Review

Michael Landreh, Jan Johansson, John Wahren[a] and Hans Jörnvall[b,*]

The structure, molecular interactions and bioactivities of proinsulin C-peptide correlate with a tripartite molecule

Abstract: Many biological roles have been assigned to proinsulin C-peptide over the years. Some appear surprisingly disparate and sometimes even contradictory, like chaperone-like actions and depository tendencies. This review summarizes recently reported biomolecular interactions of the peptide and presents how they correlate with structural and functional aspects into a partitioned molecular architecture. At the structural level, the C-peptide sequence and fold can be subdivided into three distinct parts ('tripartite'). At the functional level, its chaperone-like abilities, self-assembly, and membrane interactions, as well as interactions with relevant proteins can be separately ascribed to these three segments. At the biological level, the assignments are compatible with the suggested roles of C-peptide in granular insulin storage, chaperone-like activities on insulin oligomers, possible depository tendencies, and proposed receptor interactions. Finally, the assignments give interesting parallels to further bioactive peptides, including glucagon and neurotensin. Provided pharmaceutical and clinical trials are successfully completed, the present interpretations should supply mechanistic explanations on C-peptide as a bioactive compound of importance in health and diabetes.

Keywords: amyloid formation; diabetes; membrane binding; protein folding; receptor interactions.

[a]J.W. is employed by Cebix AB and is a minority shareholder.
[b]H.J. is a minority shareholder of Cebix AB.
***Corresponding author: Hans Jörnvall,** Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden, e-mail: Hans.Jornvall@ki.se
Michael Landreh: Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden
Jan Johansson: KI-Alzheimer's Disease Research Center, NVS Department, Karolinska Institutet, S-141 86 Stockholm, Sweden; and Department of Anatomy, Physiology and Biochemistry, Swedish University of Agricultural Sciences, S-751 23 Uppsala, Sweden
John Wahren: Department of Molecular Medicine and Surgery, Karolinska Institutet, S-171 77 Stockholm, Sweden

Introduction

Many studies during the last few years have provided novel structural knowledge of C-peptide, highlighted molecular interactions, and suggested functional interpretations. Initially, the separate reports appeared disparate and, in part, contradictory by assigning multiple functions to this peptide, once regarded as monofunctional in promoting the correct proinsulin fold. However, together with recent data, the different reports now appear interpretable into a consistent pattern and give an integrated view on the nature of C-peptide. This review is intended to summarize the present molecular data and to illustrate novel answers to two of the previously strongest arguments against any function of mature C-peptide: its apparent divergence with lack of strictly conserved residues, and its largely random tertiary structure. Instead, we now show that charge patterns are functionally important, that they are largely conserved, and that the otherwise limited conservation gives possible parallels with other peptide structures in glucagon and neurotensin, where small ordered regions suffice to induce binding. Combined, these insights reveal a consistent pattern of C-peptide molecular interactions and functional interpretations.

Sequence divergence of C-peptide

As expected from the strong conservation of the insulin genes across the biological kingdoms, the peptide linker between the B- and A chains of insulin (1) is found in a wide variety of organisms. No post-translational residue modifications have been reported, but C-peptide is subject to proteolytic cleavage in plasma and kidney cells (2–5).

The known mammalian and vertebrate C-peptide sequences differ in length, ranging from 26 residues in the bovine and ovine forms and 31 residues in human, to over 35 residues in fish. Similarly, residue divergence

is considerable [Figure 1; from Ref. (6), used with permission]. While the insulin A and B chains are highly conserved, the connecting region that encompasses C-peptide is obvious (Figure 1) by a marked interspecies variation. However, as outlined below, the sequences, although divergent, contain some conserved key features.

The mammalian C-peptide homologues contain three to four acidic and no basic residues, with the exception of a single arginine at position 8 of the canine C-peptide sequence (7). The high content of acidic residues gives an isoelectric point between pH 3 and 4. Human C-peptide has a p*I* of 3.45; canine C-peptide, 3.71; and rat C-peptide, 3.77. In addition, mammalian C-peptides contain one to three proline and three to eight glycine residues, as well as a low amount of hydrophobic and no aromatic amino acids. They have a negative hydrophobicity average, which qualifies them as hydrophilic peptides and indicates good solubility in aqueous solvents. Of the non-mammalian variants, the avian and amphibian C-peptide sequences share these features to some extent, while the fish homologues only contain a conserved glutamate residue at position 3 (Figure 1). In the following, only the mammalian C-peptide sequences, and especially human C-peptide, will be discussed.

Conserved features of the C-peptide sequence

Besides the similar amino acid compositions, the mammalian C-peptides also share a similar overall distribution of particular residues in the sequence. Most strikingly, they contain three relatively conserved glutamic acids at positions 3, 11, and 27 (numbering according to the human C-peptide sequence).

C-peptide can be subdivided into three parts (Figure 2A) (8). Most acidic residues, namely, Glu3 and Glu11, as

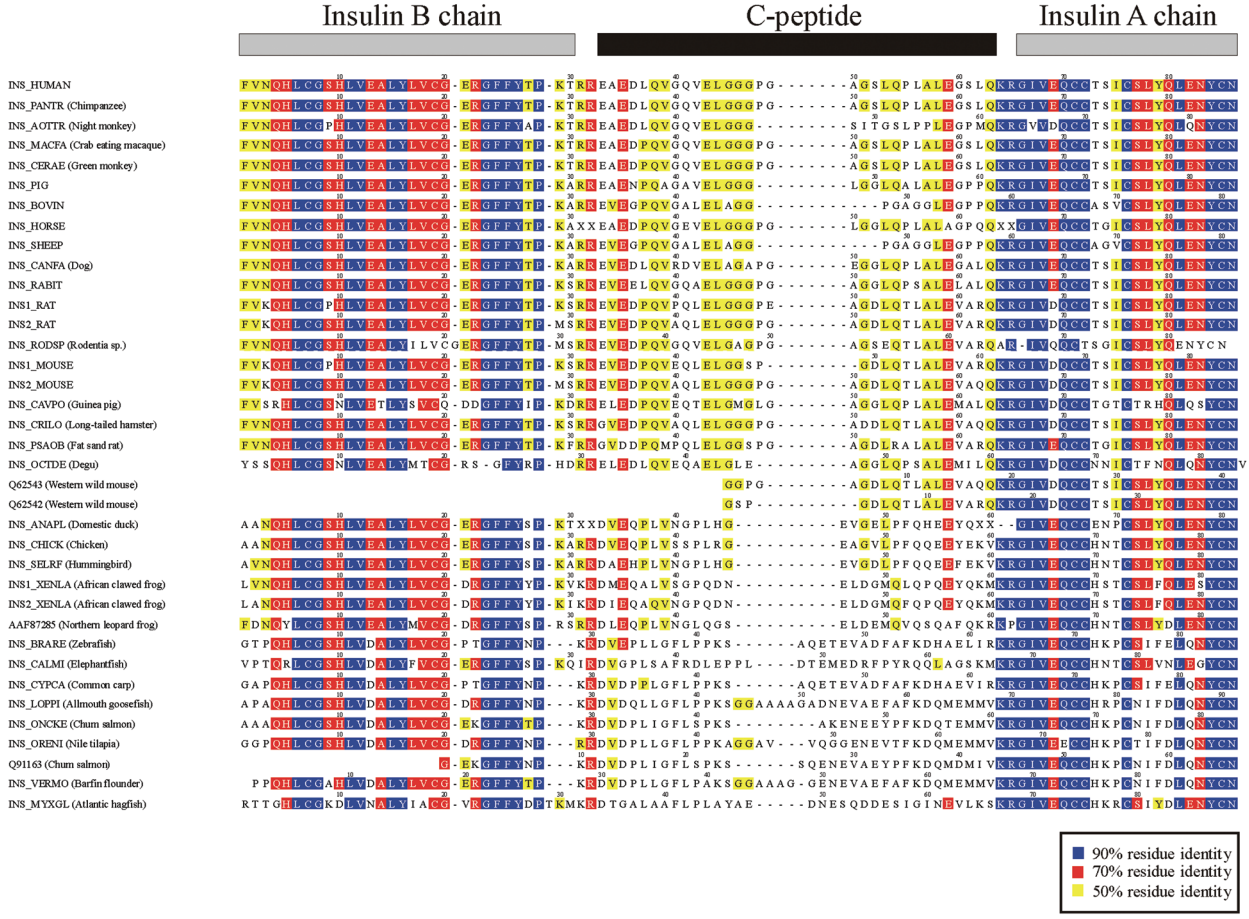

Figure 1 An alignment of proinsulin sequences from mammals, birds, amphibians and fish illustrates the comparatively low conservation of the C-peptide part in relation to the A- and B chains of insulin.
Glu3 and Glu27 (positions 35 and 59, respectively, in the human proinsulin sequence) represent the only well-conserved residues in C-peptide. [Reproduced with permission from Ref. (6).]

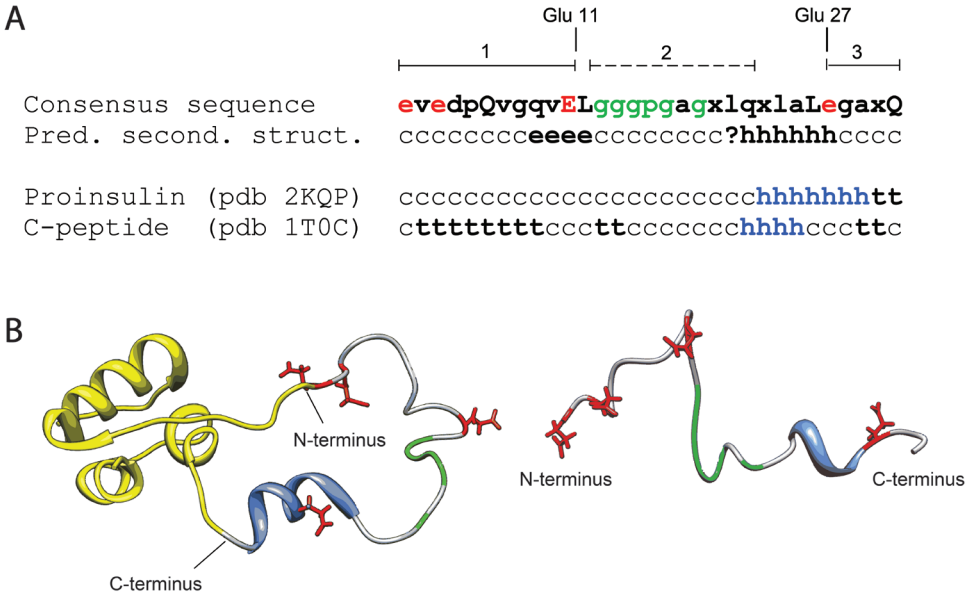

Figure 2 The three-dimensional structure of C-peptide.
(A) The consensus sequence of the mammalian C-peptides can be subdivided into three parts: (1) The N-terminal segment, which contains the Glu residues (shown in red), important for the chaperoning function of C-peptide; (2) the middle segment rich in Gly and Pro (shown in green), which confers the large conformational flexibility of C-peptide and mediates the association with phospholipids; and (3) the C-terminal pentapeptide, which contains Glu27 (shown in red) and is implicated in possible receptor binding. Large and small lettering indicate strong and intermediate conservation, respectively. A comparison of the secondary structure prediction of the consensus sequence with the experimentally determined structures shows a conserved helical segment at the C-terminal end as well as some β-strand/β-turn propensity in the N-terminal segment. (B) The NMR structures of proinsulin (PDB ID 2KQP) and C-peptide (PDB ID 1T0C) confirm the presence of a C-terminal α-helix (shown in blue). The insulin moiety of proinsulin is shown in yellow, and the C-peptide Glu residues are rendered as stick models.

well as Glu1 of human and primate C-peptides, are located in the N-terminal third of the sequence, which also constitutes its most conserved region. Pro and Gly residues are clustered in the middle segment, together with a variable content of Ala and Leu. The variability of this region is, in part, due to positional shuffling of these amino acids, which are found in all mammalian C-peptide sequences. The third segment contains the C-terminal pentapeptide that begins with a conserved Glu at position 27. Besides a C-peptide with the N-terminal dipeptide removed, the C-terminal pentapeptide is the most abundant C-peptide fragment found in plasma (4, 5).

The secondary and tertiary structures of C-peptide

The sequence features outlined above are directly coupled to the fold of the C-peptide segments and have sparked detailed investigations of their importance for the C-peptide structure.

Already, early Fourier transform infrared spectroscopy suggested that the C-peptide part of human proinsulin has a largely random coil conformation in aqueous solution (9). This was confirmed by later circular dichroism (CD) measurements as well as NMR and crystallographic studies (2, 6, 10, 11). However, some helical propensities were observed in titration experiments with 2,2,2-trifluoroethanol (TFE), which promotes a helical conformation even at low concentrations (6).

The predominance of a random coil structure is evidently due to the amino acid composition of the middle segment. Its high proline and glycine content prevents the formation of stable secondary structure elements and, instead, confers conformational flexibility. This suggests that the helical propensities must reside in the N- and C-terminal segments. Computational studies have lent support to this interpretation, depicting the C-peptide structure as a helix-turn-helix motif (12). In line with this model, proline substitutions at positions 4, 7, 8, and 10 of the N-terminal segment greatly reduce its helical propensity, while alanine residues at these positions increase helix formation (6).

Detailed molecular dynamics analyses have suggested that residues 3–6 of the human but not of the rat C-peptide can transiently adopt an α-helical conformation in solution (13). Interestingly, the same studies have provided evidence that not the C-terminal segment, but a

short stretch of the middle segment, residues 20–23, may also form an α-helical structure. Again, no such propensities were detected in rat C-peptide, which differs from human C-peptide at nine positions (13).

Experimental clarification of the exact localization of the human C-peptide α-helices was provided by Yang et al. (11), who reported the NMR structure of proinsulin, and by Munte et al. (14), who determined the solution structure of free C-peptide in 50% TFE. The proinsulin structure shows that the spatial arrangement of the insulin A- and B chains is nearly identical to that in mature insulin and, because of the three disulfide bridges, exhibits a low flexibility (14). Thus, the C-peptide linker is tethered at its N- and C-termini, but no intramolecular nuclear Overhauser effects between the C-peptide and insulin moieties were detected, which indicates that the insulin part has no further influence on the fold of the proinsulin-linked C-peptide. Interestingly, the NMR structure shows that the C-terminal pentapeptide (residues 27–31) of C-peptide is helical, in agreement with the early C-peptide models (12), while the rest of the sequence is in random coil conformation (Figure 2B).

The structure of free C-peptide largely agrees with the results obtained for proinsulin (Figure 2B) (6). The middle segment again has a random coil structure, while residues 27–31 form a type III β-turn (i.e., a single turn of a 3_{10} helix). The N-terminal segment represents the second-most structured part in the form of a type I β-turn encompassing residues 2–5, which may initiate a transient helical conformation in this region (6). Interestingly, acidic pH is able to alter the secondary structure of C-peptide towards a mild preference for β-sheet conformation in the N-terminal segment, while the majority of the peptide maintains a random coil conformation (15). However, a high-resolution structure under these conditions is not available.

Together, the computational and NMR-based structure investigations converge into the previously described three-partite C-peptide structure (8): (1) an acidic N-terminal segment with limited secondary structure propensity; (2) an unstructured, relatively hydrophobic middle segment; and (3) a helical C-terminal pentapeptide (Figure 2A). The functional implications of the structural features in each segment are discussed below.

The N-terminal segment of C-peptide mediates proinsulin folding and prevents peptide aggregation

A recurring question regarding the biological role of C-peptide concerns its co-synthesis with insulin. The resulting 1:1 ratio of the two peptides at synthesis may be of importance for the biological effects of C-peptide or for the organization of the secretory granule (below). However, a more direct relationship was uncovered by Tang et al. (16), who investigated the effects of mutations in C-peptide on proinsulin folding. Point substitutions as well as deletions of stretches of three to six residues were constructed into the C-peptide part of human proinsulin. The constructs were expressed recombinantly, purified under reducing conditions, and subjected to *in vitro* refolding. The yield of correctly refolded proinsulin variants with intact disulfide bridges was assessed by SDS-PAGE analysis and receptor-binding assays. Surprisingly, the deletion or the replacement of residues 3–6 of C-peptide with alanine resulted in a 60–80% lower ability to refold, indicating that the N-terminus of C-peptide is important for the correct folding of proinsulin (16). Similar approaches, involving the detailed monitoring of disulfide formation by LC-MS, revealed that the B-chain segment of proinsulin acts as a template for folding of the A chain. It has been speculated that the repulsion caused by the negative charges of C-peptide may help to prevent the formation of intermolecular hydrophobic interactions between the unfolded chains, which would otherwise lead to aggregation (17, 18). Analysis of C-peptide and insulin evolution supports this model and shows that specific residues in the insulin A- and B chains and the acidic charges in the C-peptide N-terminal segment have co-evolved to facilitate the interplay between the different proinsulin segments during folding (19). Apparently, this interplay still exists even between the non-covalently linked peptides, constituting the basis for peripheral chaperon-like effects on insulin (below) and possibly also for parts of their interactions during storage and secretion (also below).

Following enzymatic processing, C-peptide is stored in the secretory granules of the pancreatic β-cells alongside with insulin, islet amyloid polypeptide (IAPP), additional proteinaceous components, and metal ions (20, 21). The pH in the granule is mildly acidic and close to the isoelectric point of insulin at pH 5.5 (22). This is surprising, since the *in vitro* solubility of insulin is reduced under these conditions. Yet, it can be stored *in vivo* in mature, folded form and at extreme concentrations of >40 mmol/l (21). Ultrastructural and microscopic analyses of mature granules have shown that a part of their insulin cargo is stored in the form of Zn^{2+}-containing crystals surrounded by a 'halo' containing soluble insulin, C-peptide, and IAPP (23, 24).

While insulin tends to oligomerize at granular pH and subsequently aggregates (25), electrospray ionization mass spectrometry (ESI-MS) experiments have shown that

C-peptide can prevent the formation of these prefibrillar insulin oligomers, and the underlying direct interactions between C-peptide and insulin could be detected using surface plasmon resonance (26). The presence of an excess of insulin was found to promote the formation of heteromolecular complexes, while 'scrambled' C-peptide (i.e., a C-peptide variant with randomized sequence) was unable to bind to insulin (26). In studies using fragments covering different sections of the C-peptide sequence, the desaggregating effect could be pinpointed to the N-terminal, highly charged segment that includes Glu1, Glu3, and Glu11 (27). These charged residues mediate co-precipitation of C-peptide with insulin via mutual charge neutralization at granular pH (28). This process is reversible by elevating the pH from 5.5 to the plasma pH level of 7.5 and may hint at a role for C-peptide in balancing the soluble, crystalline, and precipitate states of insulin in the secretory granule (Figure 3) (29).

Besides being an endocrine disorder, diabetes of both types is also a protein aggregation disease, as evident from the accumulation of misfolded protein species (30). Insulin can form amyloid deposits at the site of repeated insulin injections in patients with diabetes (31). In its amyloid state, insulin is converted from its native, mostly α-helical

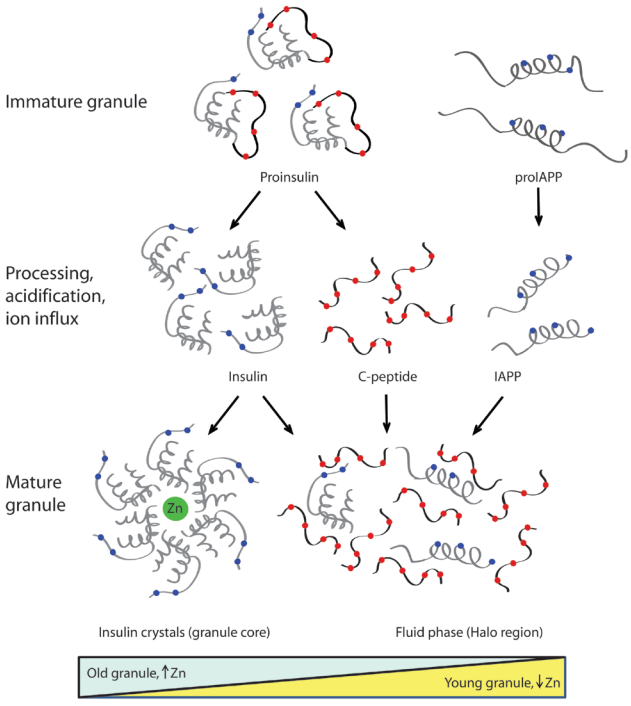

Figure 3 At granular pH, C-peptide binds insulin and IAPP, and prevents their self-association.
The resulting chaperoning activity possibly helps to balance peptide solubility in the secretory granule by partitioning insulin into the Zn^{2+}-crystalline and -soluble forms. The distribution of positive charges is indicated by blue and negative charges by red circles.

fold into a cross-β-strand conformation via an unfolded intermediate and stacked into elongated, unbranched fibrils (32, 33). C-peptide has been demonstrated to interfere with this process in a charge-dependent manner (34). Interestingly, the most aggregation-prone segment of the insulin molecule, which is located in the B chain, shares a co-evolutionary pattern with Glu3 in the C-peptide N-terminus (19), further supporting the suggestion that C-peptide interactions can prevent insulin misfolding.

The second amyloidogenic protein implicated in diabetes is the highly aggregation-prone IAPP. IAPP amyloid deposits have been found in the pancreas of over 90% of patients with diabetes and have been suggested to contribute to β-cell death in type 1 diabetes (35). C-peptide has been shown to interfere with IAPP fibrillation in a concentration-dependent manner (24, 36). However, the peptides interact only weakly due to being mostly unstructured in solution (37). Interactions with insulin, in contrast, likely induce a stable secondary structure in both IAPP and C-peptide, and, in this manner, enable complex formation (37, 38).

Co-administration of C-peptide and insulin to patients with diabetes was found to accelerate tissue glucose uptake significantly, consistent with an accelerated action of the administered insulin (26). Since oligomerization can delay the effects of the administered insulin (39), these data suggest that C-peptide can exert the same desaggregating effect *in vivo* as *in vitro*, leading to increased availability of monomeric insulin and thus faster action (26). Recently, similar conclusions have been drawn also from C-peptide administration in rats (40). Provided insulin aggregates are a part of late complications in diabetes, these chaperone-like properties of C-peptide may be useful for the treatment of diabetic complications.

The middle segment of C-peptide contributes to self-association and pH-dependent membrane interactions

Its conserved negative charges not only suggest that C-peptide is well soluble in aqueous solutions, but also indicate that the peptide conveys strong repulsive forces at neutral pH. Therefore, it is surprising that C-peptide exhibits a robust tendency to self-associate.

The first evidence for intermolecular interactions and oligomer formation came from ESI-MS spectra showing C-peptide assemblies of two to five monomers (41).

Western blotting of biotinylated C-peptide variants even revealed the presence of higher-mass species of 15–30 kDa, corresponding to 5–10 C-peptide molecules (42). These observations were confirmed by NMR and dynamic light scattering analysis (42). Self-association of C-peptide increases in a time-dependent manner when incubated in aqueous solution or in the presence of submicellar concentrations of SDS, while bivalent metal ions and insulin reduce oligomerization (41, 42). The effects of metal ions may be due to the metal ion-binding abilities involving Glu27 and to the correspondingly increased solubility of the peptide-ligand complex (43).

Self-association has been suggested to be mediated by the middle segment of C-peptide due to the absence of repulsive charges in this region (42). Further investigations of the SDS-induced assemblies by CD and attenuated total reflectance-IR spectroscopy revealed a structural transition from the previously described random coil/α-helical structure to a more β-strand-rich conformation (42). In line with these findings, the oligomers stained positively for thioflavin T (41), a dye specific for β-strand aggregates such as amyloid fibrils (44). Considering that deposits of C-peptide have been identified in atherosclerotic lesions of patients with type 2 diabetes, the self-association of C-peptide raises the possibility that it may be capable of forming amyloid-like aggregates *in vivo* (45). However, no detailed studies of the structural characteristics of these deposits or their relevance in diabetes exist to date. Nevertheless, the tendency towards structural transitions and aggregation suggest that care should be taken during storage of C-peptide in solution.

Given the ability of C-peptide to bind to cells and cross the membrane barrier (46, 47), the interactions of C-peptide with lipids are of special interest. However, the presence of lipid vesicles had no impact on the C-peptide structure, as judged by CD analysis. Furthermore, C-peptide and lipid vesicles did not co-migrate in gel filtration, and it was concluded that they do not engage in molecular interactions (48). These results are not surprising, considering the highly negative charge of C-peptide at physiological pH, which does not favor interactions with hydrophobic hydrocarbon chains or negative head groups.

These repulsive forces, however, are partially alleviated when approaching the p*I* of C-peptide. At pH 3.2, the presence of lipid bilayers was reported to induce chemical shift changes in the NMR signals from individual residues between positions 2 and 26 (15). The changes at positions 10–12 are indicative of β-strand formation in this area, while residues 17–19 adopt a more helical structure. Therefore, potential interactions with lipids occur predominantly via the middle segment of C-peptide and require the protonation of Glu11. In the light of this finding, limited, pH-dependent interactions of C-peptide with biological membranes, e.g., in acidic compartments such as lysosomes, may be possible, but have not been reported to date, although C-peptide has been reported to occur in early endosomes (49).

The C-terminal pentapeptide is a putative receptor-binding site

Early on, it became clear that some of the endocrine effects of C-peptide might be explained by the interactions of C-peptide with a receptor of the G-protein-coupled receptor (GPCR) type (50–52). In general, receptor-mediated signaling relies on high structural complementarity between ligand and receptor, and therefore usually requires a high degree of sequence conservation in both receptor and peptide ligand. While C-peptide does not exhibit the expected conservational pattern (Figure 1), it has been demonstrated that randomization of its sequence abolishes the cellular effects that are believed to be receptor mediated (53). Studies using C-peptide fragments have indicated that the C-terminal pentapeptide may be a suitable candidate for initial receptor binding (51, 52, 54). This fragment was shown to displace full-length C-peptide from binding to cellular membranes (52, 54), contains the conserved Glu27 (Figure 1), and is the only part of C-peptide that adopts a defined secondary structure (described above). *In vivo* studies also suggest a role of the C-terminal pentapeptide (40, 55), although the mechanisms involved have not been studied yet.

Recently, a surge in high-resolution structures of ligand-bound GPCRs has revealed the basis for recognition of their diverse ligands, ranging from small molecules like monoamines and lipids to peptides and proteins. Class A GPCRs recognize their target with the help of loops that connect the seven transmembrane helices on the extracellular side (56, 57). Class B GPCRs feature a large extracellular domain that is involved in ligand recognition (58). GPCRs of both of these classes mediate the endocrine activities of several diabetes-related peptide hormones that exhibit varying degrees of sequence variability (59).

A class A GPCR was recently suggested to be a possible C-peptide receptor. An RNAi knock-down study of three selected orphan GPCRs in cell lines that displayed GPCR activation in response to C-peptide treatment showed that only GPR146 ablation abolished the effects of C-peptide, making GPR146 a potential candidate receptor (60).

Comparison of the class A and class B GCPR-peptide ligand complexes reveals that, in both cases, short sequence segments play important roles in the recognition by the receptor. Due to their relatively shallow binding pocket compared to class B receptors (61), class A GCPRs may preferentially recognize short peptide ligands, as illustrated by the structure of the rat neurotensin receptor (NTSR1) in complex with the five C-terminal residues of its neurotensin ligand (62) (Figure 4B). The peptide is inserted into the extracellular binding pocket of the GPCR transmembrane domain, suggesting that this short fragment is enough to activate the receptor (Figure 4B). Such an interaction would, in theory, be compatible with C-peptide as well as with other class A GPCR ligands with low conservation, such as relaxin.

However, GPCR activation can also rely on a few key motifs that are distributed across the entire ligand sequence and interact with the receptor at multiple positions. Such a mechanism can be found in some class B GPCRs (58) as illustrated by the interactions between glucagon and its receptor GCGR (61). Like C-peptide, glucagon has a flexible structure but can adopt a helical conformation. Its helical C-terminal segment is bound by the extracellular domain of the GCPR and helps to orient the N-terminal part correctly for receptor activation (Figure 4C) (12, 63). Such an interaction for C-peptide would also be compatible with the observation that its C-terminal fragment displaces full-length C-peptide from cell membranes (54). However, the sequence conservation of glucagon is considerably higher than that of C-peptide (59), and its receptor interactions require specific residue contacts at multiple positions (61). Whether the low number of relatively conserved residues in C-peptide is sufficient for such a mode of interaction remains to be investigated.

If C-peptide can indeed specifically bind to a GPCR, it appears possible that the C-terminal C-peptide segment plays a role in its interactions, either as the only site in C-peptide that is recognized by a class A GPCR or by mediating the first interaction with a class B GPCR, which then, e.g., places and folds the bound peptide for receptor activation (12). In the light of these findings, biochemical investigations of the interactions of C-peptide variants with potential receptors are of high interest. However, independent of receptor type, the C-terminal pentapeptide region of C-peptide is concluded to be involved in the ligand interactions.

Conclusions

The structure of proinsulin C-peptide suggests that its different parts may fulfill separate molecular functions. A

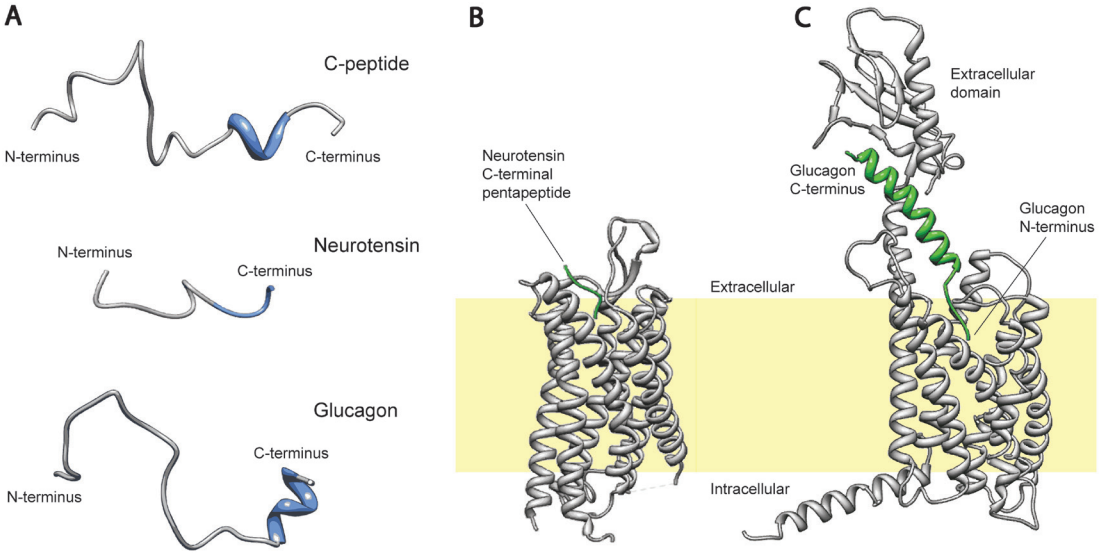

Figure 4 Possible models for C-peptide-mediated receptor signaling.
(A) A comparison of the structures of C-peptide in 50% TFE (top, PDB ID 1T0C), neurotensin in the presence of lipid bicelles (middle, PDB ID 2NLF), and glucagon at the air-water interface (bottom, PDB ID 1KX6) shows the similar inducible helix in the C-terminal region (a β-turn in neurotensin), highlighted in blue. (B) The crystal structure of the rat neurotensin receptor (PDB ID 2GRV) in complex with the C-terminal neurotensin fragment (green) shows how only a five-residue peptide segment is recognized by the extracellular loops in a class A GPCR (62). (C) The molecular model of the glucagon GGCR in complex with its ligand in helical conformation (shown in green) (61) illustrates how the C-terminal segments of the glucagon ligand are recognized by the class B GPCR extracellular domain.

tripartite layout as now described may be an evolutionary concession to accommodate chaperoning activity, self-association, and possibly even receptor interactions, all in a comparatively short peptide, and the individual segments consequently display variable degrees of sequence conservation. Since the insulin sequence is highly conserved, the residues in the N-terminal segment of C-peptide involved with insulin chaperoning have co-evolved accordingly with insulin (19). Self-association and conformational flexibility, as ascribed to the middle segment, are less dependent on a specific sequence, and the biological significance of these features remains to be elucidated. Receptor interactions, as suggested for the C-terminal pentapeptide, usually involve a conserved binding motif. The relatively defined structural propensities of the C-terminal segment may hint at the existence of such a motif, which should then display an evolutionary correlation with its putative receptor (64), the identity of which remains to be established. Nevertheless, the present correlations give a thus far coherent explanation for the multiplicities of C-peptide actions at both its structural and its bioactivity levels.

Expert opinion and outlook

Since the discovery of C-peptide, our view of its biological importance has constantly evolved, and the research field now stretches from biophysics to clinical endocrinology. In this review, we describe how the proinsulin C-peptide sequence, despite its considerable variability, contains three distinct segments that can be assigned specific molecular interactions and different biological roles. This subdivision reveals how the previously diverse and partially unrelated functions are integrated into a molecular architecture.

Among the many outstanding questions regarding C-peptide physiology, we consider its roles in protein folding and its hormone-like actions to be of special interest. As described here, there is now evidence for a connection between C-peptide and diabetes-related protein aggregation as well as a receptor, both of which may become structurally clearer in the near future. With an insight into these mechanisms, important physiological effects of C-peptide can be considered interpretable and may be harnessed for the treatment of diabetic complications. In addition, the ongoing investigations of the molecular biology of C-peptide are likely to uncover additional connections with further levels of the endocrine system. As we describe in this review, C-peptide is a part of the complex regulatory networks that are imbalanced in diabetes. Therapeutic strategies will develop from the substitution of individual components to finely tuned reconstitution of their interactions, lending additional importance to the understanding of the multiplicity of C-peptide mechanisms.

Highlights

- Proinsulin C-peptide contains conserved biophysical features despite its variable sequence.
- These features subdivide the C-peptide structure into three distinct parts, with relatively conserved Glu residues.
- The acidic N-terminal part mediates the chaperoning action on insulin and other peptides of the secretory granules, indicating a role in granular storage and secretion mechanisms.
- The flexible middle segment governs self-association and membrane interactions under acidic conditions.
- The pro-helical pentapeptide at the C-terminal end is involved in putative G-protein coupled receptor interactions as is supported also by parallels to other peptide ligands.
- Notably, although the three segments initiate each activity, end interactions are likely to involve longer parts; hence, the different activities need not be independent or simultaneous.
- The three-partite structure explains how the reported diverse biological actions of C-peptide can be accommodated in a single peptide.

Acknowledgments: The GCGR-glucagon model [described in Siu et al. (61)] was kindly provided by Stevens and co-workers. The present work was supported by The Swedish Research Council (grant O3X-3532).

Received February 18, 2014; accepted February 21, 2014

References

1. Steiner DF, Cunningham D, Spigelman L, Aten B. Insulin biosynthesis: evidence for a precursor. Science 1967; 157: 697–700.
2. Varandani PT, Shroyer LA. Degradation of proinsulin and isolated C-peptide by rat kidney neutral metallo-endopeptidase. Biochim Biophys Acta 1981; 661: 182–90.
3. Melles E, Jörnvall H, Tryggvason S, Danielsson KG, Ekberg K, Tryggvason K, Wahren J, Bergman T. Degradation of proinsulin C-peptide in kidney and placenta extracts by a specific endoprotease activity. Cell Mol Life Sci 2004; 61: 2979–82.
4. Melles E, Bergman T, Alvelius G, Jonsson A, Ekberg K, Wahren J, Jörnvall H. Proinsulin C-peptide and its C-terminal pentapeptide: degradation in human serum and Schiff base formation with subsequent CO_2 incorporation. Cell Mol Life Sci 2003; 60: 1019–25.
5. Oran PE, Jarvis JW, Borges CR, Nelson RW. C-peptide microheterogeneity in type 2 diabetes populations. Proteom Clin Appl 2010; 4: 106–11.
6. Henriksson M, Nordling E, Melles E, Shafqat J, Ståhlberg M, Ekberg K, Persson B, Bergman T, Wahren J, Johansson J, Jörnvall H. Separate functional features of proinsulin C-peptide. Cell Mol Life Sci 2005; 62: 1772–8.
7. Kwok SC, Chan SJ, Steiner DF. Cloning and nucleotide sequence analysis of the dog insulin gene. Coded amino acid sequence of canine preproinsulin predicts an additional C-peptide fragment. J Biol Chem 1983; 258: 2357–63.
8. Nerelius C, Lindahl E, Landreh M, Jörnvall H. Peptide interactions of proinsulin C-peptide. New York: Springer, 2012.
9. Xie L, Tsou CL. Comparison of secondary structures of insulin and proinsulin by FTIR. J Prot Chem 1993; 12: 483–7.
10. Steiner DF. Cocrystallization of proinsulin and insulin. Nature 1973; 243: 528–30.
11. Yang Y, Hua QX, Liu J, Shimizu EH, Choquette MH, Mackin RB, Weiss MA. Solution structure of proinsulin: connecting domain flexibility and prohormone processing. J Biol Chem 2010; 285: 7847–51.
12. Ido Y, Vindigni A, Chang K, Stramm L, Chance R, Heath WF, DiMarchi RD, Di Cera E, Williamson JR. Prevention of vascular and neural dysfunction in diabetic rats by C-peptide. Science 1997; 277: 563–6.
13. Mares-Guia TR, Maigret B, Martins NF, Maia AL, Vilela L, Ramos CH, Neto LJ, Juliano MA, dos Mares-Guia ML, Santoro MM. Molecular dynamics and circular dichroism studies of human and rat C-peptides. J Molec Graph Model 2006; 25: 532–42.
14. Munte CE, Vilela L, Kalbitzer HR, Garratt RC. Solution structure of human proinsulin C-peptide. FEBS J 2005; 272: 4284–93.
15. Unnerståle S, Mäler L. pH-Dependent interaction between C-peptide and phospholipid bicelles. J Biophys 2012; 2012: 185–907.
16. Chen LM, Yang XW, Tang JG. Acidic residues on the N-terminus of proinsulin C-peptide are important for the folding of insulin precursor. J Biochem 2002; 131: 855–9.
17. Qiao ZS, Min CY, Hua QX, Weiss MA, Feng YM. In vitro refolding of human proinsulin. Kinetic intermediates, putative disulfide-forming pathway folding initiation site, and potential role of C-peptide in folding process. J Biol Chem 2003; 278: 17800–9.
18. Min CY, Qiao ZS, Feng YM. Unfolding of human proinsulin. Intermediates and possible role of its C-peptide in folding/unfolding. Eur J Biochem 2004; 271: 1737–47.
19. Wang S, Wei W, Zheng Y, Hou J, Dou Y, Zhang S, Luo X, Cai X. The role of insulin C-peptide in the coevolution analyses of the insulin signaling pathway: a hint for its functions. Plos ONE 2012; 7: e528–47.
20. Brunner Y, Coute Y, Iezzi M, Foti M, Fukuda M, Hochstrasser DF, Wollheim CB, Sanchez JC. Proteomics analysis of insulin secretory granules. Mol Cell Proteom 2007; 6: 1007–17.
21. Foster MC, Leapman RD, Li MX, Atwater I. Elemental composition of secretory granules in pancreatic islets of Langerhans. Biophys J 1993; 64: 525–32.
22. Davidson HW, Rhodes CJ, Hutton JC. Intraorganellar calcium and pH control proinsulin cleavage in the pancreatic beta cell via two distinct site-specific endopeptidases. Nature 1988; 333: 93–6.
23. Michael DJ, Ritzel RA, Haataja L, Chow RH. Pancreatic beta-cells secrete insulin in fast- and slow-release forms. Diabetes 2006; 55: 600–7.
24. Westermark P, Li ZC, Westermark GT, Leckström A, Steiner DF. Effects of beta cell granule components on human islet amyloid polypeptide fibril formation. FEBS Lett 1996; 379: 203–6.
25. Nettleton EJ, Tito P, Sunde M, Bouchard M, Dobson CM, Robinson CV. Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry. Biophys J 2000; 79: 1053–65.
26. Shafqat J, Melles E, Sigmundsson K, Johansson BL, Ekberg K, Alvelius G, Henriksson M, Johansson J, Wahren J, Jörnvall H. Proinsulin C-peptide elicits disaggregation of insulin resulting in enhanced physiological insulin effects. Cell Mol Life Sci 2006; 63: 1805–11.
27. Nerelius C, Alvelius G, Jörnvall H. N-terminal segment of proinsulin C-peptide active in insulin interaction/desaggregation. Biochem Biophys Res Commun 2010; 403: 462–7.
28. Landreh M, Alvelius G, Willander H, Stukenborg JB, Söder O, Johansson J, Jörnvall H. Insulin solubility transitions by pH-dependent interactions with proinsulin C-peptide. FEBS J 2012; 279: 4589–97.
29. Landreh M, Johansson J, Jörnvall H. C-peptide: a molecule balancing insulin states in secretion and diabetes-associated depository conditions. Horm Metab Res 2013; 45: 769–73.
30. Hayden MR, Tyagi SC, Kerklo MM, Nicolls MR. Type 2 diabetes mellitus as a conformational disease. J Pancreas 2005; 6: 287–302.
31. Dische FE, Wernstedt C, Westermark GT, Westermark P, Pepys MB, Rennie JA, Gilbey SG, Watkins PJ. Insulin as an amyloid-fibril protein at sites of repeated insulin injections in a diabetic patient. Diabetologia 1988; 31: 158–61.
32. Ivanova MI, Sievers SA, Sawaya MR, Wall JS, Eisenberg D. Molecular basis for insulin fibril assembly. Proc Natl Acad Sci USA 2009; 106: 18990–5.
33. Ahmad A, Uversky VN, Hong D, Fink AL. Early events in the fibrillation of monomeric insulin. J Biol Chem 2005; 280: 42669–75.

34. Landreh M, Stukenborg JB, Willander H, Söder O, Johansson J, Jörnvall H. Proinsulin C-peptide interferes with insulin fibril formation. Biochem Biophys Res Commun 2012; 418: 489–93.
35. Westermark P, Andersson A, Westermark GT. Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. Physiol Rev 2011; 91: 795–826.
36. Janciauskiene S, Eriksson S, Carlemalm E, Ahrén B. B cell granule peptides affect human islet amyloid polypeptide (IAPP) fibril formation in vitro. Biochem Biophys Res Commun 1997; 236: 580–5.
37. Landreh M, Alvelius G, Johansson J, Jörnvall H. Insulin, islet amyloid polypeptide and C-peptide interactions evaluated by mass spectrometric analysis. Rapid Commun Mass Spectrom 2014; 28: 178–84.
38. Wiltzius JJ, Sievers SA, Sawaya MR, Eisenberg D. Atomic structures of IAPP (amylin) fusions suggest a mechanism for fibrillation and the role of insulin in the process. Protein Sci 2009; 18: 1521–30.
39. Brange J, Andersen L, Laursen ED, Meyn G, Rasmussen E. Toward understanding insulin fibrillation. J Pharm Sci 1997; 86: 517–25.
40. Kubota M, Sato Y, Khookhor O, Ekberg K, Chibalin AV, Wahren J. Enhanced insulin action following subcutaneous co-administration of insulin and C-peptide in rats. Diabetes Metab Res Rev 2014; 30: 124–31.
41. Jörnvall H, Lindahl E, Astorga-Wells J, Lind J, Holmlund A, Melles E, Alvelius G, Nerelius C, Mäler L, Johansson J. Oligomerization and insulin interactions of proinsulin C-peptide: threefold relationships to properties of insulin. Biochem Biophys Res Commun 2010; 391: 1561–6.
42. Lind J, Lindahl E, Peralvarez-Marin A, Holmlund A, Jörnvall H, Mäler L. Structural features of proinsulin C-peptide oligomeric and amyloid states. FEBS J 2010; 277: 3759–68.
43. Keltner Z, Meyer JA, Johnson EM, Palumbo AM, Spence DM, Reid GE. Mass spectrometric characterization and activity of zinc-activated proinsulin C-peptide and C-peptide mutants. Analyst 2010; 135: 278–88.
44. Biancalana M, Koide S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochim Biophys Acta 2010; 1804: 1405–12.
45. Marx N, Walcher D, Raichle C, Aleksic M, Bach H, Grub M, Hombach V, Libby P, Zieske A, Homma S, Strong J. C-peptide colocalizes with macrophages in early arteriosclerotic lesions of diabetic subjects and induces monocyte chemotaxis in vitro. Arterioscl Thromb Vasc Biol 2004; 24: 540–5.
46. Wahren J, Ekberg K, Jörnvall H. C-peptide is a bioactive peptide. Diabetologia 2007; 50: 503–9.
47. Lindahl E, Nyman U, Melles E, Sigmundsson K, Ståhlberg M, Wahren J, Obrink B, Shafqat J, Joseph B, Jörnvall H. Cellular internalization of proinsulin C-peptide. Cell Mol Life Sci 2007; 64: 479–86.
48. Henriksson M, Shafqat J, Liepinsh E, Tally M, Wahren J, Jörnvall H, Johansson J. Unordered structured of proinsulin C-peptide in aqueous solution and in the presence of lipid vesicles. Cell Mol Life Sci 2000; 57: 337–42.
49. Luppi P, Geng X, Cifarelli V, Drain P, Trucco M. C-peptide is internalised in human endothelial and vascular smooth muscle cells via early endosomes. Diabetologia 2009; 52: 2218–28.
50. Ohtomo Y, Aperia A, Sahlgren B, Johansson BL, Wahren J. C-peptide stimulates rat renal tubular Na+, K(+)-ATPase activity in synergism with neuropeptide Y. Diabetologia 1996; 39: 199–205.
51. Ohtomo Y, Bergman T, Johansson BL, Jörnvall H, Wahren J. Differential effects of proinsulin C-peptide fragments on Na+, K+-ATPase activity of renal tubule segments. Diabetologia 1998; 41: 287–91.
52. Rigler R, Pramanik A, Jonasson P, Kratz G, Jansson OT, Nygren P, Stahl S, Ekberg K, Johansson B, Uhlén S, Uhlén M, Jörnvall H, Wahren J. Specific binding of proinsulin C-peptide to human cell membranes. Proc Natl Acad Sci USA 1999; 96: 13318–23.
53. Zhong Z, Davidescu A, Ehren I, Ekberg K, Jörnvall H, Wahren J, Chibalin AV. C-peptide stimulates ERK1/2 and JNK MAP kinases via activation of protein kinase C in human renal tubular cells. Diabetologia 2005; 48: 187–97.
54. Pramanik A, Ekberg K, Zhong Z, Shafqat J, Henriksson M, Jansson O, Tibell A, Tally M, Wahren J, Jörnvall H, Rigler R, Johansson J. C-peptide binding to human cell membranes: importance of Glu27. Biochem Biophys Res Commun 2001; 284: 94–8.
55. Sato Y, Oshida Y, Han YQ, Morishita Y, Li L, Ekberg K, Jörnvall H, Wahren J. C-peptide fragments stimulate glucose utilization in diabetic rats. Cell Mol Life Sci 2004; 61: 727–32.
56. Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. Ann Rev Pharmacol Toxicol 2013; 53: 531–56.
57. Granier S, Kobilka B. A new era of GPCR structural and chemical biology. Nat Chem Biol 2012; 8: 670–3.
58. Hollenstein K, de Graaf C, Bortolato A, Wang MW, Marshall FH, Stevens RC. Insights into the structure of class B GPCRs. Trends Pharm Sci 2014; 35: 12–22.
59. Wahren J, Shafqat J, Johansson J, Chibalin A, Ekberg K, Jörnvall H. Molecular and cellular effects of C-peptide – new perspectives on an old peptide. Exp Diabesity Res 2004; 5: 15–23.
60. Yosten GL, Kolar GR, Redlinger LJ, Samson WK. Evidence for an interaction between proinsulin C-peptide and GPR146. J Endocrinol 2013; 218: B1–8.
61. Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, Zhou C, Xu Q, Wacker D, Joseph JS, Liu W, Lau J, Cherezov V, Katritch V, Wang MW, Stevens RC. Structure of the human glucagon class B G-protein-coupled receptor. Nature 2013; 499: 444–9.
62. White JF, Noinaj N, Shibata Y, Love J, Kloss B, Xu F, Gvozdenovic-Jeremic J, Shah P, Shiloach J, Tate CG, Grisshammer R. Structure of the agonist-bound neurotensin receptor. Nature 2012; 490: 508–13.
63. Authier F, Desbuquois B. Glucagon receptors. Cell Mol Life Sci 2008; 65: 1880–99.
64. Lovell SC, Robertson DL. An integrated view of molecular coevolution in protein-protein interactions. Mol Biol Evol 2010; 27: 2567–75.