans, encoded by three mRNA variants produced by alternative splicing of the 3' end of the pre-IGF-I mRNA. Exon 4 to exon 6 mRNA splicing encodes the Ea-peptide, which contains 35 aa. The first 16 aa, which are common in all the E-peptides, are encoded by the exon 4 and the remaining 19 are encoded by exon 6 (44,66). Splice variant of exon 4 to exon 5 yields the Eb-peptide which, apart from the 16 common aa encoded by the exon 4, contains 61 additional aa encoded by exon 5, resulting in the 77 aa long Eb-peptide (44,67). The third mRNA splice variant, which contains exon 4, only 49 bp from exon 5, and then exon 6, produces the Ec-peptide with a predicted length of 40 aa, that is, 16 aa from the exon 4, 16 aa from the exon 5 and 8 aa from the exon 6 (22,44), (see Figures 1A–C). It is noted, that the last 8 aa of Ec-peptide are encoded by exon 6, however they differ from the corresponding Ea-peptide sequence, because of a frameshift at the splice point. Ec-peptide is thought to
occur by use of a cryptic IGF 633 donor splice site, which is located 49 bp down-
stream from the 5′ end of the exon 5. When this cryptic IGF 633 donor splice
site is not used, the alternative splicing of exon 4–5 occurs, that is, the Eb pep-
tide (22,43). The predicted molecular mass and the residues of the three differ-
ent isoforms of the human IGF-I precursor polypeptide as well as of their vari-
ous forms and partial peptides (based on the amino acid sequence derived from
the various IGF-I mRNA transcripts) are summarized in Table 1.

In general, the complexity introduced by the transcriptional and splicing vari-
ants, posttranscriptional regulation and posttranslational modifications of the IGF1 gene (77), giving rise to various IGF-I isoforms, probably indicate their different biological roles under various conditions or pathologies and following different stimuli (78). The development of epitope-specific antibodies for distinguishing the different IGF-I E-peptides (20,38,79–82) can contribute to a more definitive analysis of IGF-I isoforms ex-
pression in various tissues and physio-
logical or pathophysiological conditions.

**IGF-I PROCESSING, SECRETION AND GLYCOSYLATION**

Post translational processing of IGF-I precursor protein may be a regulatory
mechanism of the IGF-I activity, as indicated by the unique processing features of IGF-I precursor polypeptides that have been described (25,78). Posttransla-
tional endoproteolysis of those polypep-
tides produces the signal, the mature and
the E-peptides (E domains). It is thought that the signal peptide at the start of a precursor is removed after facilitating the passage of the polypeptide into the en-
doplasmic reticulum and the secretory pathway, with possibly no further bio-
logical significance (23,83). The se-
quences of the signal peptides and the E-peptides are less strongly conserved compared with mature IGF-I peptide, though to a variable extent (23).
Table 1. Predicted molecular mass and residues of the three different isoforms (i.e., IGF-Ia, IGF-Ib, IGF-Ic), as well as of their various forms and partial peptides, of the human IGF-I precursor polypeptide.

<table>
<thead>
<tr>
<th>Class</th>
<th>IGF-I form</th>
<th>Molecular mass (kDa)</th>
<th>Number of residues (aa)</th>
<th>NCBI reference sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 or 2</td>
<td>Signal peptide</td>
<td>5.35 or 2.73</td>
<td>48 or 25</td>
<td>CAAD24998/AAAB52543</td>
</tr>
<tr>
<td>1 or 2</td>
<td>Signal peptide</td>
<td>3.51</td>
<td>32</td>
<td>NPC_001104754</td>
</tr>
<tr>
<td>1 or 2</td>
<td>Mature IGF-I</td>
<td>7.65</td>
<td>70</td>
<td>CAAB9054</td>
</tr>
<tr>
<td>1</td>
<td>Pre-pro-IGF-I Ea</td>
<td>17.03 or 14.41</td>
<td>153 or 130</td>
<td>CAAD24998/AAAB52543</td>
</tr>
<tr>
<td>1 or 2</td>
<td>Pre-pro-IGF-I Ea</td>
<td>15.18</td>
<td>137</td>
<td>NPC_001104754</td>
</tr>
<tr>
<td>1 or 2</td>
<td>Pre-pro-IGF-I Ea</td>
<td>11.69</td>
<td>105</td>
<td>NPC_001104754</td>
</tr>
<tr>
<td>1</td>
<td>Pre-pro-IGF-I Ea</td>
<td>4.05</td>
<td>35</td>
<td>AAAB52543/NPC_001104754</td>
</tr>
<tr>
<td>1</td>
<td>Pre-pro-IGF-I Eb</td>
<td>21.84 or 19.22</td>
<td>195 or 172</td>
<td>NPC_001104755</td>
</tr>
<tr>
<td>1</td>
<td>Pre-pro-IGF-I Eb</td>
<td>19.99</td>
<td>179</td>
<td>NPC_001104755</td>
</tr>
<tr>
<td>1</td>
<td>Pre-pro-IGF-I Ec</td>
<td>16.51</td>
<td>147</td>
<td>NPC_001104755</td>
</tr>
<tr>
<td>1</td>
<td>Pre-pro-IGF-I Ec</td>
<td>8.87</td>
<td>77</td>
<td>NPC_001104755</td>
</tr>
<tr>
<td>1</td>
<td>Pre-pro-IGF-I Ec</td>
<td>17.16 or 15.14</td>
<td>158 or 135</td>
<td>NPC_001104753/EAW97695</td>
</tr>
<tr>
<td>2</td>
<td>Pre-pro-IGF-I Ec</td>
<td>15.91</td>
<td>142</td>
<td>NPC_001104753/EAW97695</td>
</tr>
<tr>
<td>1 or 2</td>
<td>Pre-pro-IGF-I Ec</td>
<td>12.43</td>
<td>110</td>
<td>NPC_001104753/EAW97695</td>
</tr>
<tr>
<td>1 or 2</td>
<td>Pre-pro-IGF-I Ec</td>
<td>4.79</td>
<td>40</td>
<td>NPC_001104753/EAW97695</td>
</tr>
</tbody>
</table>

The mature peptide comprises four domains, that is, the B amino-terminal domain, C and A domain and D carboxy-terminal domain, of IGF-I polypeptides (I-84). In addition, two other protein products have been identified in the human brain: the tripeptide glycyl-prolyl-glutamate (GPE) corresponding to the NH2-terminal of the B domain of mature IGF-I and a truncated IGF-I form (~3N:IGF-I) that lacks the first three amino acids of the amino terminal end of mature peptide, probably due to alternate signal peptides or the combined action of some peptidases (78,85,86). Removal of the NH2-terminal tripeptide could be a mechanism for increasing the biological potency and availability of IGF-I, since the truncated ~3N:IGF-I has less affinity for IGF-binding proteins than mature IGF-I, thus, increasing its bioactivity (78), (see Figure 1A) (see below: IGF-I Receptors and Binding Proteins).

The mature IGF-I is a 70 aa long single-chain peptide and a highly conserved sequence among primate species (23,24). Cleavage of pro-IGF-I removes the carboxyl-terminal E domain and can occur at the highly conserved, unique pentabasic motif K^71^X-X-X-K^74^X-X-R^77^, X-R^74^, X-X-R^77^, X-X-R^77^. More specifically, the Arg^71^Ser^72^ bond is cleaved followed by the removal of the Arg residue by the action a carboxypeptidase (25), (see Figures 1A–C). In general, propoproteins can be processed at this specific motif, usually residing at the end of their pro regions, by propoprotein convertases (PCs) such as furin (25,87,88). Furin belongs to the subtilisin-related PCs (SPCs), a major family of endoproteolytic processing enzymes of the secretory pathway in mammals (89,90). Seven mammalian PCs have been identified, namely PC1, PC2, furin, PC4, PC5, paired basic amino acid cleaving enzyme 4 (PACE4) and PC7, and a method of prediction of the general PC-specific or furin-specific cleavage sites has been proposed (88). PCs process precursors at sites usually containing the specific consensus sequence [R/K]-Xn-[R/K], where X indicates any amino acid residue, and n, the number of spacer amino acid residues, which is 0, 2, 4 or 6 (91), (see Figures 1A–C). However, furin appears to have a more stringent specificity and preferentially recognizes sites that contain the sequence motif R-X-[R/K]-R (87), while R-X-X-R is its minimal cleavage sequence (88). Thus, apart from the cleavage site Arg^71^Ser^72^ for mature IGF-I, another furin-mediated processing at Arg^77^ has been observed in overexpression studies, which produces an extended, 76 aa long mature IGF-I (25,78), (see Figures 1A–C).

PC-mediated processing of pro-IGF-I to mature peptide has been shown to occur intracellularly (92), as expected for intracellular converstases such as furin, which are located in the secretory pathway (92,93). Nevertheless, evidence has been provided that the E domains are not cleaved intracellularly (94), and the secretion of unprocessed pro-IGF-Ia isoform, both glycosylated and nonglycosylated, has been reported (25,30,95–97), while there are potential proprotein convertases that could process pro-IGF-I extracellularly (97,98).

Conversion of pro-IGF-I to mature peptide cleaves of the E domains of the IGF-I precursors and one peptide is produced from the human Ea domain after posttranslational processing of pro-IGF-Ia isoform (23,80), (see Figure 1A). However, the Eb domain contains potential processing sites (see Figure 1B), and at least two putative peptide products of the human Eb peptide have been identified: the EBa peptide (residues 103–124 aa) with a C-terminal amide and the EBb peptide amide (residues 129–142) (20). The Ec domain also contains potential processing sites, from which one corresponds to the EBa peptide cleavage site and cleaves off the last eight residues of the Ec domain encoded by the exon 6 (26,82), (see Figures 1B, C). However, so far only the proform of IGF-Ib isoform and not the Ec peptide or other cleavage products have been detected (31,33–35,73,82).

The Ea-peptide of human IGF-I contains an N-linked glycosylation site at Asn^73^ based on the consensus sequence Asn-X-Ser/Thr, where X represents any encoded amino acid except proline (99,100), (see Figure 1A). However, it has been found that the human IGF-Ia precursor with a signal sequence of 48 aa (Met-48) is not glycosylated, implying that the greater the number of the amino acids contained in the signal peptide the lesser the extent of a glycosylation process (101). The human Eb- and Ec-peptides...
lack this cotranslational modification due to the reading frame shift resulted from the inclusion of the exon 5 in these two IGF-I isoforms (see Figures 1B, C). N-linked glycosylation involves the transfer of a lipid-linked tetradecasaccharide (GlcNAc<sub>2</sub>-Man<sub>α</sub>-Glc<sub>3</sub>) to an asparagine side chain (100,102), (see Figure 1A). It occurs in endoplasmic reticulum and the subsequent diversification of the conjugates occurs both in the endoplasmic reticulum and Golgi apparatus (100,102). Recently, it has been found that the glycosylation status of pro-IGF-I Eα does not affect its processing (97), while the secretion of a peptide similar or identical to Eα from a human B-lymphocyte cell line, in a partially glycosylated form has been reported (80). Although the deglycosylation process is expected to occur intracellularly, whether the deglycosylation of pro-IGF-I Eα or the Ea-peptide also could be clipped extracellularly is not known.

Considering the unique role of glycosylation in the protein biosynthesis process (100), it is possible that the Eα-peptide glycosylation might play a role in interactions with chaperones in the endoplasmic reticulum (78), or in regulation of the bioavailability of the different species of this IGF-I isoform (that is, pro-IGF-I Eα, mature IGF-I or Ea-peptide) (78,97,103). Thus, the existence of an N-linked glycosylation site in the Ea-peptide, which is absent in the E<sub>c</sub> and E<sub>b</sub>-peptide, might reflect a differential and specific biological action of the IGF-I Eα isoform mediated by this posttranslational modification of its Ea-peptide (25,43,97). Recently, it has been shown that the species of IGF-I produced by the IGF-I Eα isoform has a differential ability to activate IGF-IR. Glycosylated pro-IGF-I Eα is less efficient at receptor activation than pro-IGF-I and mature IGF-I (97), resembling the decreased receptor-binding affinity of pegylated IGF-I forms (104,105) and implying that glycosylated pro-IGF-I Eα may serve as a reservoir for IGF-I that can be stored until needed (97). The strong conservation observed in the sequence of the Ea domain also suggests a specific biological function for the Ea-peptide (23).

Nevertheless, whether the E-peptides are more stable and/or bioactive within their pro-IGF-I forms, or they are processed to act directly on their targets remains to be elucidated (see further discussion in next sections). Moreover, it would be essential to verify whether the IGF-I isoforms could be released in the circulation as different proforms or E-peptides (3,79), or the final peptide that enters the circulation after extracellular endoproteolysis of the IGF-I prohormone is only the mature peptide (103,106).

**IGF-I RECEPToRS AND BINDING PROTEINs**

IGF-I actions are mediated through its binding to several receptors, such as IGF-IR (or type I IGF receptor) and IGF-II R (or type II IGF receptor), insulin receptor (IR), and some atypical receptors such as the hybrid IR/IGF-IR (107–109). More specifically, the mature IGF-I peptide, which is responsible for binding to the receptors, binds IGF-IR with the highest affinity, IGF-IR with low affinity and is also able to interact with IR. The IGF-IR exhibits a high degree of homology to IR (110) and, given the significant structural similarity between IGF-I and insulin, these ligands can cross-activate both receptors, while the IGF-IR signaling pathways share multiple intracellular mediators with the insulin signaling cascade (19,52). The IGF-IR/IR hybrid receptor is thought to function predominantly as an IGF-I receptor, since its binding affinity for insulin is lower than that for IGF-I, however the functional importance of IGF-IR/IR hybrid receptor remains poorly understood (74,109,111).

It is widely recognized that most of the observed IGF-I biological effects on cell growth, differentiation, invasion and survival depend on the binding and activation of IGF-IR, which is a ligand-activated receptor tyrosine kinase (4,112). Specifically, functional epitope mapping of IGF-I has revealed that the IGF-IR interacts with the residues 21, 23, 24, 44 as well as the tyrosines 31 and 60, which are located in the C and A domains (113), (see Figure 1A). IGF-IR is a transmembrane protein consisting of two extracellular α-subunits, which contain the cysteine-rich ligand binding site, and two transmembrane β-subunits that have a cluster of three tyrosine residues, which undergo phosphorylation and activation upon IGF-I binding (2,110,114). A structural rearrangement in the transmembrane β subunits of the receptor is caused by binding of IGF-I to IGF-IR, resulting in transautophosphorylation of the cytoplasmic tyrosine kinase domain of the receptor, as one kinase domain phosphorylates the other, and thus destabilizing the autoinhibitory conformation within the kinase domain (115,116).

This conformational change permits unrestricted access to the binding sites for protein substrates (117), thus recruiting specific cytoplasmic molecules, such as insulin receptor substrate (IRS) proteins, and activating specific intracellular pathways including Ras/mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinases 1 and 2 (ERK1/2) and phosphatidylinositol 3-kinase (PI3K)/Akt (118).

Biological actions of IGF-I are modulated by a family of at least six IGFBPs (52,75,119–21), which interact mainly with the residues 1–3 and 49–51 of mature IGF-I (113), (see Figure 1A). In general, IGFBPs transport IGF-I and increase its half-life in the circulation; most of the circulating IGF-I is protected from proteolytic degradation by forming a ternary complex with IGFBP-3 and the glycoprotein acid-labile subunit (ALS) (74,122). IGFBPs also would be expected to modulate and control, both in the circulation and in the extracellular environment, the extent of IGF-dependent cellular effects via regulation of free IGF-I concentration and its local bioavailability in the tissue, since IGFBPs provide tissue specificity for the local action of IGF-I (74,119,123–126). In addition, IGFBPs compete with IGF-IR and normally have higher binding affinity to IGF-I than IGF-IR does. Therefore, binding of IGFBPs to IGF-I prevents the ligand from interacting with the receptor and, thus, suppresses IGF-I actions (74,124,125).
However, some IGFBPs can exhibit IGF-I potentiating effects and their IGF-I-inhibitory or stimulatory activities are dictated by factors such as the tissue-specific distribution of particular IGFBPs and the ratio between free (active) IGF-I and IGFBP-IGF-I bound (4,11). Moreover, it has been shown that certain IGFBPs have IGF-independent activities, implying that they can modulate cell survival and apoptosis, or inhibit tumor growth in the absence of the ligand (11,127). In addition, proteolytic fragments of IGFBP-3 were reported to possess mitogenic activity in the peritoneal fluid of women with endometriosis (128).

There is also a group of cysteine-rich proteins, known as IGFBP-related proteins (IGFBP-rPs), that share important structural similarities with the IGFBPs but they have low binding affinity to IGFs. It has been proposed that these proteins and the IGFBPs constitute an IGFBP superfamily (129,130), however the functions of the IGFBP-rPs regarding the IGFs actions are as yet unclear (131).

**IGF-I PEPTIDES ACTIONS AND SIGNALING**

Although, by the general consensus, IGF-I is thought to exert its biological actions predominantly through mature peptide, differential biological activities have been reported for the different IGF-I isoforms (propeptides), or for their E-peptides, exogenously administered or overexpressed in various in vivo (28,29,36,132–134) and in vitro models (20,27,31–33,135,136), implying that there are peptides other than the IGF-I ligand that also possess bioactivity and, thus, both common and unique or complementary pathways exist for the IGF-I isoforms to promote biological effects (31,36).

A differential expression profile of the IGF-I isoforms have been shown in various conditions or pathologies in humans, such as skeletal muscle damage (31,68), endometriosis (35) or prostate (34), cervical (69) and colorectal cancer (137). The diversity and the patterns of differential expression have been proposed to reflect potential biological activities associated with the E domain peptides (27,52). Divergent actions and signaling of the different pro-IGF-I forms or mature IGF-I lacking any E-peptide have been reported after viral-mediated expression of the IGF-I isoforms (IGF-IA and IGF-IB) in mouse skeletal muscle (36,132). Specifically, it has been shown that overexpression of mature IGF-I in skeletal muscle did not promote muscle hypertrophy in young mice, suggesting that the pro-IGF-I forms are required for this effect and that E-peptides, either as a part of pro-IGF-I or independent of mature IGF-I, may be necessary for IGF-I-mediated muscle hypertrophy (36). Moreover, by using the MKR transgenic mouse model where functional IGF-IRs are lacking in muscle fibers (as a dominant negative IGF-IR is expressed specifically in skeletal muscle), it was shown that, regardless of which isoform is overexpressed, IGF-I receptors on muscle fibers are required for IGF-I-mediated hypertrophy (36). Interestingly, however, after viral-mediated delivery of murine IGF-I isoforms into skeletal muscle, even though both isoforms caused increased phosphorylation of the IGF-IR, increased expression of IGF-IB (murine IGF-IBe) drove both main pathways downstream of IGF-IR, that is, the PI3K/Akt pathway and the MAPK pathway, whereas IGF-IA (murine IGF-IEa) overexpression resulted in increased Akt phosphorylation only (132). These findings imply distinguishing IGF-I isoform-specific actions, regardless of any potential receptor(s) activated (see also discussion in following sections). In addition, constitutively overexpression of the IGF-IEa isoform specifically in cardiac muscle was shown to protect the heart from oxidative stress via SirT1 and mammalian target of rapamycin (mTOR) signaling pathway was not induced in the transgenic mice overexpressing IGF-IEa (138,140). Instead, this specific IGF-I propeptide activated alternate signaling intermediates 3-phosphoinositide-dependent protein kinase-1 (PDK1) and serum/glucocorticoid regulated kinase 1 (SGK1), as well as SirT1. Thus, it was suggested that this downstream of IGF-I receptor(s) signaling activated by IGF-IEa employs novel pathways and that the divergent signaling mechanisms between the two IGF-I forms (that is, mature IGF-I and IGF-IEa propeptide) may account for their opposing effects on the heart (136,138,140).

Considering particularly the bioactivity of the E-peptides, in vitro studies have suggested that the E-peptides of the human IGF-I precursors may act as independent growth factors, since their synthetic analogs, generated from unique regions within the E domains, were demonstrated to possess mitogenic (20,27,34,35,73,141), angiogenic (142) and migratory activity (28,141,143), and regulate cell differentiation (27,28) in various human cells or cell lines. Antitumor activity of human Eb-peptide also has been reported in some cancer cells (144).

Similar to the results derived from human cell lines, studies using animal cell lines models have shown that exogenous administration or overexpression of synthetic peptides, generated from different regions within the E domain of human (Ec) (28,31,33,145,146), trout (Ea) (144,147) or rodent (Ea and Eb) sequences (38) in human, rodent or porcine cells in culture showed unique though inconsistent effects in promoting cell proliferation and migration, and in delaying or inhibiting cell differentiation. In particular, actions associated with peptides derived from mammalian Ea domain have only recently been established, with rodent Ea-peptide reported to possess bioactivity (38). Furthermore, it was shown that murine Ea- and Eb-peptides increase cell entry of IGF-I from the media, providing evidence that, in addition to having independent activity, they may modulate IGF-I (30).
The differential biological effects of the synthetic Ec-peptide compared with mature IGF-I peptide, such as cell proliferation versus differentiation, and the lack of suppression of the synthetic E-peptide bioactivity after blocking (mature) IGF-I signaling with IGF-IR neutralizing antibodies, makes it tempting to postulate that the Ec-peptide acts via a different receptor (28,135,145). However, concerns have been raised about the effectiveness of the IGF-IR neutralizing antibodies to block IGF-I signaling, since they could internalize and, in this way, even activate the receptor, or change its localization, thus facilitating an E-peptide action (38). Nevertheless, the mature IGF-I bioactivity appeared to be suppressed in those cells where IGF-IR neutralizing antibodies were used.

Interestingly, an autonomous, IGF-IR-independent bioactivity has been reported for specific regions of the human Eb domain of IGF-I; the synthetic analog of the EB1 peptide (see above: IGF-I Processing, Secretion and Glycosylation) was found not only to possess mitogenic activity in human bronchial epithelial cells but also to bind to specific high-affinity receptors on those cells. Furthermore, neither ligand binding was inhibited by recombinant IGF-I or recombinant insulin, nor did a monoclonal antibody antagonist to the IGF-IR suppress the proliferative response induced by the synthetic EB1 peptide. It was suggested that IB is a growth factor that mediates its effect through a specific receptor (20). Similarly, in accordance with the well established notion that the molecular action of a bioactive peptide initiates through its specific binding to a cell surface receptor, it was demonstrated that the human Eb-peptide binds to common cell surface molecules on human neuroblastoma cells (27). Again, the saturation of mature IGF-I or insulin did not displace the binding of the human Eb-peptide, suggesting the existence of distinct putative receptor components on those cells (148).

Further evidence for distinct, IGF-IR-independent bioactivity of the human Ec domain was provided from its divergent signaling compared with mature IGF-I. Our group (31,33,34,65) and others (39) have shown that a synthetic analog of the human Ec peptide possesses distinct signaling compared with IGF-IR ligand. Specifically, to distinguish any unique biological effect of the Ec domain from a potential bioactivity possessed by the common part (that is, the first 16 residues) of the IGF-I E domains, we utilized a synthetic Ec-peptide corresponding to the region beyond that common sequence, that is, a peptide similar to the C-terminal 24-residues, of which 16 are encoded by exon 5 and the last 8 by exon 6. It was documented that this sequence of the human Ec domain possesses distinct signaling, in contrast to mature IGF-I, it activates only ERK1/2 and not Akt (31,33,34,65). Moreover, the selective activation of only one of the two main signaling pathways downstream of IGF-I/IGF-IR and the IGF-IR- and IR-independent bioactivity of this synthetic part of the Ec domain was further documented by using the siRNA knock-out model in various human cell lines (34,35,73).

More recently, synthetic E-peptides corresponding to the rodent Ea and Eb domain sequences were utilized in a kinase receptor activation (KIRA) assay to test IGF-I-dependent and -independent activation of IGF-IR by these E-peptides, showing that they do not directly induce IGF-IR phosphorylation in mouse fibroblasts (P6 cells) (38). Interestingly, however, the presence of either of those E-peptides increased IGF-IR activation by IGF-I in the murine C2C12 cell line, suggesting that they may modulate IGF-I activity (38). Moreover, from the findings of that model, it was concluded that E-peptides signaling, as well as mitogenic and motogenic effects are dependent upon a functional IGF-IR, and their activity plausibly reflects actions of pro-IGF-I (38). Further evidence has been provided recently, suggesting that IGF-I splice variants may exert their actions through mature IGF-I and not the E-peptides (149), or by supporting the bioactivity of pro-IGF-I forms (97). Furthermore, there is controversial evidence regarding the role of proforms (103,136,149) or E-peptides (38) of the murine IGF-I isoforms, particularly in cell differentiation. Whether it reflects a differential mode of action of the free E-peptides compared with the pro-IGF-I isoforms, or potential differences between their exogenous administration and overexpression (or natural endogenous production), possibly should be addressed.

Collectively, the in vitro models utilizing the sequences of the rodent IGF-I E domains have provided specific information regarding the bioactivity and the mode of action of those E-peptides in murine cell lines. The findings of these studies suggest that the E-peptides have little or no independent activity and, instead, they modulate mature IGF-I (ligand) activity and signaling (see review in [150]). Nevertheless, since these domains are very variable and much less conserved, it has been indicated that the species specificity must clearly be taken into account when assessing the activity of the human IGF-I Eb and Ec domains, from which peptides with important biological activities have been reported (23). Thus, it remains to be further elucidated whether the autonomous, IGF-IR-independent and IR-independent bioactivity, and the specific high-affinity receptor binding observed for human Eb- and Ec-peptides in various human cell lines (20,34,35,73,148) reflect an alternative, species-specific ligand/receptor mechanism of action for these human E domains. Species-specific models, in terms of utilizing peptide sequences that correspond to the IGF-I E domains of the species or the cells used in the model, are of particular importance to investigate physiological mechanisms of actions, such as a ligand/receptor mechanism.

Moreover, it is still an elusive fact where the E-peptides signaling diverges from that of mature IGF-I (given the distinct activation of ERK1/2 but not Akt pathway by the E-peptide). It has been proposed that the E-peptides may increase the phosphorylation of ERK1/2 by increasing the IGF-IR internalization,
which would act in favor of the MAPK arm of the IGF-IR downstream signaling (38). However, since the possibility that the E-peptides also activate ERK1/2 through an IGF-IR independent mechanism cannot be excluded, it remains to be further elucidated whether they affect the ERK1/2 pathway only at the level of the IGF-IR per se (that is, possibly by increasing its internalization, or by affecting its conformational change after its binding to IGF-I and, thus, activating exclusively a specific pathway) or if the E-peptide–induced activation of ERK1/2 occurs at a level downstream of the IGF-IR, or even upstream, in that the E-peptide utilizes another “receptor” or cellular uptake mechanism, given the activation of ERK1/2 after exogenous administration of the E-peptide. The intracellular signaling pathways initiated by IGF-IR ligation may interact with signaling via G protein-coupled receptors or other mediators, modulating some responses (151). However, the existence of such a putative, canonical or noncanonical receptor or internalization mechanism for the E-peptide remains to be determined and characterized. Alternatively, the exclusive phosphorylation of ERK1/2 has been proposed to be a result of a possible tuning of the IGF-IR signaling cascade by the E-peptide toward MAPK (38).

The identification of the signaling pathways and effectors upstream and downstream of the E-peptide–induced activation of the extracellular-regulated kinases ERK1 and ERK2, the best-characterized members of the MAP kinase family (152), would provide information of particular importance regarding the mechanisms of E-peptide bioactivity. Furthermore, it is tempting to speculate that the selective activation of MAPK/ERK1/2 pathway up to a critical level by the E-peptide might further exert an inhibitory regulation of its competitive PI3K/Akt pathway, since it has been shown that the activity of one pathway might inactivate portions of the other (153,154), thus consisting of a regulatory mechanism of competitive biologic actions such as cell proliferation and differentiation.

It should be also mentioned that, given the subcellular as well as differential localization of the E-peptides (94,141,155,156), an intracellular signaling mechanism that mediates E-peptides bioactivity, as proposed for some peptide growth factors (157), cannot be excluded.

CONCLUSION
IGF-I regulation occurs at multiple levels due to the alternative splicing of its pre-mRNA both at the 5′ and 3′ end, and the isoform-specific co- and posttranslational modifications, which appear to play key roles in modulating the bioavailability and bioactivity of this growth factor. New evidence supports the concept of IGF-I isoforms— or IGF-I various peptides—specific functions, possibly via specific and divergent signaling. Nevertheless, it remains to be determined where the distinct E-peptide signaling diverges from that of mature IGF-I, or whether this signaling is induced uniquely and autonomously by the E-peptides or synergistically with the mature IGF-I bioactive molecule, and in the form of processed or pro-IGF-I molecules. Thus, the regulatory mechanism(s) of a potentially specific and apportioned bioactivity of the various IGF-I peptides is of particular interest within the context of a revisited characterization of the multiple IGF-I actions.

DISCLOSURE
The authors declare that they have no competing interests as defined by Molecular Medicine, or other interests that might be perceived to influence the results and discussion reported in this paper.

REFERENCES
20. Siegfried IM, et al. (1992) A mitogenic peptide amide encoded within the E peptide domain of


vated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. EMBO J. 16:5572–81.