Poster Presentation

7th Symposium of the Protein Society - July, 1993

Poster #291-M

Identification of Tryptic Peptides After HPLC Separation Using Spectral Libraries From High-Resolution Photodiode Array (PDA) Analysis

P. Young & M. Gorenstein
Millipore Corporation, Waters Pharmaceutical Division
34 Maple Street, Milford, MA. 01757 U.S.A.

MILLIPORE

Waters Chromatography Division
34 Maple Street
Milford, MA 01757
508 478-2000

ABSTRACT

Protein integrity is routinely determined by subjecting the intact molecule to enzymatic digestion and resolving the resultant peptides by reversed phase HPLC. Typically, identification is based on retention time with detection at a single low-ultraviolet wavelength (214 nm). Use of PDA detection combined with narrow-bore HPLC separations provides high-sensitivity spectral data regarding peptide integrity. Additional spectral information and detection sensitivity can be obtained through use of hydrochloric acid, an optically clear mobile phase modifier. Alterations in peptide structure, such as deamidations or residue substitutions, can be easily and accurately detected using software-driven, high resolution spectral algorithms. On-line PDA detection provides a reliable tool to assist the researcher in better characterizing tryptic-derived and synthetic peptides.

INTRODUCTION

- Enzymatic digestion followed by peptide mapping is normally the first step in characterization of complex proteins.
- The standard method for resolution of the complex mixture of peptides is reversed phase HPLC in shallow gradients of increasing acetonitrile with trifluoroacetic acid (TFA) as an ion-pair reagent.
- Typically, peptides are detected by the of the absorbance of the peptide bond at <220 nm.
- With single wavelength detection, retention time is the sole method of peptide identification.
- Monitoring a separation at 280 nm can detect the presence of aromatic compounds.
- Photodiode array (PDA) detection provides full-spectrum data that can more fully characterize peptides when combined with peak purity and spectral matching capabilities.
- Due to the increased optical clarity of hydrochloric acid (HCl) in the low UV, HCl can provide increased detection sensitivity.
- Since HCl is effective < 200 nm, full spectral data can be acquired that permits qualitative identification of peptides differing by a single amino acid.

Materials and Methods

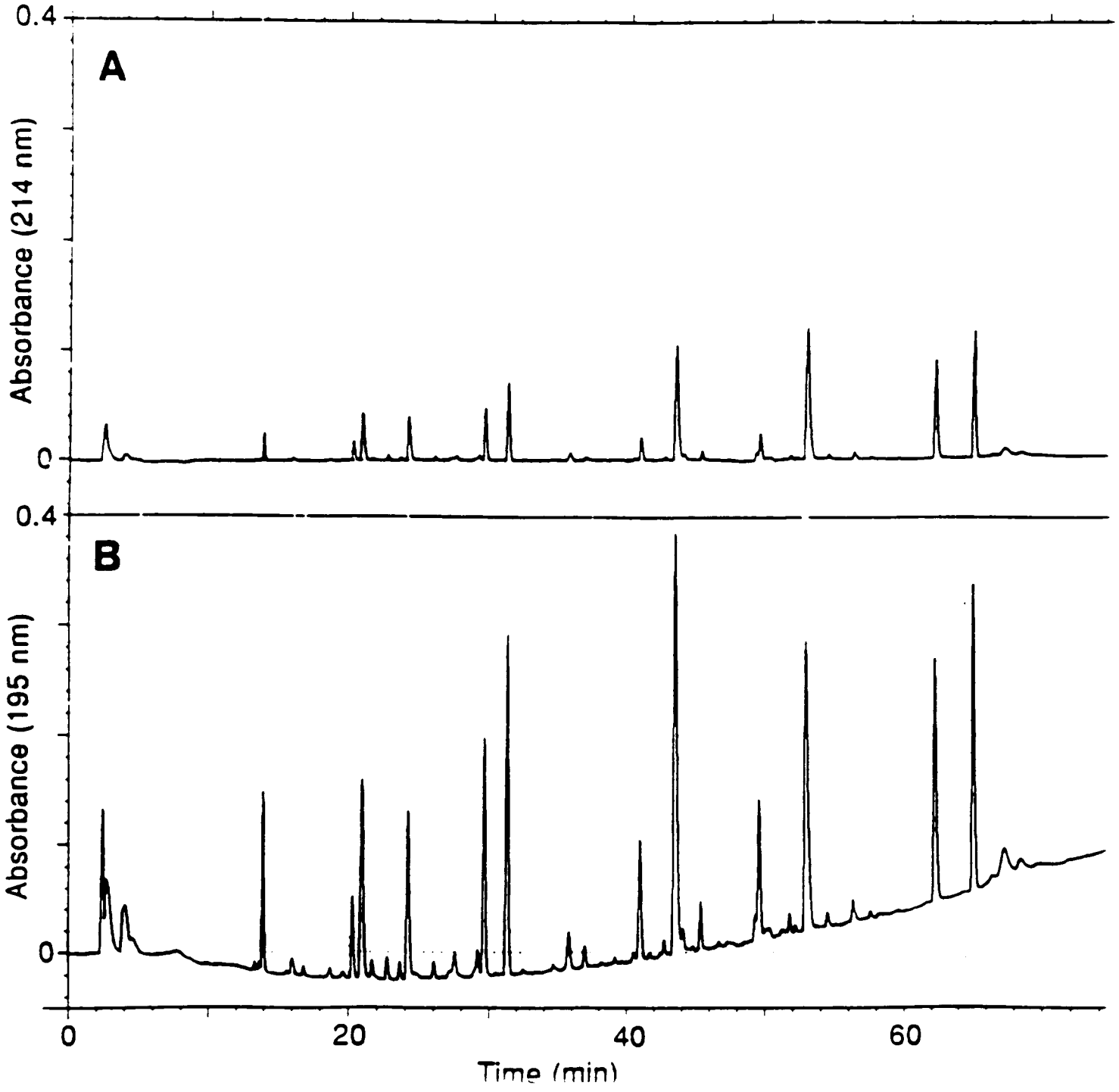
Tryptic Digestion

Cytochrome *c* (Sigma) from bovine and chicken heart mitochondria (1 mg/1 ml) was dissolved in 0.1M ammonium bicarbonate (Sigma) buffer, pH 8.0. TPCK-treated trypsin (Worthington Biochemical) was added at a 1:10 weight ratio with respect to cytochrome *c*. The digestion mixture was incubated for 24 h at 37°C. Following incubation, trypsin was deactivated by heating at 100°C for 5 min. Aliquots of 100 µl were frozen (-20°C). Prior to HPLC separation, the digests were diluted 1:10 with aqueous TFA.

HPLC Conditions

System:	Waters BioDiscovery with 625 LC System
Detection:	Photodiode Array (Waters model 996) 190-425 nm at 1.2 nm resolution
Sample:	Tryptic Digest of Bovine or Chicken Cytochrome c (400 pmol)
Column:	Delta-Pak HPI C_{18}, 5 µm, 300 A (2 mm x 150 mm)
Eluants:	A= water/ 6 mM HCl B= acetonitrile/ 6 mM HCl
Gradient:	0-60% B / 120 min
Flow rate:	180 µl/ min
Temperature:	35°C

Enhanced sensitivity at 195 nm with HCl-based mobile phases. The tryptic digest of bovine cytochome c was separated as described in the methods section. Full spectral data was acquired from 190-425 nm at 1.2 nm resolution. Two wavelength extractions at 214 nm (panel A) and 195 nm (panel B) illustrate a 4-fold increase in peak heights at 195 nm when compared to detection at 214 nm. Increased dynamic range in the low UV can be used for high resolution spectral characterization.



Amino Acid Sequence for Tryptic Peptides derived from Bovine and Chicken Cytochrome *c*

Peptide Number	Sequence	Bovine	Chicken
1	1-5 *	GD**V**EK	GD**I**EK
2	74-79	YIPGTK	YIPGTK
3	73-79	KYIPGTK	KYIPGTK
4	9-13	IFVQK	IFVQK
5	39-53 *	KTGQA**P**GFSYTDANK	KTGQA**E**GFSYTDANK
6	40-53 *	TGQA**P**GFSYTDANK	TGQA**E**GFSYTDANK
7	92-97 *	**E**DLIAY	**V**DLIAY
8	80-86	MIFAGIK	MIFAGIK
9	28-38	TGPNLHGLFGR	TGPNLHGLFGR
10	92-99 *	**E**DLIAYLK	**V**DLIAYLK
11	14-22 *	C**A**QCHTVEK	C**S**QCHTVEK
12	56-73 *	GITWGE**E**TLMEYLENPKK	GITWGE**D**TLMEYLENPKK
13	56-72 *	GITWGE**E**TLMEYLENPK	GITWGE**D**TLMEYLENPK

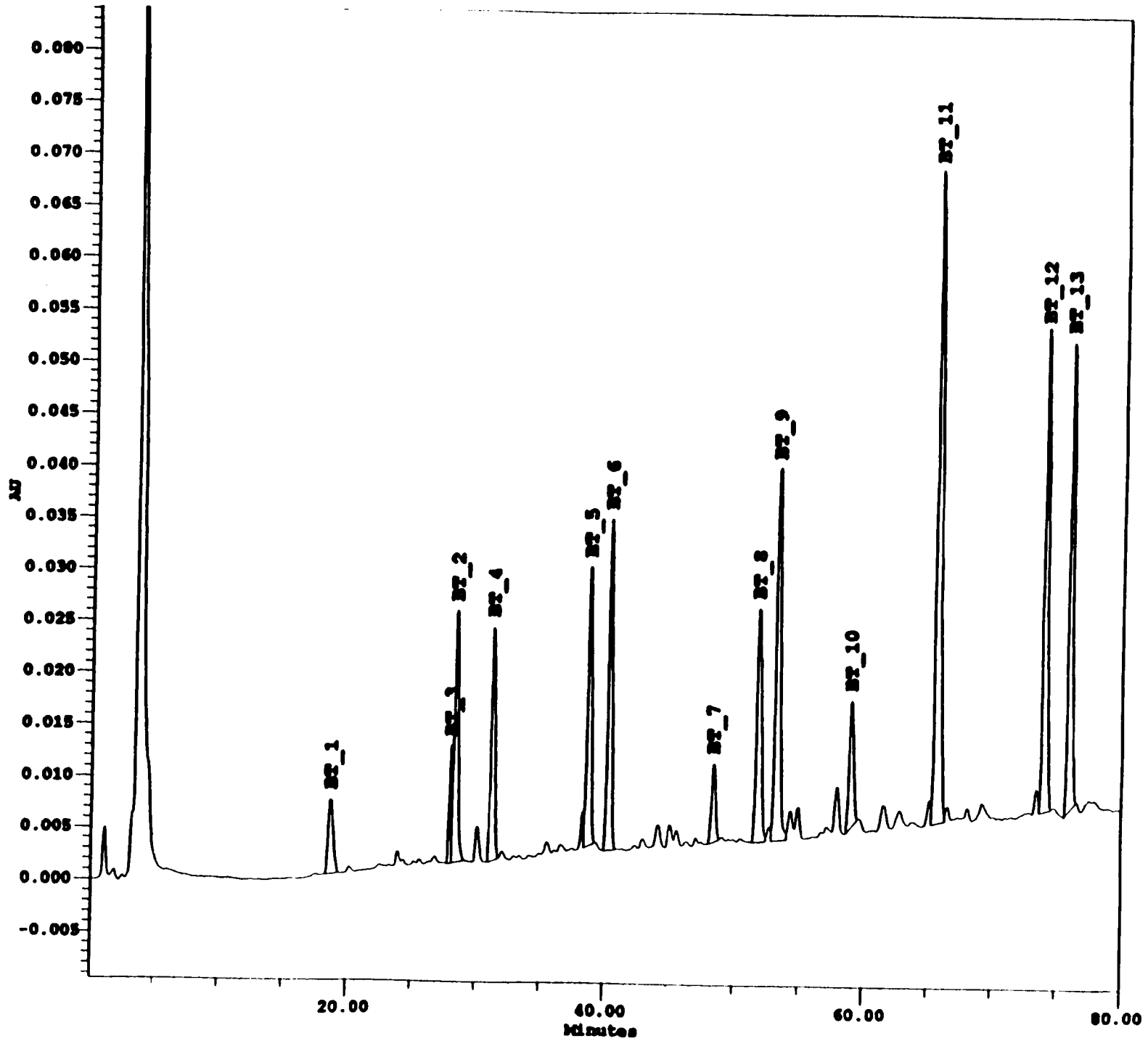
* Denotes peptides having amino acid substitutions; the amino acid difference is shown in bold-type.

While bovine and chicken cytochrome c each have 104 amino acids, the molecules differ at seven sites. These amino acid substitutions affect peptide sequences marked with an *. For more information regarding the alterations in reversed phase retention times, consult the following publication:

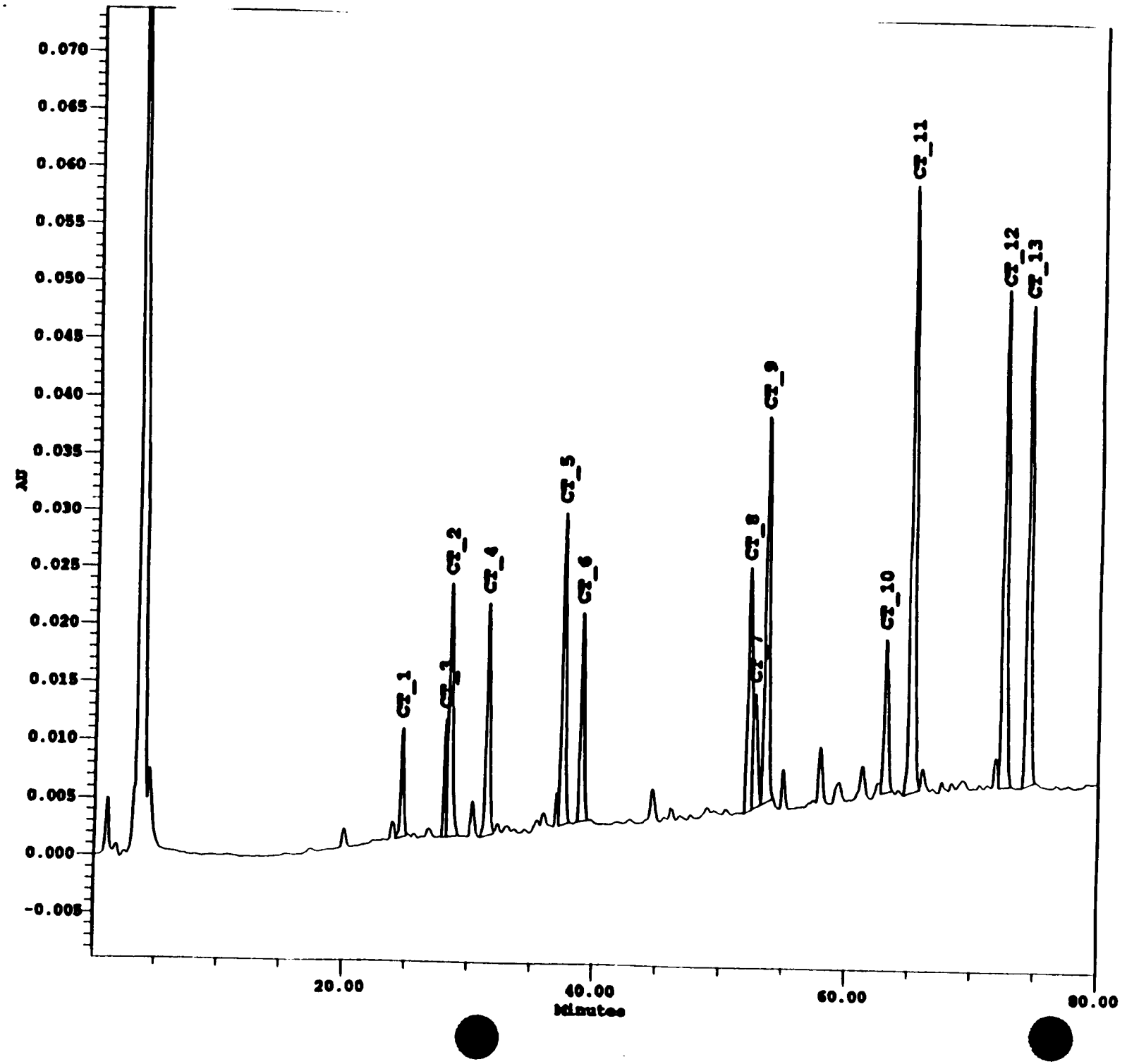
Young, P. M., et al. "Effects of pH and Buffer Compositions on Peptide Separations by HPLC and Capillary Electrophoresis," *LC·GC* 10:26-32, 1992.

Separation of Tryptic Digest o[redacted]ovine Cytochrome c in HCl-based Mo[redacted] Phases

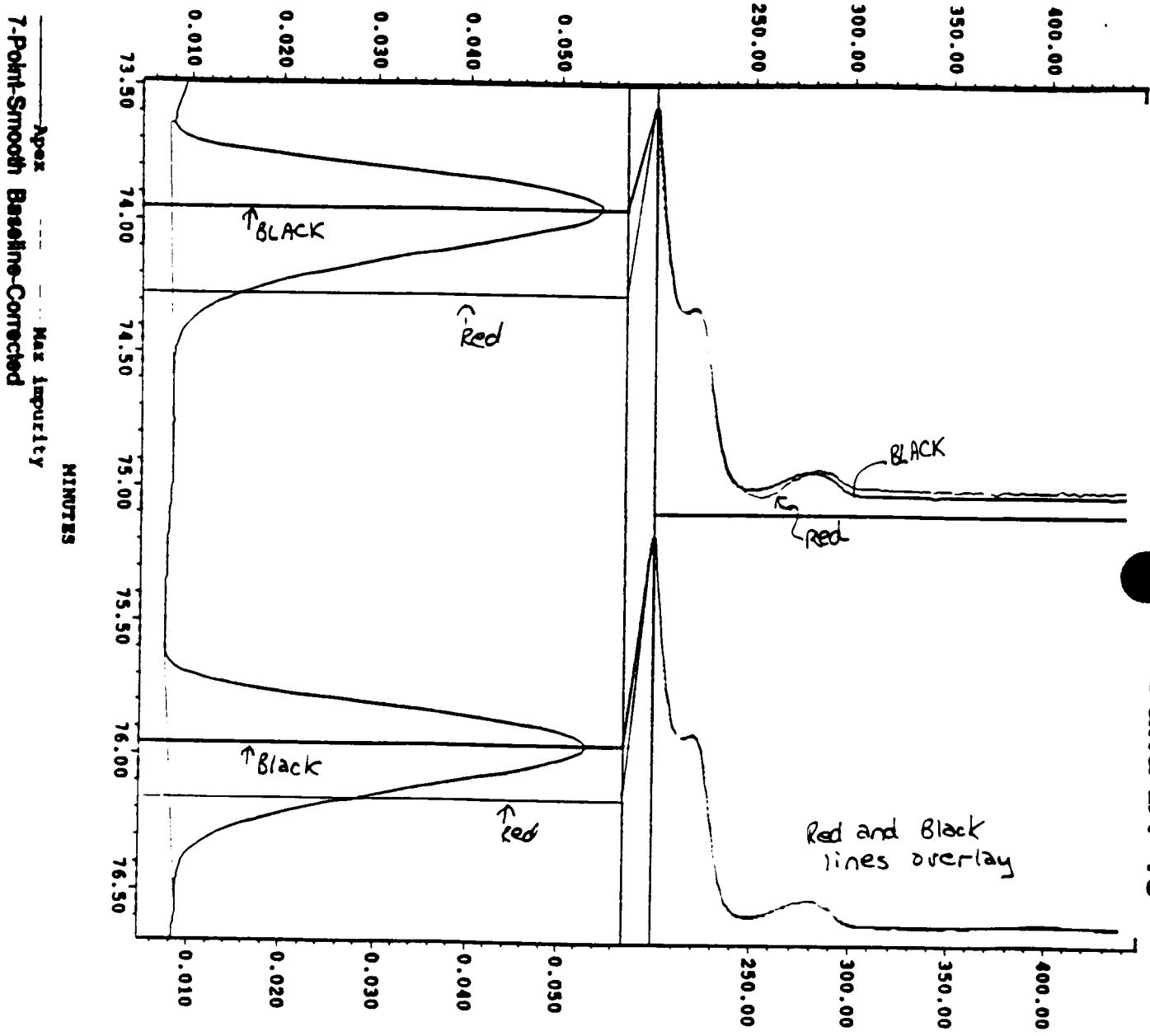
Millennium Results Report: Big_Plot_cc Proc Chan: 214nm_PDA Page 2
For Sample: HCL_BT34A_400 Vial: 5 Inj: 1 Chan: 996 Date Processed 07/01/93 04:18 PM
Processing Method BT_CT_HCL_bmth Project Name SP3_CC_TFA_HCL
Channel Descr: PDA 214.0 nm



Separation of Tryptic Digest of Chicken Cytochrome *c* in HCl-based Mobile Phases



Millennium Results Report: PDA_SIP_ONLY_COLOR Proc Chan: 214nm_PDA Page 1
For Sample: HCL_BT34A_400 Vial: 5 Inj: 1 Chan: 996 Date Processed 07/16/93 11:18 AM
Processing Method bt12_13only
Channel Descr: PDA 214.0 nm Project Name THREE_SP_ABER



Spectrum Index Plot of BT-12 and BT-13

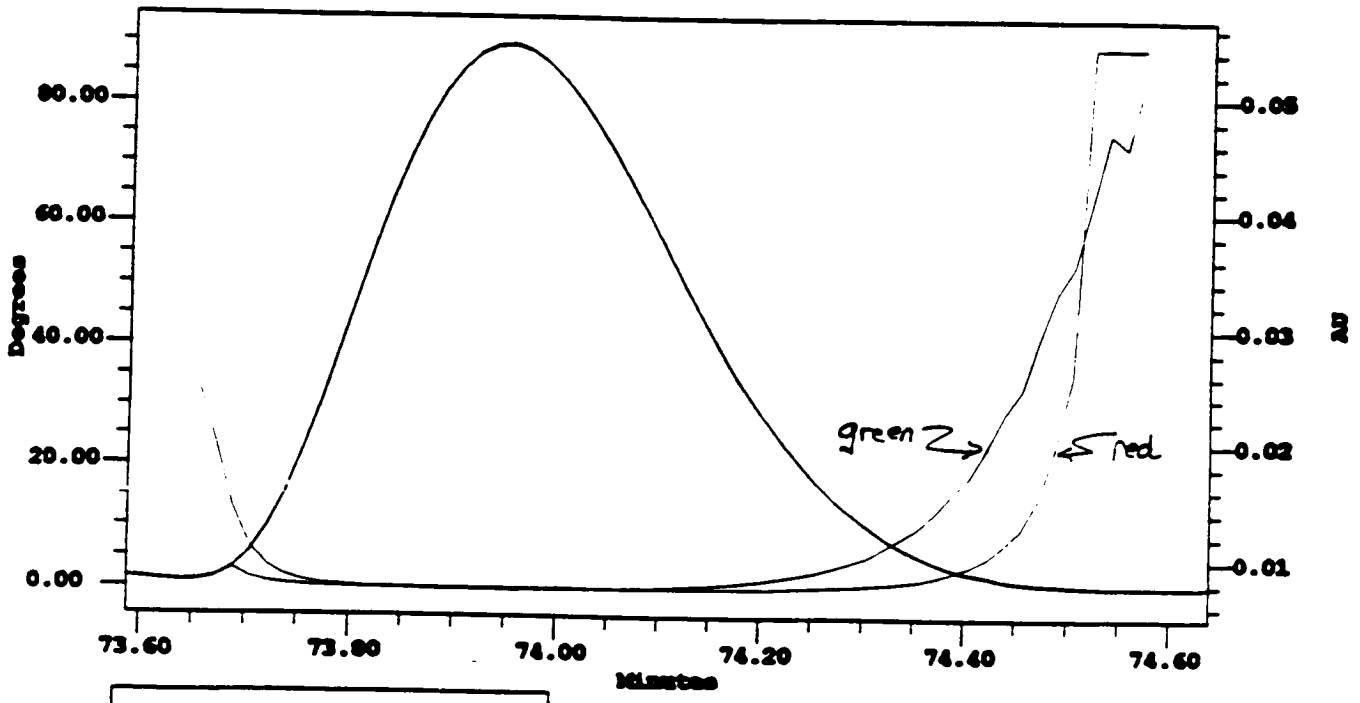
Peptides BT-12 (73.96 min) and BT-13 (75.96) that differ by a C-terminal Lys (K) are shown in the Spectrum Index Plot. In the reference chromatogram at 214 nm in the lower panel, the lines drawn in each peak indicate where two spectra have been taken. The apex spectrum (black) and the spectra calculated to contain the maximum impurity (red) are shown in the upper panel. As seen by the apex spectra, there is no visible difference between the spectra of these peptides. For BT-12, there appears to be an impurity near touchdown.

Purity Result for Peak 12: BT_12

Retention Time: 73.96 Purity Angle: 0.85 Threshold Angle: 1.19

Matching Spectra List

#	Match Angle	Match Threshold	Spectrum Name	Library Name	Ideal	Wvln Diff	Match Errors
1	0.310	0.640	BT12_HCL	BT_HCL	Yes	0.0000	None
2	0.816	0.640	BT13_HCL	BT_HCL	Yes	0.0000	None
3	1.154	0.666	CT12_HCL	CT_HCL	Yes	0.0000	None



Peak
Purity Green
Noise red

Purity Plot and Match Data for BT-12 (GITWGEETLMEYLENPKK)

Purity Plot and Match Data for BT-12 (GITWGEETLMEYLENPKK)

The spectrum for BT-12 was matched against libraries containing spectra from all the peptides in the bovine and chicken cytochrome c tryptic digests. Matching data for BT-12 are tabulated in rank order where the previously recorded BT-12 spectrum was selected (Match angle 1 < Match Threshold) as the best possible match from all the peptides in the bovine and chicken cytochrome *c* tryptic digest libraries. The second-level match (Match Angle 2 > Match Threshold) is for BT-13, differing from BT-12 by a C-terminal Lys (K). The third best match for BT-12 is CT-12 with a Match Angle 3 > match Threshold. These peptides differ at position 62 where BT-12 has a Glu (E) and CT-12 has an Asp (D).

The Purity plot indicates reference peak at 214 nm (black), the Noise (red) and Purity (green). When the Purity Plot exceeds the Noise Plot, that section of the peak is considered spectrally inhomogeneous. This can be seen in the area of the peak near touchdown. This is consistent with the data shown in the Spectrum Index Plot.

The Mathematics Behind the Peak Purity Plot

The calculations that lead to a Peak Purity Plot are based on complex vector algebra. The basic principles behind these calculations can be described as follows:

1. Each spectrum within a peak is baseline corrected. This is done by interpolating liftoff and touchdown spectra to obtain a series of baseline spectra. The resultant baseline spectra are substracted from corresponding peak spectra.

2. Each baseline-corrected spectrum is converted into vectors.

3. Each baseline-corrected spectrum is compared to the apex spectrum by measuring the differences in the directions of the vectors. The resulting value is a Spectral Contrast Angle for *that particular timepoint on the peak* and is plotted as such on the graph. The Spectral Contrast Angle can be between 0° and 90°, where 0° implies virtual identity and 90° implies no correlation. These angles correspond roughly to an average fractional difference in peak shape. For instance, a 10° Spectral Contrast Angle would mean the spectra were, on average, different by 10%. The point where you can visually see a difference between spectra is at a 1° Spectral Contrast Angle.

4. The noise associated with each spectral comparison is similarly calculated and the resulting Noise Angle for *that specific comparison* is similarly plotted on the graph.

   If there is a trend across the peak where the Spectral Contrast Angle is significantly larger than the Noise Angle, there is reason to believe that the peak is inhomogeneous. To make the comparison even simpler, however, a Purity Angle is calculated for the entire peak and compared to a noise calculation called the Threshold Angle (the threshold at which signal can be separated from noise). The calculations are as follows:

5. The Purity Angle is the weighted sum (with other math operations) of the individual Spectral Contrast Angle comparisons across the peak. Spectra with more absorbancy are more highly weighted.

6. The Noise Angle is the sum of the individual noise angles across the peak. To this is added the Solvent Angle, a weighted measure of the spectral changes that can be attributed to elution conditions. The sum of the overall Noise Angle and the Solvent Angle equals the Threshold Angle.

The Purity Angle and the Threshold Angle are reported along with the Purity Plot. When the Purity Angle exceeds the Threshold Angle, the peak is not homogeneous. When the Purity Angle is less than the Threshold Angle, the peak can be considered spectrally pure.

Peak Purity and Spectral Matching of Bovine Cytochrome *c* Tryptic Peptides

Matched Against Tryptic Libraries of Bovine and Chicken Cytochrome *c*

#	Peak Name	R T (min)	Purity Angle	Purity Thresh.	Match1 Angle	Match1 Thresh.	Match1 Spect. Name	Match1 Lib. Name
1	BT_1	18.755	0.72	1.85	0.46	1.15	BT1_HCL	BT_HCL
2	BT_3	28.022	0.51	1.58	0.40	1.06	BT3_HCL	BT_HCL
3	BT_2	28.388	0.98	1.35	0.19	0.75	BT2_HCL	BT_HCL
4	BT_4	31.255	0.41	1.41	0.21	0.80	BT4_HCL	BT_HCL
5	BT_5	38.705	0.45	1.29	0.29	0.71	BT5_HCL	BT_HCL
6	BT_6	40.305	0.23	1.26	0.17	0.69	BT6_HCL	BT_HCL
7	BT_7	48.422	1.07	1.97	0.49	1.20	BT7_HCL	BT_HCL
8	BT_8	51.805	1.32	1.37	0.31	0.80	BT8_HCL	BT_HCL
9	BT_9	53.222	1.03	1.26	0.23	0.68	BT9_HCL	BT_HCL
10	BT_10	59.005	1.26	1.56	0.72	0.91	BT10_HCL	BT_HCL
11	BT_11	65.588	0.83	1.06	0.15	0.55	BT11_HCL	BT_HCL
12	BT_12	73.955	0.85	1.19	0.31	0.64	BT12_HCL	BT_HCL
13	BT_13	75.955	0.34	1.20	0.14	0.64	BT13_HCL	BT_HCL

#	Peak Name	R T (min)	Match2 Angle	Match2 Thresh.	Match2 Spect. Name	Match2 Lib. Name	Match3 Angle	Match3 Thresh.	Match3 Spect. Name	Match3 Lib. Name
1	BT_1	18.755	0.652	1.180	CT1_HCL	CT_HCL	9.600	1.046	BT8_HCL	BT_HCL
2	BT_3	28.022	0.766	1.195	CT_3_HCL	CT_HCL	2.449	0.970	BT2_HCL	BT_HCL
3	BT_2	28.388	0.307	0.810	CT_2_HCL	CT_HCL	2.532	0.905	BT3_HCL	BT_HCL
4	BT_4	31.255	0.296	0.868	CT4_HCL	CT_HCL	2.516	0.859	CT8_HCL	CT_HCL
5	BT_5	38.705	0.630	0.703	BT6_HCL	BT_HCL	2.087	0.814	CT6_HCL	CT_HCL
6	BT_6	40.305	0.594	0.697	BT5_HCL	BT_HCL	2.256	0.804	CT6_HCL	CT_HCL
7	BT_7	48.422	1.866	1.233	CT7_HCL	CT_HCL	3.490	1.111	BT10_HCL	BT_HCL
8	BT_8	51.805	0.622	0.862	CT8_HCL	CT_HCL	2.483	0.805	BT4_HCL	BT_HCL
9	BT_9	53.222	0.709	0.722	CT9_HCL	CT_HCL	2.998	0.740	BT4_HCL	BT_HCL
10	BT_10	59.005	2.597	1.081	CT7_HCL	CT_HCL	2.766	0.972	CT10_HCL	CT_HCL
11	BT_11	65.588	0.420	0.557	CT11_HCL	CT_HCL				
12	BT_12	73.955	0.916	0.640	BT13_HCL	BT_HCL	1.154	0.666	CT12_HCL	CT_HCL
13	BT_13	75.955	0.768	0.643	BT12_HCL	BT_HCL	0.960	0.669	CT12_HCL	CT_HCL

Peak Purity and Spectral Matching of Bovine Cytochrome *c* Tryptic Peptides

Matched Against Tryptic Libraries of Bovine and Chicken Cytochrome *c*

After HPLC separation of the tryptic peptides from bovine and chicken cytochrome *c*, the apex spectra from each peptide was stored in a library. Subsequent preparations of tryptic digests of bovine cytochrome *c* were chromatographed as decscribed. Resultant spectra were analyzed for purity and matching using the original libraries containing spectra from chicken and bovine cytochrome *c* tryptic peptides. Peaks are considered spectrally homogeneous when the purity angle < purity threshold. Similarly, when the match angle < match threshold, the peaks are considered good matches.

Data tabulated in Match 1 Angle, Match 1 Threshold, and Match 1 Spectral Name, indicate that the bovine peptides are identified by their distinctive spectral signature. Within the bovine peptides (BT), data tabulated clearly identify pairs varying by a single amino acid. BT-3 and BT-2, that differ by an N-terminal Arg (R) and pairs BT-5 / BT-6 and BT-12 / BT-13, that differ by C-terminal Lys (K), are correctly identified.

It is not until the Match 2 comparisons that chicken peptides (CT) are reported. The subtle difference between BT-1 (GDVEK) vs. CT-1 (GDIEK) was discernible by rank order of the spectral matching. Similarly, BT-5 was identified as a better second match to BT-6 than the either of the CT-5 / CT-6 pair. BT-5 and BT-6, while differing by an Arg both contain Pro (P) at position 44, while the corresponding CT-5 and CT-6 peptides have GLU (E) at position 44. CT-5 and CT-6 are reported in the Match 3 data with Match Angle > Match Threshold. BT-7 and CT-7 have a Glu (E) and Val (V) at position 92, respectively. CT-7 was selected from all the other tryptic peptides as the best second match to BT-7, but with a Match Angle > Match Threshold. Like peptide pairs BT-12 / BT-13, peptides CT-12 / CT-13 differ from each other by a C-terminal Lys (K). BT-12 / BT-13 have a Glu (E) at position 62 while CT-12 / CT-13 have an Asp (D). CT-12 and CT-13 are poorly matched for the corresponding BT-12 and BT-13 as seen in Match Angle 3 where CT-12 is identified with match angle > match threshold.

These data indicate that single amino acid substitutions can be discerned from spectral matching. Also, the matching algorithm can be used to screen a peptide separation for structurally similar peptides.

Conclusions

- Photodiode Array detection combined with high-resolution spectral libraries can be used to qualitatively identify peptides differing by a single amino acid. This amino acid difference can reside at the N-terminus (BT-5 /BT-6) , the C-terminus (BT-12 / BT-13) , or in the interior of the peptide (BT-12 / CT-13).

- Photodiode Array detection and the matching algorithm can augment the identification of routinely assayed peptides. This may prove beneficial to those researchers studying recombinant protein reproducibility and storage stability.