Identification of hybrid insulin peptides (HIPs) in mouse and human islets by mass spectrometry

*T. Aaron Wiles†, Roger Powell†, Cole Michel†, Scott Beard‡, Anita Hohenstein†□, Brenda Bradley□, Nichole Reisdorph†, Kathryn Haskins□, Thomas Delong†**

†Department of Pharmaceutical Sciences, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus
‡Barbara Davis Center for Childhood Diabetes
□Department of Immunology and Microbiology, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO

Table of Contents

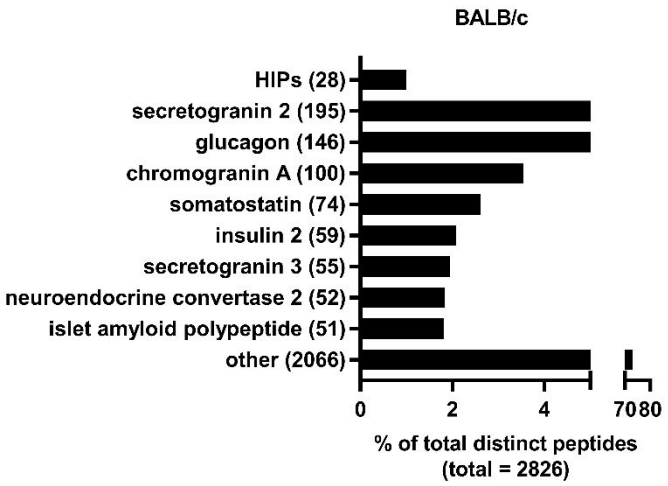


Figure S-1: Summary of peptides identified in BALB/c mouse islets by mass spectrometry. Islets were harvested and pooled from several BALB/c mice. Islets were processed and proteins were digested with AspN. The resultant peptide mixtures were analyzed by LC-MS/MS. Data were searched against a standard proteome database then against a custom HIP database. Proteins were ranked by the number of distinct peptides identified for the protein and the top eight proteins are reported. HIPs constituted less than 1% of the total number of distinct peptides identified. MS/MS spectra matched to HIPs were later subjected to further analysis to assess the validity of each match. Data are from one experiment.

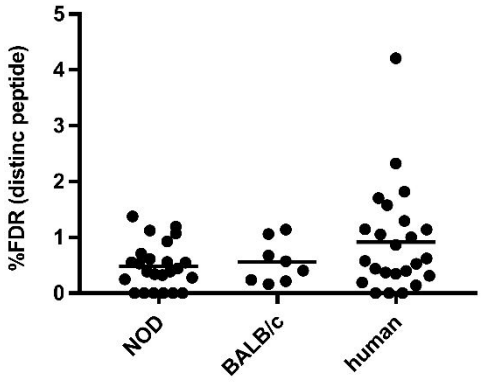


Figure S-2: False discovery rates (FDRs) for database searches. A reverse sequence database search was included in all searches and results were used to determine an FDR for each search. FDRs at the distinct peptide level (rather than the spectrum level) are reported for analyses of NOD, BALB/c, and human islets. Each data point represents the combined FDR for the SwissProt search and the HIP database search for a single sample. Lines indicate the mean %FDR for searches within each group.

TABLE S-1: VALIDATION OF PUTATIVE HIPS IDENTIFIED IN THE ISLETS OF BALB/C MICE.

		endogenous peptide											synthetic peptide			
		Rule 1[b]			**Rule 2[b]**			**Rule 3[c]**	**Rule 4[d]**		**Rule 5[e]**		**Rule 6[f]**			**Rule 7[g]**
		SwissProt database			HIP database				# of residues		b/y-ion coverage		spectral similarity			ΔRT
Sequence[a]	R pep	%SPI	score	R1-R2	%SPI	score	R1-R2	AspN	L pep	R pep	L pep	R pep	%SPI	score	PCC	(min)
EVEDPQVAQLELGGGPGAGDLQTLALEVAQQ																
EVED-LPVNSPMTKG	ChgA	61.4	7.9	1.3	72.3	13.4	5.5	Yes	4	10	2	3	62.9	18.2	0.80	0.13
EVED-TPVRSGSNPQM	IAPP	36.2	5.7	0.7	81.6	16.9	10.8	Yes	4	11	3	5	n.d.	n.d.	n.d.	n.d.
D-LPVNSPMTKG	ChgA	80.0	9.8	0.3	87.6	14.0	4.2	Yes	1	10	0	5	n.d.	n.d.	n.d.	n.d.
D-TPVRSGSNPQM	IAPP	64.1	5.9	1.2	83.4	12.3	6.4	Yes	1	11	0	6	n.d.	n.d.	n.d.	n.d.
DPQVAQLELGG-EVEDPQVAQLELGGGPGAG	C-pep	54.7	8.2	2.2	89.8	18.8	5.9	Yes	11	19	8	8	n.d.	n.d.	n.d.	n.d.
DLQTL-WSRM	ChgA	63.0	8.4	0.6	82.7	15.4	7.0	Yes	5	4	4	3	88.7	23.4	0.93	0.00
DLQTLAL-EVE	C-pep	80.0	10.2	0.6	98.1	14.6	4.5	Yes	7	3	6	3	97.2	24.0	0.98	n.d.
DLQTLAL-NAAG	IAPP	61.5	5.7	0.2	92.0	11.6	5.9	Yes	7	4	5	3	n.d.	n.d.	n.d.	n.d.
DLQTLAL-WSRM	ChgA	48.1	6.5	0.2	93.6	17.8	11.3	Yes	7	4	6	3	85.0	21.4	0.91	0.39
EVEDPQVAQLELGGGPGAGDLQTLALEVAQQ																

Results are from one experiment. (a) Insulin 2 C-peptide sequence is shown in white letters as a reference. (b) Best match is reported for the SwissProt database search (methionine oxidation considered as a variable modification) and the HIP database search. No digest was specified for either search. Thresholds: %SPI > 70%, score > 10, R1-R2 > 2.5. To satisfy rule 1, the best match from the SwissProt database search was required to fall below at least one threshold. The HIP match was required to surpass all three thresholds to satisfy Rule 2. (c) The match was required to be a sequence predicted for an AspN digest. A maximum of two missed cleavages were allowed. Only those matches that satisfied this criterion are shown in the table. (d) More than one residue required for each side of the proposed HIP. (e) Two or more b- and/or y-ions corresponding to fragmentation within the left peptide region and two or more corresponding to fragmentation within the right peptide region were required. If more than one detected ion corresponded to fragmentation at the same peptide bond, only one was counted. Fragmentation of the peptide bond at the left peptide-right peptide junction was counted for both L pep and R pep coverage. (f) Spectral similarity was assessed using the Spectrum Mill Spectrum Matcher utility. %SPI > 80 and score > 20 were required. The Pearson Correlation Coefficient (PCC) was calculated using GraphPad Prism software; a PCC > 0.9 was required for validation. (g) All samples and standards were spiked with a retention time calibration mixture containing forty standard peptides and observed retention times for the islet samples and synthetic peptide runs were aligned based on the retention times of these standards. Delta retention time threshold was 0.5 minutes. L pep = left peptide. R pep = right peptide. SPI = scored peak intensity. R1-R2 = rank1 minus rank2 score. RT = retention time. C-pep = insulin C-peptide. ChgA = chromogranin A. IAPP = islet amyloid polypeptide. n.d. = not determined. Values that satisfied the corresponding validation thresholds are highlighted in green, and those that did not are highlighted in yellow. For matches that validated across all categories, the HIP sequence is highlighted in green.

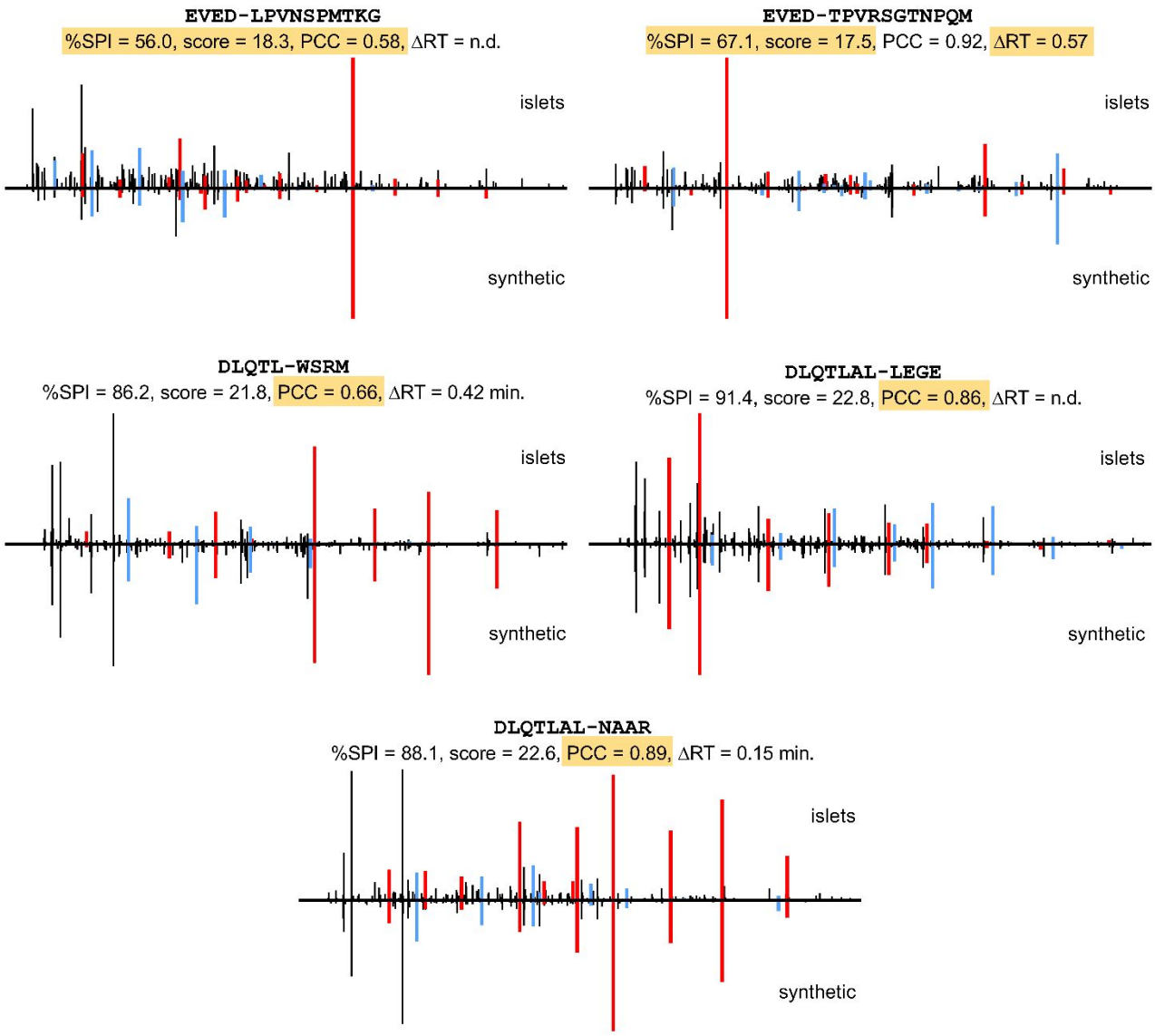


Figure S-3: Additional mirror plots demonstrating validation of HIPs identified in NOD islets. Following identification of potential HIPs in NOD islets, the putative HIP sequences were synthetized and the synthetic peptides were analyzed by tandem MS. Mirror plots comparing the MS/MS spectra for the endogenous islet peptide (positive y-axis) and the synthetic peptide standard (reflected y-axis) are shown. B- and y-ions (singly- and doubly-charged) are displayed in blue and red, respectively. A mass tolerance of +/- 20 ppm was used for labeling of b- and y-ions. Percent scored peak intensity (%SPI), score, and Pearson Correlation Coefficient (PCC) are indicated for each comparison. Values that did not satisfy criteria for validation are highlighted in yellow.

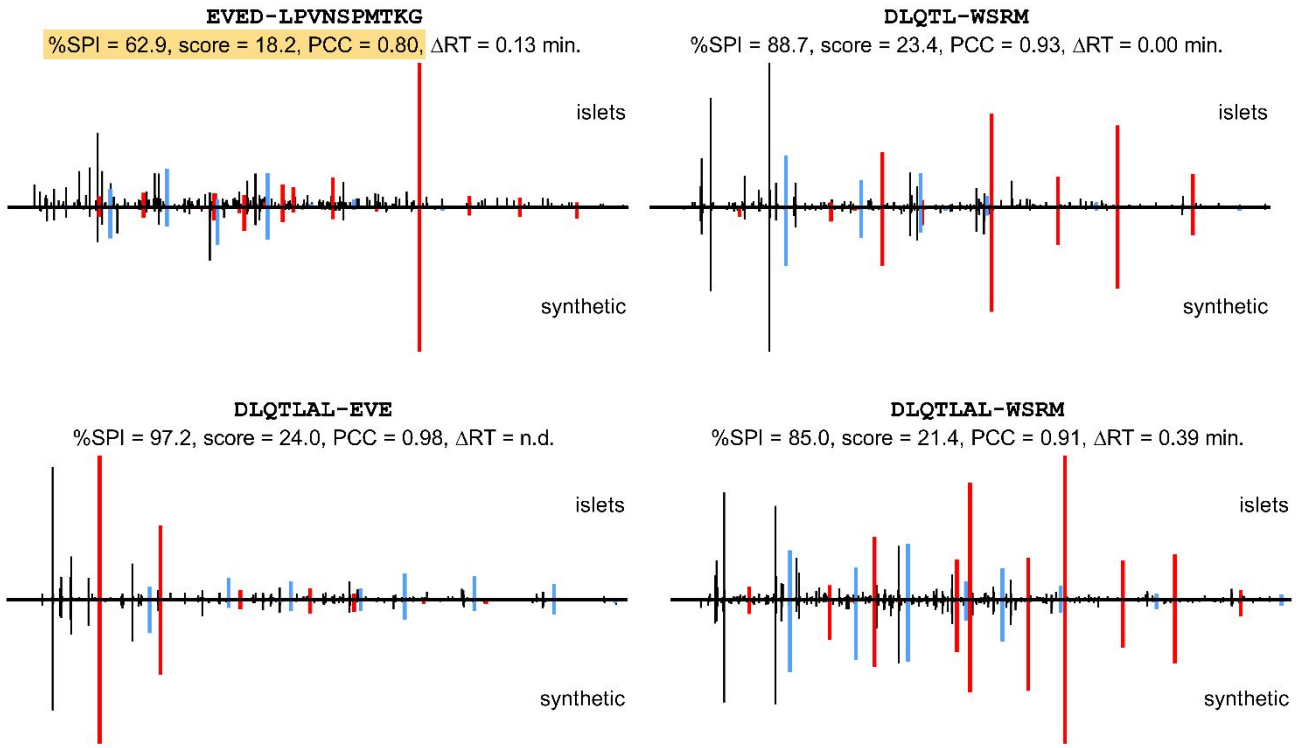
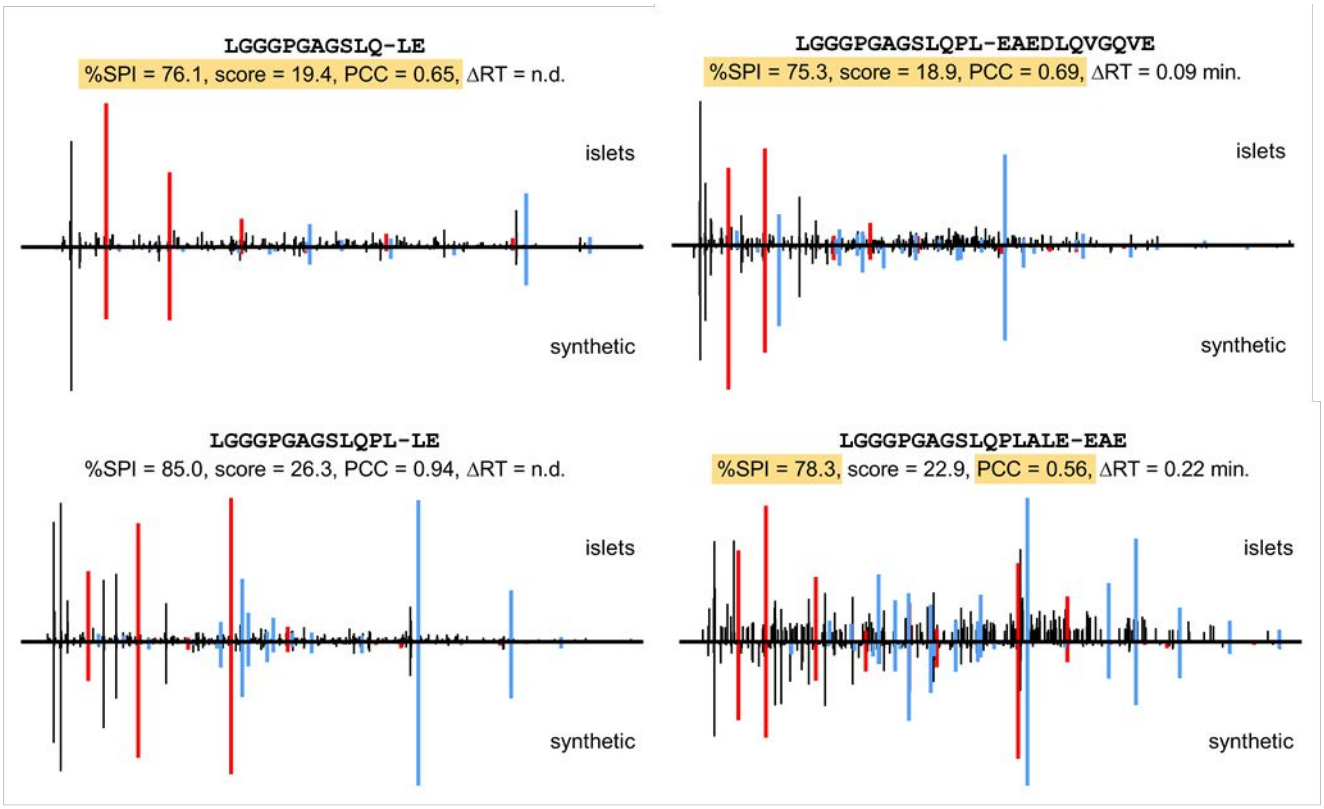


Figure S-4: Putative HIPs identified by mass spectrometry in islets from BALB/c mice validate with synthetic peptide standards. Following identification of potential HIPs in BALB/c islets, the putative HIP sequences were synthetized and the synthetic peptides were analyzed by tandem MS. Mirror plots comparing the MS/MS spectra for the endogenous islet peptide (positive y-axis) and the synthetic peptide standard (reflected y-axis) are shown. B- and y-ions (singly- and doubly-charged) are displayed in blue and red, respectively. A mass tolerance of +/- 20 ppm was used for labeling of b- and y-ions. Percent scored peak intensity (%SPI), score, and Pearson Correlation Coefficient (PCC) are indicated for each comparison. Values that did not satisfy criteria for validation are highlighted in yellow.



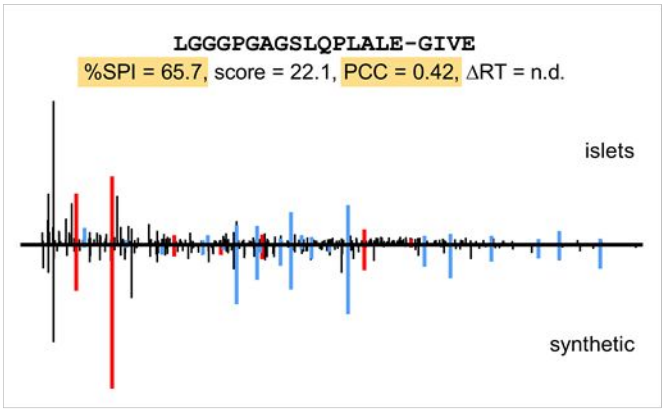


Figure S-5: Additional mirror plots demonstrating validation of HIPs identified in human islets. Following identification of potential HIPs in human islets, the putative HIP sequences were synthetized and the synthetic peptides were analyzed by tandem MS. Mirror plots comparing the MS/MS spectra for the endogenous islet peptide (positive y-axis) and the synthetic peptide standard (reflected y-axis) are shown. B- and y-ions (singly- and doubly-charged) are displayed in blue and red, respectively. A mass tolerance of +/- 20 ppm was used for labeling of b- and y-ions. Percent scored peak intensity (%SPI), score, and Pearson Correlation Coefficient (PCC) are indicated for each comparison. Values that did not satisfy criteria for validation are highlighted in yellow.

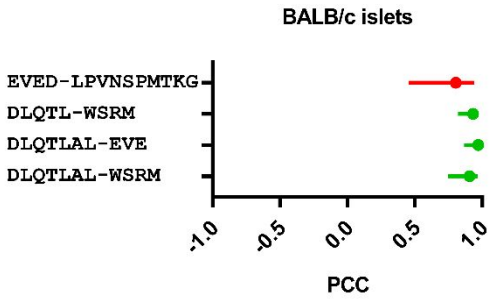


Figure S-6: Pearson correlation coefficients (PCC) obtained by comparison of mass spectra of endogenous peptides and synthetic HIPs. Fragmentation spectra for putative HIPs identified in BALB/c islets were compared to spectra obtained by analysis of synthetic versions of the peptides. For each comparison, peaks with an abundance greater than three standard deviations above the average in either spectrum were included. Error bars indicate 95% confidence intervals. Values that surpassed the PCC validation threshold of 0.9 are shown in green and those that did not are shown in red.