C-peptide microheterogeneity in type 2 diabetes populations

Paul E. Oran, Jason W. Jarvis, Chad R. Borges, and Randall W. Nelson
Molecular Biosignatures Analysis Unit, The Biodesign Institute at Arizona State University, Tempe, AZ, USA

Abstract

Purpose—The purpose of this study was to investigate naturally occurring C-peptide microheterogeneity in healthy and type 2 diabetes (T2D) populations.

Experimental design—MS immunoassays capable of simultaneously detecting intact C-peptide and variant forms were applied to plasma samples from 48 healthy individuals and 48 individuals diagnosed with T2D.

Results—Common throughout the entire sample set were three previously unreported variations of C-peptide. The relative contribution of one variant, subsequently identified as C-peptide (3–31), was found to be more abundant in the T2D population as compared to the healthy population. Dipeptidyl peptidase IV is suspected to be responsible for this particular cleavage product, which is consistent with the pathophysiology of T2D.

Conclusions and clinical relevance—C-peptide does not exist in the human body as a single molecular species. It is qualitatively more heterogeneous than previously thought. These results lay a foundation for future studies devoted to a comprehensive understanding of C-peptide and its variants in healthy and diabetic populations.

Keywords
C-peptide; Microheterogeneity; Population proteomics; Variants

In response to elevated blood glucose levels within the beta cells of the pancreatic islets of Langerhans, preproinsulin is produced in the endoplasmic reticulum and is cleaved by microsomal enzymes to generate proinsulin. Proinsulin is subsequently transported to the Golgi apparatus and packaged into clathrin-coated secretory granules, where it is processed by a cascade of proconvertases and carboxypeptidases, resulting in the portal circulation of equimolar amounts of C-peptide and insulin [1–3]. The reported half-life of C-peptide in plasma is between 20 and 30 min and its concentration is three to five times that of insulin [2, 4–6].

Conventionally, C-peptide functions as a mediator/connector in the protein folding of the proinsulin precursor molecule to facilitate two disulfide bridges between the α- and β-chains of the insulin molecule. Excluding this well-established role in the insulin biosynthesis pathway, the physiological functions of C-peptide have remained largely undefined. Recently, however, C-peptide has been shown to have alleviating physiological effects toward diabetes. Diabetic rats with pharmacological doses of C-peptide demonstrated a
restoration toward normal of the diabetes-induced decrease in cellular sodium–potassium adenosine triphosphatase activity and impaired nerve conduction, and a decrease in the diabetes-induced increase in vascular permeability and blood flow [7]. Moreover, structural studies indicate the C-terminal pentapeptide segment (27–31 residues, particularly Glu27) as the site of receptor–ligand binding, which is thought to be responsible for increasing intracellular Ca^{2+} concentrations, stimulation of MAP kinase signaling pathways, and the stimulation of Na^+, K^+-adenosine triphosphatase and endothelial eNOS. However, it is possible that there are other signaling pathways involved [2, 8–11].

In clinical application, C-peptide is a useful indicator for differentiating between type 1 and type 2 diabetes (T2D) [2, 4, 12–14]. Low fasting levels of C-peptide are used in combination with other biomarkers to classify diabetics as having idiopathic type 1 diabetes of nonautoimmune or autoimmune origin. C-peptide concentrations are generally measured using immunobased assays such as enzyme-linked immunosorbent assays [15] and radioimmunoassay [16], or fluorescence [17, 18]. However, a recent study investigating the interassay and interlaboratory reproducibility of these industry-standard assays has reached the conclusions that results generated by different methods and laboratories do not always agree, and that calibrating C-peptide measurement to a reference method can increase comparability between laboratories [17]. The reference method indicated in this recent study is isotope-dilution MS (ID-MS), which is a natural extension of a significant body of work using MS for the accurate quantification of C-peptide [19–25]. In particular, two methodologies, either 2-D LC/MS operating in select ion monitoring mode [19] or LC-MS/MS operating in a multiple reaction-monitoring mode [20], have demonstrated excellent analytical metrics (e.g. in accuracy 94.6–104.1% [20]), precision (CV<1.5% [19], CV<4% [20]), LOD 10 pM [20],] and LOQ ~50 pM [19, 20]) for the quantification of intact C-peptide. Accordingly, these approaches are likely to become “gold standard” assays for C-peptide [17, 23].

One caveat of the ID-MS approaches (as they are described, e.g. [19, 20] without full-scan data), as well as the conventional immunometric assays, is the inability to differentiate between wild-type and variant forms of C-peptide that might be encountered in plasma. Formally, the consideration of such microheterogeneity may help to explain measurement discrepancies found between ID-MS and conventional assays. Alternatively, from a biological perspective, the study of microheterogeneity is important when considering that genetic and posttranslational variants of proteins contain additional information regarding the personal makeup of an individual, which in turn may have significant implications toward personalized medicine. For example, inactivation of peptide hormones by N-terminal dipeptide truncation via dipeptidyl peptidase IV (DPP-IV), such as occurs with glucagon-like peptide-1 (GLP-1), results in multiple endogenous isoforms that can begin to describe an individual’s state of enzymatic activity (e.g. DPP-IV activity) and peptide signaling activity (e.g. the relative abundance of the agonist form of GLP-1) [26]. In addition, in acknowledging that increased activity levels of protein regulatory enzymes have been in many cases disease correlated, such as with T2D and DPP-IV [27], it is essential to take an in-depth look at the breadth of protein micro-heterogeneity.

Thus, future studies stand to benefit from a more complete survey of qualitatively different variants of C-peptide as they exist in healthy and diseased populations. To this end, this report focuses on the use of MS immunoassay (MSIA) [28] to map structural differences in C-peptide found in healthy and T2D populations.

1,1′-Carbonyldiimidazole-activated affinity pipette tips were prepared and derivatized with mouse antihuman C-peptide antibody (AbD Serotec), as previously described for other antibodies [29]. For the development of the assay, healthy human female plasma was
purchased from ProMedDx (Norton, MA). Ninety-six additional human plasma samples (equaling 48 healthy individuals, 12 insulin-dependent T2D, and 36 noninsulin-dependent T2D) were purchased from ProMedDx. Age and gender information was available for 47 of the 48 samples from each disease cohort. The average age of healthy patients was 38.6 years with a gender distribution of 23 females and 24 males. The average age of the T2D cohort was 53.4 years with a gender distribution of 22 females and 25 males. Five hundred microliters of human plasma was pretreated with 250 μL of a solution containing: 4.5% Tween 20, 150 mM octyl-β-glucopyranoside, 1.5 M ammonium acetate, and concentrated PBS (0.67 M sodium phosphate, 1 M sodium chloride), for a total analytical volume of 0.75 mL.

C-peptide was extracted with the aid of a Beckman Multimek 96 pipetting robot by repeatedly (250 repetitions) drawing and expelling (back into the analytical volume) 125 μL aliquots of the analytical volume through an anti-C-peptide affinity pipette. After extraction, the pipettes were rinsed using PBS (0.1 M sodium phosphate, 0.15 M sodium chloride, pH 7.2), HBS-P (0.01 M HEPES, 0.15 M NaCl, 0.05% v/v surfactant P20, pH 7.4), H2O, 100 mM Tris-HCl (pH 4.6), and H2O (in this order, each rinse = 10 repetitions of 150 μL), after which C-peptide was eluted and prepared for MALDI-TOF MS by drawing 4 μL of MALDI matrix solution (saturated aqueous solution of CHCA, in 33% v/v ACN, 0.45% v/v TFA) into the pipette and depositing onto a MALDI target.

MALDI-TOF MS was performed using a Bruker Ultraflex MALDI-TOF instrument operating in the negative ion, delayed-extraction mode; reflector engaged with “ion source 1” at 20.00 kV, “ion source 2” at 17.45 kV, lens at 7.50 kV, “reflector” at 21.00 kV, “reflector 2” at 10.95, 60 ns delayed extraction, deflection signal suppression up to m/z 1500, and 2 GS/s sample rate. Two hundred laser-shots were signal averaged for each mass spectrum. Single stage MALDI-TOF mass spectra were externally calibrated with a mixture of six peptides supplied by Bruker (Cat. No. 208241) ranging in monoisotopic m/z from 1044.53 (Angiotensin II [M–H]−) to 3145.46 (Somatostatin [M–H]−)

Individual mass spectra were baseline subtracted and smoothed prior to peak integration. Peaks representing the five most abundant C-peptide and C-peptide-variant isotopes were integrated and tabulated in a spreadsheet for the determination of percent relative abundances.

LIFT-TOF/TOF mass spectra of C-peptide were acquired on a Bruker Ultraflex MALDI-TOF/TOF instrument in positive ion mode using sinapic acid as a matrix. Precursor ions were selected within a 30 Da mass window. The LIFT-TOF/TOF instrument control method was calibrated according to the instrument manufacturer’s protocol and was run without CID gas.

Figure 1 shows a MSIA spectrum qualitatively representative of those obtained for the individuals investigated in this study. In addition to intact C-peptide (observed monoisotopic m/z 3017.50), signals registering at m/z 2888.49, 2817.45, and 2688.41 (labeled I, II, and III, respectively) were observed at varying relative intensity in many but not all samples throughout both the healthy and T2D cohorts (n=48 each). The sequence of human C-peptide is H2N-EAEDLQVGQVELGGPGAGSLQPLALEGSLQ-CO2H.

Two major considerations arise from protein surveys such as this one that are conducted across human populations. First, the concept of populational frequency addresses the question of the number of samples (of the total), in which a particular protein variant was observed with signal-to-noise ratio greater than 3: Of the 48 T2D, isoform I was observed in 1 subject, isoform II in 44 subjects, and isoform III in 1 subject. In the healthy cohort,
isoform I was observed in nine subjects, isoform II in 43 subjects, and isoform III was not observed.

Second, the relative percent abundance (RPA) of each unique variant may be obtained (without the use of an internal standard) by integrating all mass spectral peak areas corresponding to variant forms of the target protein, then dividing the peak area of each variant form by the summed areas of all forms and multiplying by 100. The use of affinity capture ensures that spurious peaks unrelated to the target protein do not arise in the mass range of interest. Using this form of relative quantification, the RPA for C-peptide variant II was evaluated with respect to the presence of T2D by grouping data from individuals into their respective cohorts. Table 1 summarizes the average RPA with standard deviation as well as RPA range of intact C-peptide and variant II within the T2D and healthy cohorts. (Variants I and III were not tabulated in analogous manner because they were infrequently detected throughout both cohorts.) Interestingly, the relative contribution of variant II was found to be comparatively different between the T2D and the healthy cohorts. Figure 2 shows histograms comparing the frequency of occurrence between the two cohorts for the RPA of variant II. A broad distribution averaging 10% RPA (average of all individuals in the cohort) was observed for the T2D cohort, as compared with a narrow distribution averaging 5% RPA observed for the healthy cohort. Taking these differing variances into account, the average RPA for variant II in the T2D cohort was found by t-test to be significantly different from the healthy cohort at \( p < 0.01 \).

To verify repeatability of the assay, plasma from a single donor was divided into 96 aliquots and analyzed as described earlier. The average RPA value for unmodified C-peptide was 93±1.0% SD and the average RPA for variant II was 7.4±1.0% SD.

Given accurate mass determination (within 10 ppm), variants I, II, and III can be tentatively identified as truncated forms of C-peptide, respectively, lacking one, two, or three N-terminal amino acids \( \text{i.e.} \) C-peptide(2–31), C-peptide(3–31), and C-peptide(4–31)]. Because of its observed greater contribution in the diabetic population, MS/MS was used to confirm the identity of variant II \textit{via} partial sequencing. For reference, Fig. 3A shows a MALDI-TOF/TOF-MS spectrum obtained from intact C-peptide (positive mode \( m/z =3019.50 \)). Signals are observed to align predominantly with y-ions (with \( \sim 0.2 \text{ Da mass accuracy} \) yielding a fragmentation signature for intact C-peptide. Figure 3B shows an MS/MS obtained for variant II aligned to the sequence of C-peptide(3–31). Although of lower signal intensity, the y-ion fragmentation series registers the same pattern and mass as the intact C-peptide series (note the three overlaid enlarged glycine registers) which is expected for N-terminally truncated peptide forms. Notably, the signal for the b-ion of C-peptide(1–31) terminating at Q22 is observed at \( m/z \ 2093.1 \), but the b-ion of the putatively truncated C-peptide terminating at the same glutamine residue is shifted in mass by \( \sim 200.2 \text{ Da} \), which corresponds to the mass difference associated with the loss of the N-terminal Glu-Ala dipeptide. Together, the affinity extraction, mass mapping, and MS/MS data firmly support the conclusion that variant II is C-peptide(3–31).

Previous to this study, another C-peptide variant – in the form of C-peptide(1–26) – had been reported in the literature. This variant form was first discovered in rats but was later found to exist at very low relative abundance in humans. Furthermore, it was not observed to increase in individuals with T2D or insulinoma [30, 31]. To our knowledge, however, N-terminally truncated forms have yet to be reported or studied. The consequences and utility of the truncated variants are thus not understood, but are intriguing—particularly, the observation of elevated levels of C-peptide(3–31) in the diabetic population. Structural and biophysical characterization of DPP-IV (DPP-IV; EC 3.4.14.5) [32], a multifunctional type II transmembrane serine peptidase, reveals a preferred enzyme specificity of Xaa-Pro or

Proteomics Clin Appl. Author manuscript; available in PMC 2013 September 04.
Xaa-Ala from the N-termini of peptide hormones. Moreover, DPP-IV is a critical enzyme in regulating physiological glucose homeostasis via N-terminal cleavage of peptide hormones such as GLP-1 and glucose-dependent insulinotropic polypeptide [26]. As such, control of DPP-IV activity via drug formulations is a major focus of new T2D drug development [33–35]. Observation of C-peptide(3–31) is thus consistent with DPP-IV cleavage at the N-terminal Glu-Ala- of C-peptide, and the elevated levels observed in the diabetic population may be explained by abnormal DPP-IV activity in the diabetic population. If C-peptide is indeed an in vivo substrate for DPP-IV, these results are in alignment with previous work that showed an increase in DPP-IV activity in T2D individuals with moderate to severe hyperglycaemia relative to newly diagnosed diabetics, impaired glucose tolerance subjects, and normal glucose tolerance subjects [27]. Consequently, the measurement of C-peptide(3–31) as a companion marker to C-peptide is a potential avenue to improve the already gold-standard ID-MS assays, and, moreover, the measurement of both C-peptide forms may offer an assay simultaneously reflective of insulin production and modulation of DPP-IV activity during therapy. Further investigations in larger clinically well-defined healthy and diabetic cohorts are necessary to gauge the relative behavior of C-peptide and C-peptide(3–31) and their individual implications in assay improvement and disease stratification.

Acknowledgments

Funding was provided by Arizona State University’s Technology and Research Initiative Fund and by the Agilent Technologies Foundation in the form of Research Project Gift 08US-422UR.

Abbreviations

- DPP-IV: dipeptidyl peptidase IV
- GLP-1: glucagon-like peptide-1
- ID-MS: isotope-dilution MS
- MSIA: MS immunoassay
- RPA: relative percent abundance
- T2D: type 2 diabetes

References


Clinical Relevance

Clinical assays based solely on antibody recognition often lack the specificity to account for protein/peptide microheterogeneity. As such, indeterminate peptide variants encountered throughout the general population represent biological “noise” that may have a negative technical impact on the quantitative analysis of a peptide, potentially contributing to the mechanical failure of a conventional assay or producing otherwise misleading results. B-type natriuretic peptide represents a striking case in point [36]. Thus, for the sake of biomedical accuracy, there is a pressing clinical need to catalog the populational presence of variant forms of clinically assayed peptides and proteins such as C-peptide.
Figure 1.
C-peptide MSIA spectrum qualitatively representative of the subjects in this study. The mass spectrum contains the intact form of C-peptide as well as C-peptide (2–31) (calc. monoisotopic $m/z$ 2888.47), C-peptide (3–31) (calc. monoisotopic $m/z$ 2817.44), and C-peptide (4–31) (calc. monoisotopic $m/z$ 2688.39). These N-terminally truncated variants had signals registering at $m/z$ = 2888.49, 2817.45, and 2688.41 (labeled as I, II, and III, respectively). The loss of ammonia from the N-terminus is observed on all native and variant forms and is frequently observed under reflector-TOF-MS conditions.
Figure 2.
Expanding on the last row of Table 1, these histograms provide the frequency of occurrence within the healthy and T2D populations for the RPA of variant II. A broad distribution averaging 10% RPA (average of all individuals in the cohort) was observed for the T2D cohort, as compared with a narrow distribution averaging 5% observed for the healthy cohort.
Figure 3.
(A) MALDI-TOF/TOF-MS spectrum of intact C-peptide (positive mode $m/z = 3019.50$; calc. mono-isotopic $m/z = 3019.52$). (B) MALDI-TOF/TOF-MS spectrum of variant II aligned to the sequence of C-peptide(3–31). The y-ion fragmentation series registers as the same pattern and mass as the intact series (note the three overlaid enlarged glycine registers), which is expected for N-terminally truncated peptide forms. The signal for the b-ion of C-peptide(1–31) terminating at Q22 is observed at $m/z = 2093.1$, but the b-ion of the putatively truncated C-peptide terminating at the same glutamine residue is shifted in mass by ~200.2 Da, which corresponds to the mass difference associated with the loss of the N-terminal Glu-Ala dipeptide.
Table 1

Average RPA and RPA range of C-peptide(1–31) and C-peptide(3–31) for the T2D and healthy cohorts

<table>
<thead>
<tr>
<th>Variant</th>
<th>T2D</th>
<th></th>
<th>Healthy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average RPA (%)</td>
<td>RPA range (%)</td>
<td>Average (%)</td>
<td>RPA range (%)</td>
</tr>
<tr>
<td>C-peptide (1–31)</td>
<td>90±12</td>
<td>20–100</td>
<td>95±4</td>
<td>68–99</td>
</tr>
<tr>
<td>C-peptide (3–31)a)</td>
<td>10±12b)</td>
<td>0–80</td>
<td>5±4</td>
<td>1–32</td>
</tr>
</tbody>
</table>

RPA is obtained following peak area integration by dividing each individual peak area by the sum of all peak areas representing C-peptide and multiplying by 100. RPA range provides the low and high values observed for each variant in the two 48 sample cohorts described here.

a) The population distribution of C-peptide(3–31) for the RPA ranges of healthy and diabetic individuals is shown in Fig. 2.

b) Statistically significant from the healthy cohort at p<0.01.