Hypothesis

Amyloidogenic determinant as a substrate recognition motif of insulin-degrading enzyme

Igor V. Kurochkin*

Chugai Research Institute for Molecular Medicine, Inc., 153-2 Nagai, Niihari, Ibaraki 300-4101, Japan

Received 9 March 1998

Abstract Insulin-degrading enzyme (IDE) is an evolutionarily conserved neutral thiol metalloprotease expressed in all mammalian tissues whose biological role is not well established. IDE has highly selective substrate specificity. It degrades insulin, glucagon, atrial natriuretic peptide, transforming growth factor α but does not act on related hormones and growth factors. The structural properties determining whether a peptide is an IDE substrate are essentially unknown. The reported cleavage sites are not consistent with simple peptide-bond recognition and it was proposed that IDE recognizes in its substrates some elements of tertiary structure. We noticed that although IDE substrates are functionally unrelated, the majority of them share a specific property, an ability to form under certain conditions amyloid fibrils. Utilizing the residue pattern recognition procedure, this study reveals a common motif in the sequences of IDE substrates, HNHHHPSH, where H is wholly or partly hydrophobic character, N is small and neutral, P is polar, and S is polar and/or small amino acid residue. It is proposed that this sequence motif predetermines a structure recognized by IDE. The identified motif appears to be essentially the same as the proposed earlier consensus sequence for amyloid-forming peptides [Turnell and Finch, J. Mol. Biol. 227 (1992) 1205–1223]. The study suggests that IDE may play a role in elimination of potentially toxic amyloidogenic peptides.

Key words: Insulin-degrading enzyme; Amyloid; Insulin; Protein structure; Alzheimer's disease

1. Introduction

The insulin-degrading enzyme (IDE, EC 3.4.24.56) is an evolutionarily conserved non-lysosomal thiol metalloendopeptidase, also referred to as insulinase. IDE is a member of a newly recognized family of metalloendopeptidases, the pitrilysins, which are distinct from classical metalloproteinases in that they contain an unusual Zn^{2+} binding site [1,2]. This family also includes *Escherichia coli* protease III (pitrilysin) and *Drosophila* IDE, both of which share a high degree of homology to mammalian IDEs [3,4] and are distantly related to IDE mitochondrial [5] and chloroplast [6] processing enzymes. IDE is expressed in all mammalian tissues [7], however, its physiological role remains unknown. It was postulated almost 50 years ago but not firmly established yet that the enzyme is responsible for proteolysis of insulin, thus terminating its biological action (reviewed in [8]). IDE is developmentally regulated [7] and is apparently involved in the morphological and biochemical differentiation of sperm cells [9] and cultured myoblasts [10].

The remarkable feature of IDE is its highly restricted substrate specificity. The enzyme degrades with high affinity several polypeptide hormones including insulin [11], atrial natriuretic peptide (ANP) [12], transforming growth factor α [13] and glucagon [14]. On the other hand, IDE does not act on several other related peptides, such as glucagon-like peptide 1, glucagon-(19–29), epidermal growth factor, platelet-derived growth factor, nerve growth factor, and vasoactive intestinal peptide [15]. Highly similar to ANP in structure, brain natriuretic peptide and C-type natriuretic peptide are very poor IDE substrates [16]. Analysis of cleavage sites in IDE substrates has revealed no specificity for a particular peptide bond, although some preference for large hydrophobic or basic amino acids on the carboxyl side of the cleavage site has been noticed [8]. In order to extend the number of potential IDE substrates, Werlen at al. prepared different peptides by tryptic digestion of BSA and CNBr degradation of cytochrome *c* [17]. Only a few peptides could serve as IDE substrates and again, as in the case of natural substrates, the specificity of IDE was difficult to categorize. It was demonstrated that metal chelators could inhibit degradation but not binding of the substrates [16,18]. Mutations that abolished catalytic activity of IDE did not prevent binding of insulin [19]. These results are consistent with a model whereby IDE first binds to the substrate recognition site prior to acting as a catalyst. Thus, the specificity of the enzyme is realized at the stage of substrate binding. Amino acid sequence comparisons have not revealed any significant similarity between different IDE substrates, suggesting that the enzyme recognizes some unidentified higher order structure. This idea is supported by the observation that the IDE cleavages are generally grouped on the same side of substrate molecules when viewing the three-dimensional structure [20,21]. In this report, it is proposed that IDE specificity is based on recognition of the structural elements that predetermine the ability of peptides to form amyloid fibrils. Furthermore, it is demonstrated that IDE substrates have in their sequence a common motif found previously in amyloid-forming peptides [22].

2. Hypothesis

Although peptide substrates of IDE are functionally unrelated, it is of interest to note that many of them share an ability to form under certain conditions amyloid fibrils. Indeed, amyloid deposits were found at sites of repeated insulin injections in diabetic patients [23,24]. In the single case where the amyloid fibril protein was extracted and analyzed, it was

*Fax: (81) (298) 30-6270.
E-mail: igork@tk.chugai-pharm.co.jp

PII S0014-5793(98)00422-0

shown to contain intact insulin molecule [24]. Another high-affinity IDE substrate, ANP, forms amyloids in the atria of 83% of persons aged 80 and over [25,26]. One study demonstrated that IDE is able to degrade calcitonin [27]. Calcitonin is a major fibril protein in amyloid deposits associated with medullary carcinoma of the thyroid [28]. We demonstrated previously that the amyloid β peptide, the principal component of amyloid deposits in Alzheimer's disease brain, could be specifically cross-linked to IDE in rat brain extracts and was efficiently degraded by the purified enzyme [29]. In addition, insulin, ANP, calcitonin, amyloid β peptide and glucagon can readily form amyloid fibrils under variety of conditions in vitro (reviewed in [30]). These facts taken together imply that the peptides having amyloidogenic potential are somehow recognized by IDE. Thus, it would be not unreasonable to propose that the amyloidogenic determinant and IDE substrate recognition motif are similar if not identical.

Although the primary structures of amyloid proteins differ widely, inherent amyloidogenic amino acid sequences probably exist. This would explain why only some polypeptides give rise to amyloid fibrils. Notwithstanding different solution structures in the monomer state, amyloidogenic proteins assume a remarkably similar structure when in amyloid fibrils, which is a β-pleated sheet conformation [30]. Drawing a parallel with the IDE recognition mechanism, it is speculated that, although IDE substrates do not share common structural properties, they nevertheless may adopt a similar conformation upon binding to the enzyme. Several studies have demonstrated that polypeptides, IDE substrates, display a high degree of conformational flexibility and can readily assume β-sheet conformation upon self-association and/or when bound to lipid. For example, ANP in aqueous solution has a random structure [31]. However, binding to bilayer vesicles of acidic phospholipids results in adoption of a highly ordered structure of β-sheets. A transition to a similar β-structure occurs upon self-association of the peptide [31]. Amyloid β peptides (1–28) and (1–42) in solution are mixtures of α-helices, β-sheets and random coil structures with the conversion to a β-sheet conformation upon peptide self-association [32]. Another IDE substrate, glucagon, has extended flexible conformation in aqueous solution [33] with transition into an all-β-sheet conformation upon self-aggregation. Unlike these smaller peptides, insulin, a globular protein with stably folded elements of secondary structure, dimerizes through β-strands preexisting in the monomer [34].

As noted above, amyloidogenic peptides do not share significant homology at the level of amino acid sequence. Nevertheless, the residue pattern recognition procedure developed by Turnell [35] allowed him and Finch to satisfactorily align sequences of different amyloid-forming peptides [22]. In this approach, the alignment was performed on the basis of similar amino acid residue types, e.g. polar versus polar etc., and their similar relative positions and proportions. The common patterns revealed in this way imply structural rather than simply sequence homology between peptides. Implementation of these rules for peptides distant in primary structure, such as insulin and amyloid β peptide, revealed that the sequence of residues 21–30 along the insulin B-chain has homology with residues 16–25 of amyloid β peptide ([22], Table 1). Notably, the residues within the common pattern, namely residues 24B–26B of insulin, are involved in insulin dimerization and binding Congo red dye [22], a specific stain for amyloid deposits. The residues 17–20 of amyloid β peptide are crucial for the formation of β-sheet structure and the amyloid properties of

Table 1
Multiple sequence alignment of IDE peptide substrate fragments

Peptide	Amyloidogenicity	Segment	Sequence
Amyloid β peptide	+	14 - 28	H Q K L V F F A E D V G S N K | | * | | | : + | | + + |
Insulin	+	19B - 30B	C G E R G F F Y T P K A | | * | | | : + | |
Glucagon	+	8 - 20	S D Y S K Y L D S R R A Q | * | | | : + | | + +
Cytochrome C	ND	42 - 52	K T G Q A P G F T Y T D A N K N K G I T W K E E T L M | | * | | | : + | + |
Cytochrome C	ND	82 - 94	I F A G I K K K T E R E D L I A Y L K K A T N E | | * | | | : + |
Calcitonin	+	1 - 25	C S N L S T C V L S A Y W K D L N N Y H R F S G M | | * | | | : + | +
Amylin¶	+	20 - 35	S N N F G A I L S S T N V G S N | * | | | : + |

Peptide segments shown are derived from the amino acid sequences of human amyloid β peptide, porcine insulin, human glucagon, horse cytochrome *c*, human calcitonin and human amylin. Alignment of sequences was performed according to the residue pattern recognition procedure [22,35]. Amino acid residues of the same type are marked as (|) wholly or partly hydrophobic; (*) small and neutral; (:) polar; (+) polar, and/or small [22]. The residue hydrophobicities are estimated according to Eisenberg and McLachlan [44]. In this method, the hydrophobic character of a given amino acid residue is not determined by simply a single number, whole residue hydrophobicity, but combines calculations both of the solvation free energy and of the solvent-accessible surface of each atom of the residue. As a result, amino acid residues such as Arg and Lys, having both polar and apolar parts, are regarded as partly of hydrophobic character. This is consistent with the experimental data indicating that the apolar parts of these residues make a significant contribution to the stability of proteins [44].

¶ Amylin binds to but is not degraded by IDE (see text).

ND, not determined.

Fig. 1. Alignment of the amino acid sequence of glucagon with that of glucagon-like peptide 1. The residues that do not fit into a common pattern of the residue types defined in Table 1 are black boxed.

the peptide [36]. Remarkably, applying the same procedure it was possible to align sequence segments of other IDE peptide substrates – glucagon, calcitonin and two cytochrome *c* fragments (Table 1). Two other IDE substrates, ANP and BSA fragment (503–518-S-S-543–571), do not fit into a common pattern. The conformation of these peptides is constrained by the intrachain S-S bond and, therefore, the rules for the alignment of linear sequences might not be applicable in their case. Alignment of peptide sequences provided in Table 1 indicates that IDE substrates have in their sequence a common motif that, at its core, has a consensus sequence of HNHHHPSH, where H is wholly or partly hydrophobic character, N is small and neutral, P is polar, and S is polar and/or small amino acid residue. The distribution of hydrophobic residues within the motif is characteristic both of amphipathic, α-helical and of β-strand within a β-sheet [22].

That the defined motif serves as a recognition determinant for IDE is consisted with several experimental data. Affholter et al., using a variety of insulin mutants, have established that the residues B24–B26 which are within the motif (Table 1) are critical for the high binding affinity of insulin to IDE [37]. Insulin analogues substituted, for example, at position B25 (Phe) with the non-aromatic non-hydrophobic amino acids Asp or His resulted in respective 16- and 20-fold decreases in affinity to IDE [37]. From the inhibition of IDE activity by different fragments of amyloid β peptide, it appears that the main interaction site lies in the region of residues 17–24 of the peptide [38], the same residues that fit into the common pattern (Table 1). Again as in case of insulin, the residues of amyloid β peptide involved in binding to IDE are crucial for the amyloid properties of the peptide [36]. Also, competition studies demonstrated that while glucagon was a potent inhibitor of insulin binding to IDE, glucagon-(19–29) and glucagon-like peptide-1 were without effect [15]. Residues 19–29 of glucagon lie outside the proposed consensus sequence and in the case of glucagon-like peptide-1, the consensus is disrupted at several positions (Fig. 1).

While it now seems clear that IDE substrates do share a common motif, that alone cannot explain fully the substrate specificity of the enzyme. For instance, although insulin, glucagon and amyloid β peptide appear to share a common motif, the affinity of IDE for these substrates differs significantly. While IDE binds to insulin with high affinity ($K_m = 0.1$ μM) [13], the affinities for glucagon and amyloid β peptide are relatively low ($K_m = 5.3$ μM and 2.7 μM, respectively [14,38]). Therefore, it is likely that the residues outside of proposed motif might participate in modulation of substrate binding.

3. Conclusions

Analysis of substrate specificity of IDE revealed that the enzyme cleaves peptides, which share an ability to form amyloid fibrils in vivo and/or in vitro. We postulated, therefore, that the structural features predetermining the amyloidogenic property of the peptides would serve as the substrate recognition determinant for IDE. A common pattern of residue types proposed earlier for amyloid-forming peptides [22] was found in sequence segments of the majority of known IDE substrates. Remarkably, existing experimental data indicate that it is the amyloidogenic part of the IDE peptide substrates that is responsible for their binding to the enzyme [37,38]. The lack of a specific cleavage motif and negligible homology between IDE substrates at the level of the amino acid sequence imply that the higher order structure plays a role in substrate specificity. Precedence for proteolytic enzymes that appear to recognize a structural motif rather than a substrate-specific amino acid sequence does exist. Mitochondrial processing peptidase cleaves the signal sequence from a variety of mitochondrial precursor proteins yet there is no obvious cleavage motif. Recent studies established the importance of a particular structure in the vicinity of the processing site for signal peptide recognition by the protease [39]. Analysis of substrate specificity of the endopeptidase magaininase from *Xenopus laevis* revealed that the enzyme recognizes a particular substrate conformation [40]. As in the case of IDE, *Xenopus* magaininase acts on peptides with negligible amino acid identity yet experiments with numerous synthetic peptide analogues have established that all the peptidase substrates share an ability to adopt an amphipathic, α-helical motif [40]. Interestingly, *Xenopus* magaininase has the same inhibition profile and molecular mass (about 110 kDa) as IDE [40]. Determination of the primary sequence of this endopeptidase would make it possible to establish its relationship with IDE.

The recognition mechanism of IDE proposed here is verifiable. If correct, experiments should demonstrate diminished affinity to IDE and compromised ability to form amyloid fibrils of substrate analogues with the substitutions within the HNHHHPSH motif. The proposed mechanism predicts that any peptide having amyloidogenic potential and sharing the recognition motif would bind to the protease. Recently we tested whether amyloid-forming peptide amylin (Table 1) could serve as a substrate for IDE. Although this peptide was found to resist cleavage, it effectively blocked insulin-degrading activity of IDE in a competitive manner [41].

Despite numerous studies on IDE, the biological role of the protease remains to be determined. Cytosolic/peroxisomal localization of IDE is not consistent with its involvement in degradation of internalized insulin located within the endosomes. The fact that the highest levels of IDE are found in testes, tongue and brain [7,9], while the major insulin metabolizing tissues are liver, kidney and muscle, suggests that IDE may have other functions beside degrading insulin. The presence of IDE in all tissues tested and the high evolutionary conservation of the enzyme imply its involvement in performing some kind of housekeeping function. We would like to propose here that IDE might function as a scavenger of potentially toxic protein fragments prone to aggregation and amyloid formation. As cellular proteins are being constantly degraded, it is inevitable that some intermediate proteolytic products with high aggregation potential would appear. In this respect, it is interesting to note that IDE was found to be associated with the proteasome [42], a major cellular proteolytic complex, whose generated products range in size from 4 to 25 residues [43].

Finally, this work may provide a strategy for the design of specific peptide-based inhibitors of IDE, which will be useful in elucidating the biological role of the enzyme.

References

[1] Rawlings, N.D. and Barrett, A.J. (1991) Biochem. J. 275, 389–391.
[2] Becker, A.B. and Roth, R.A. (1992) Proc. Natl. Acad. Sci. USA 89, 3835–3839.
[3] Affholter, J.A., Fried, V.A. and Roth, R.A. (1988) Science 242, 1415–1417.
[4] Kuo, W.-L., Gehm, B.D. and Rosner, M.R. (1990) Mol. Endocrinol. 4, 1580–1591.
[5] Paces, V., Rosenberg, L.E., Fenton, W.A. and Kalousek, F. (1993) Proc. Natl. Acad. Sci. USA 90, 5355–5358.
[6] Vere, P.S.V., Bennett, T.M., Oblong, J.E. and Lamppa, G.K. (1995) Proc. Natl. Acad. Sci. USA 92, 7177–7181.
[7] Kuo, W.-L., Montag, A.G. and Rosner, M.R. (1993) Endocrinology 132, 604–611.
[8] Authier, F., Posner, B.I. and Bergeron, J.J.M. (1996) Clin. Invest. Med. 19, 149–160.
[9] Baumeister, H., Müller, D., Rehbein, M. and Richter, D. (1993) FEBS Lett. 317, 250–254.
[10] Kayalar, C. and Wong, W.T. (1989) J. Biol. Chem. 264, 8928–8934.
[11] Shii, K., Yokono, K., Baba, S. and Roth, R.A. (1986) Diabetes 35, 675–683.
[12] Müller, D., Baumeister, H., Buck, F. and Richter, D. (1991) Eur. J. Biochem. 202, 285–292.
[13] Garcia, J.V., Gehm, B.D. and Rosner, M.R. (1989) J. Cell. Biol. 109, 1301–1307.
[14] Rose, K., Savoy, L.-A., Muir, A.V., Davies, J.G., Offord, R.E. and Turcatti, G. (1988) 256, 847–851.
[15] Authier, F., Bergeron, J.J.M., Ou, W.-J., Rachubinsky, R.A., Posner, B.I. and Walton, P.A. (1995) Proc. Natl. Acad. Sci. USA 92, 3859–3863.
[16] Müller, D., Schulze, C., Baumeister, H., Buck, F. and Richter, D. (1992) Biochemistry 31, 11138–11143.
[17] Werlen, R.C., Offord, R.E. and Rose, K. (1994) Biochem. J. 302, 907–911.
[18] Ogawa, W., Shii, K., Yonezava, K., Baba, S. and Yokono, K. (1992) 267, 1310-1316.
[19] Perlman, R.K. and Rosner, M.R. (1994) J. Biol. Chem. 269, 33140–33145.
[20] Duckworth, W.C., Hamel, F.G., Peavy, D.E., Liepnieks, J.J., Ryan, M.P., Hermodson, M.A. and Frank, B.H. (1988) J. Biol. Chem. 263, 1826–1833.
[21] Hamel, F.G., Gehm, B.D., Rosner, M.R. and Duckworth, W.C. (1997) Biochim. Biophys. Acta 1338, 207–214.
[22] Turnell, W.G. and Finch, J.T. (1992) J. Mol. Biol. 227, 1205–1223.
[23] Störkel, S., Schneider, H.-M., Müntefering, H. and Kashiwagi, S. (1983) Lab. Invest. 48, 108–111.
[24] Dische, F.E., Wernstedt, C., Westermark, G.T., Westermark, P., Pepys, M.B., Rennie, J.A., Gilbey, S.G. and Watkins, P.J. (1988) Diabetologia 31, 158–161.
[25] Cornwell III, G.G., Murdoch, W.L., Kyle, R.A., Westermark, P. and Pitkänen, P. (1983) Am. J. Med. 75, 618–623.
[26] Johansson, B., Wernstedt, C. and Westermark, P. (1987) Biochem. Biophys. Res. Commun. 148, 1087–1092.
[27] Kirschner, R.J. and Goldberg, A.L. (1983) J. Biol. Chem. 258, 967–976.
[28] Sletten, K., Westermark, P. and Natvig, J.B. (1976) J. Exp. Med. 143, 993–998.
[29] Kurochkin, I.V. and Goto, S. (1994) FEBS Lett. 345, 33–37.
[30] Sipe, J.D. (1992) Annu. Rev. Biochem. 61, 947–975.
[31] Surewicz, W.K., Mantsch, H.H., Stahl, G.L. and Epand, R.M. (1987) Proc. Natl. Acad. Sci. USA 84, 7028–7030.
[32] Barrow, C.J., Yasuda, A., Kenny, P.T. and Zagorski, M.G. (1992) J. Mol. Biol. 225, 1075–1093.
[33] Boesch, C., Bundi, A., Oppliger, M. and Wüthrich, K. (1978) Eur. J. Biochem. 91, 209–214.
[34] Badger, J., Harris, M.R., Reynolds, C.D., Evans, A.C., Dodson, E.J., Dodson, G.G. and North, A.C.T. (1991) Acta Crystallogr. Sect. B 47, 127–136.
[35] Turnell, W.G. (1988) in: Methodological Surveys in Biochemistry and Analysis (Reid, E., Cook, G.M.W. and Luzio, J.P., Eds.), Vol. 19, pp. 9–101, Royal Society of Chemists, Cambridge.
[36] Hilbich, C., Kisters-Woike, B., Reed, J., Masters, C.L. and Beyreuther, K. (1992) J. Mol. Biol. 228, 460–473.
[37] Affholter, J.A., Cascieri, M.A., Bayne, M.L., Brange, J., Casaretto, M. and Roth, R.A. (1990) Biochemistry 29, 7727–7733.
[38] McDermott, J.R. and Gibson, A.M. (1997) Neurochem. Res. 22, 49–56.
[39] Waltner, M. and Weiner, H. (1995) J. Biol. Chem. 270, 26311–26317.
[40] Resnick, N.M., Maloy, W.L., Guy, H.R. and Zasloff, M. (1991) Cell 66, 541–554.
[41] Kurochkin, I.V. and Turnell, W.G. (1997) FASEB J. 11, A1276.
[42] Duckworth, W.C., Bennett, R.G. and Hamel, F.G. (1994) J. Biol. Chem. 269, 24575–24580.
[43] Coux, O., Tanaka, K. and Goldberg, A.L. (1996) Annu. Rev. Biochem. 65, 801–847.
[44] Eisenberg, D. and McLachlan, A.D. (1986) Nature 319, 199–203.