SUPPLEMENTARY DATA

Iodination of proteins by IPy_2BF_4, a new tool in protein chemistry

Gemma Espuña [‡], *David Andreu* [‡], *José Barluenga* [§], *Xavier Pérez*[‖], *Antoni Planas*[‖],

Gemma Arsequell[⊥], *Gregori Valencia**[,⊥].

[‡] Proteomics Unit, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Dr. Aiguader 80, 08003 Barcelona, Spain, [§]Instituto Universitario de Química Organometálica "Enrique Moles", Universidad de Oviedo, 33071 Oviedo, Spain, [‖] Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull, Via Augusta 390, 08017 Barcelona, Spain, and [⊥] Unit for Glycoconjugate Chemistry, IIQAB-CSIC, Jordi Girona 18-26, 08034 Barcelona, Spain.

Running title: Iodination of proteins by IPy_2BF_4

*Address correspondence to: Gregorio Valencia, Unit for Glycoconjugate Chemistry, IIQAB-CSIC, Jordi Girona 18-26, E08034 Barcelona, Spain. Tel. 34934006113; Fax: 34932045904, E-mail: gvpqbp@iiqab.csic.es

Keywords: Iodine, electrophilic iodination, IPy_2BF_4, mass spectrometry, protein modifications

SUPPLEMENTARY DATA

Iodination of Proteins by IPy_2BF_4, a New Tool in Protein Chemistry

Gemma Espuña, David Andreu, José Barluenga, Xavier Pérez, Antoni Planas, Gemma Arsequell, Gregori Valencia*

SUPPLEMENTARY MATERIAL

Iodination of Proteins by IPy_2BF_4, a New Tool in Protein Chemistry

Gemma Espuña, David Andreu, José Barluenga, Xavier Pérez, Antoni Planas, Gemma Arsequell, Gregori Valencia*

INSULIN. Supplementary material:

- **a) Insulin HPLC results**
- **b) MALDI-TOF-MS spectra**
- **c) Edman sequencing for iodinated insulin chain A**
- **d) Edman sequencing for iodinated insulin chain B**

LYSOZYME: Supplementary material.

- **A) Tryptic digestion: calculated values**
- **B) MS/MS sequence analysis:**

GLUCANASE: Supplementary material

- **A) Calculated values of recombinant 1,3-1,4-beta-D-glucanohydrolase tryptic peptides**
- **B) MALDI-TOF-MS spectrum of tryptically digested recombinant glucanase (after treatment with IPy_2BF_4)**

INSULIN. Supplementary material:

e) Insulin HPLC results

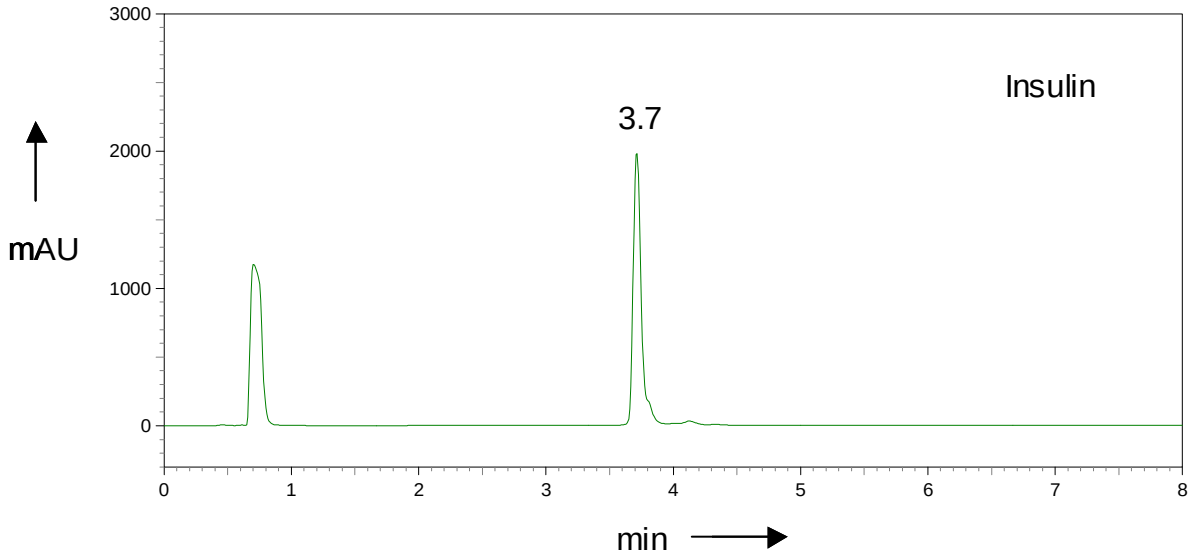
f) MALDI-TOF-MS spectra

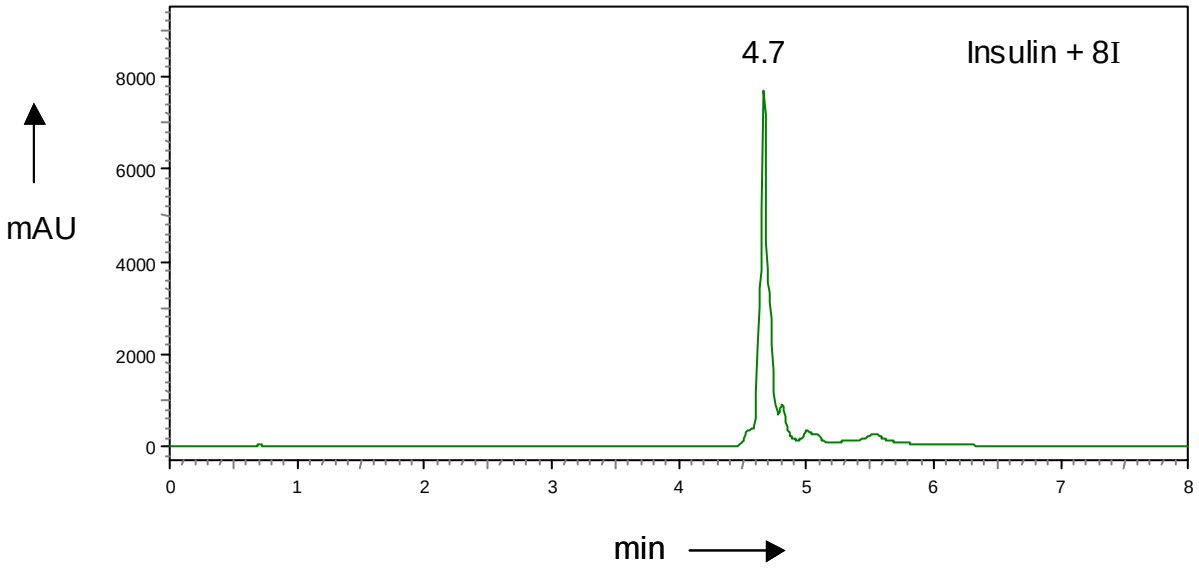
g) Edman sequencing for iodinated insulin chain A

h) Edman sequencing for iodinated insulin chain B

INSULIN HPLC RESULTS:

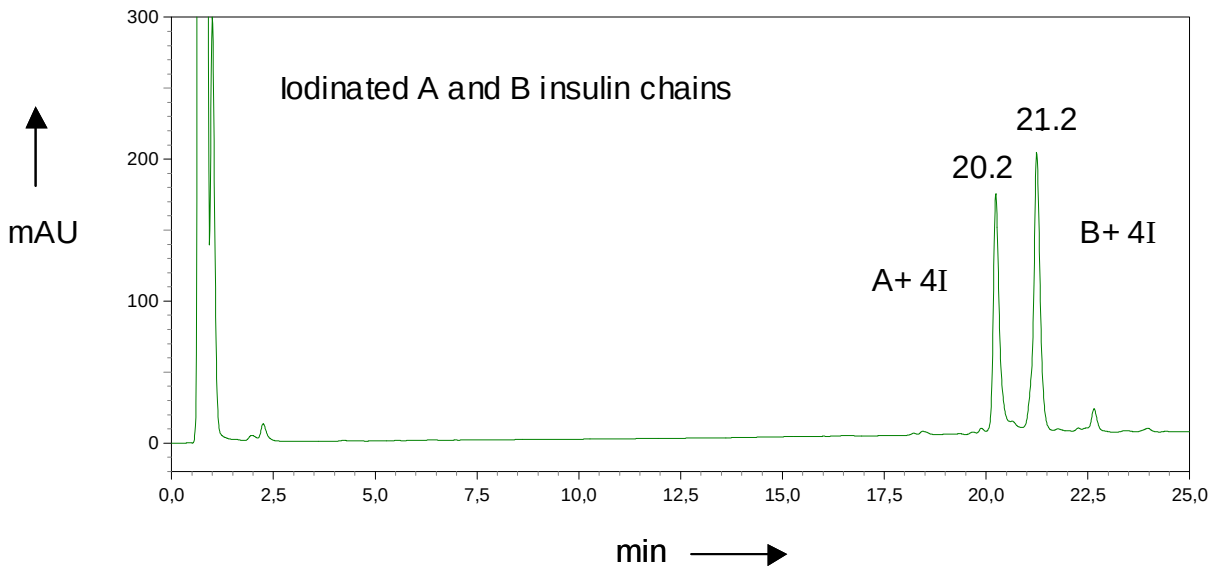
A)





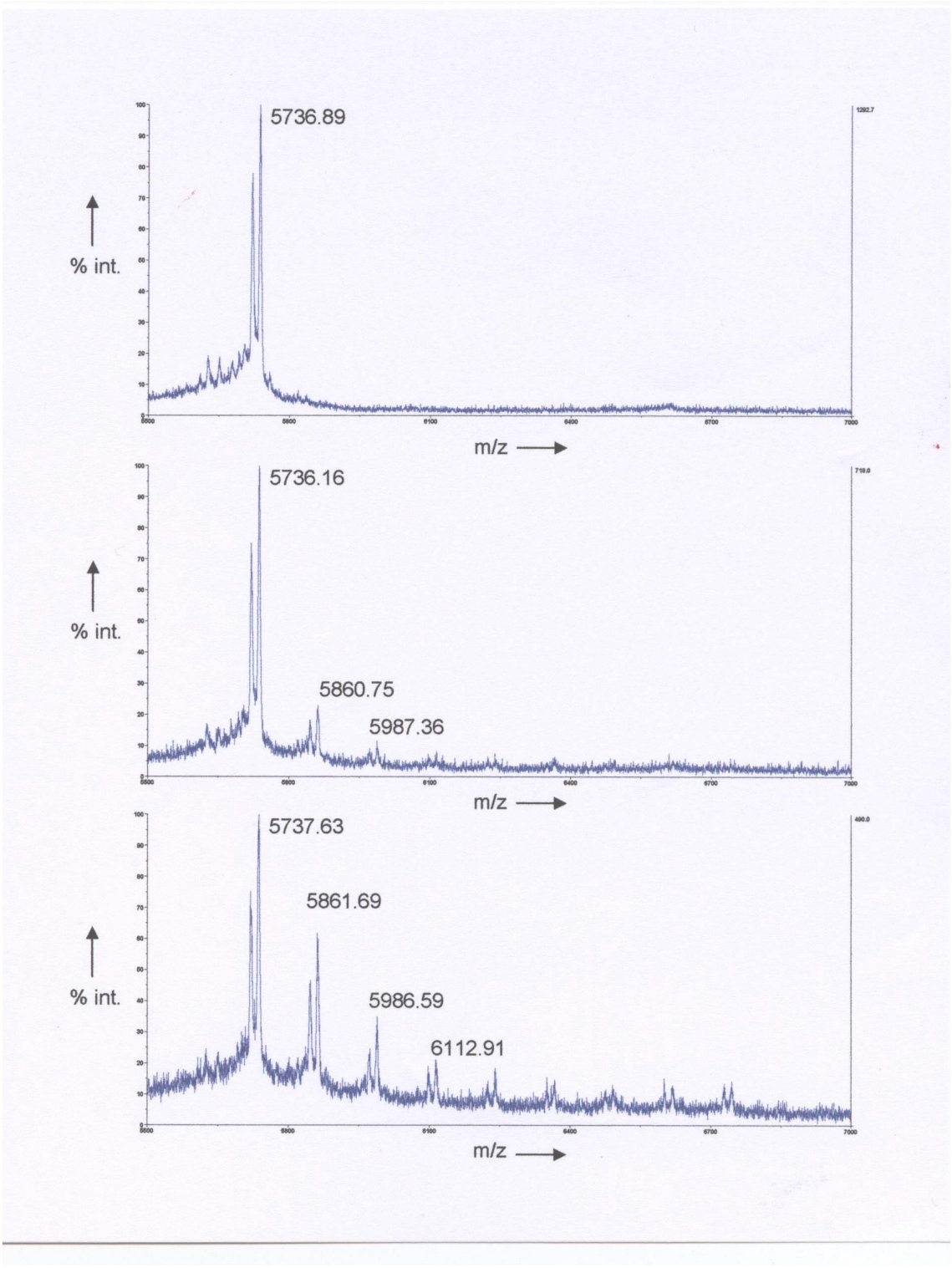
A) Bovine insulin (retention time = 3.7 min) and B) Octaiodinated bovine insulin with 8.8 equiv. of IPy_2BF_4 (retention time = 4.7 min), using a 20 to 52% gradient of acetonitrile (+0.036% trifluoroacetic acid) into water (0.045% trifluoroacetic acid) at 40°C over 8 min.

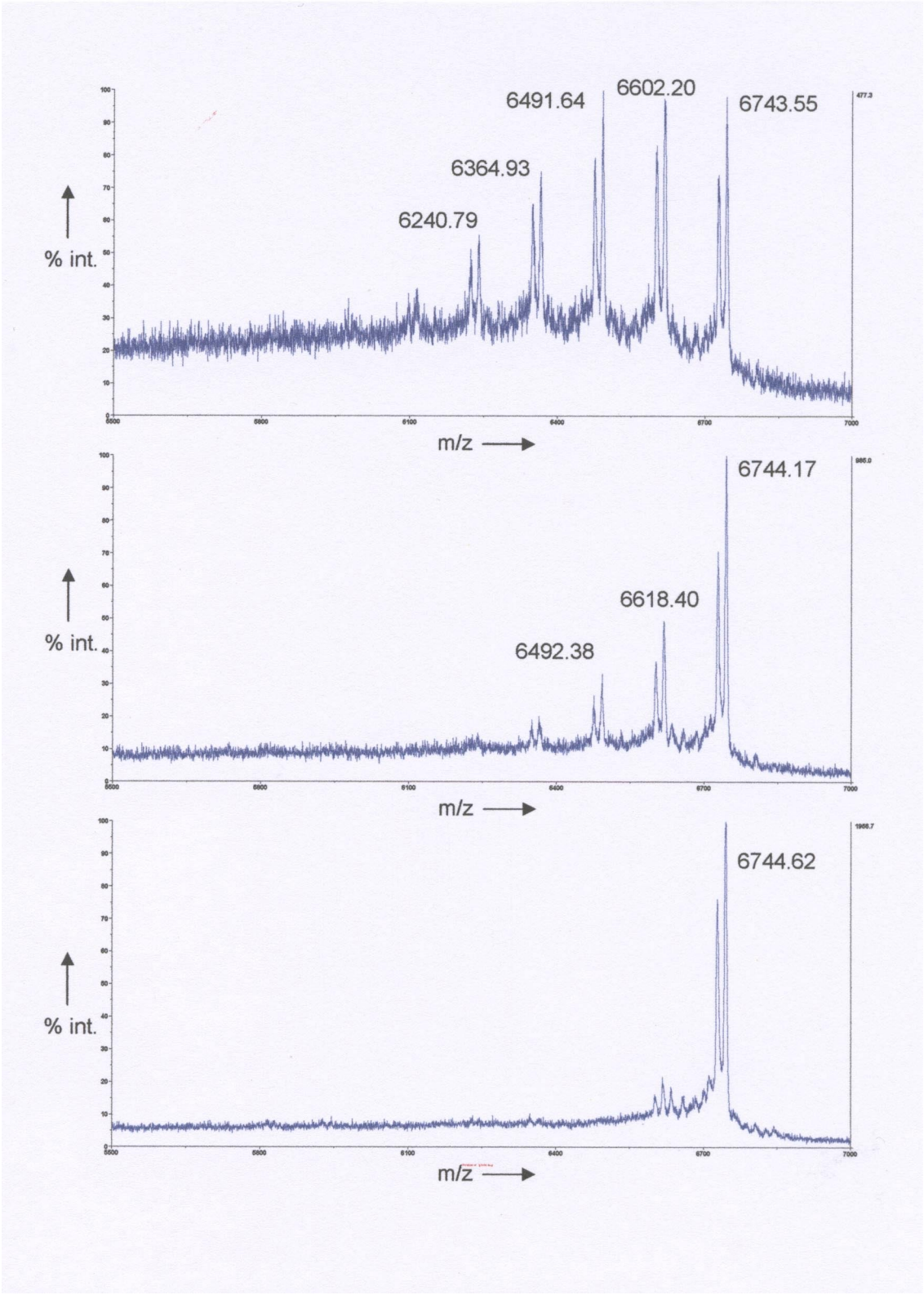
C)



C) Iodinated A and B insulin chains with retention times of 20.2 and 21.2 min, respectively, using a 10 to 40% gradient of acetonitrile (+0.036% trifluoroacetic acid) into water (0.045% trifluoroacetic acid) at 40°C over 30 min. These peaks were isolated and Edman sequenced, and their masses were confirmed by MALDI-TOF mass spectrometry.

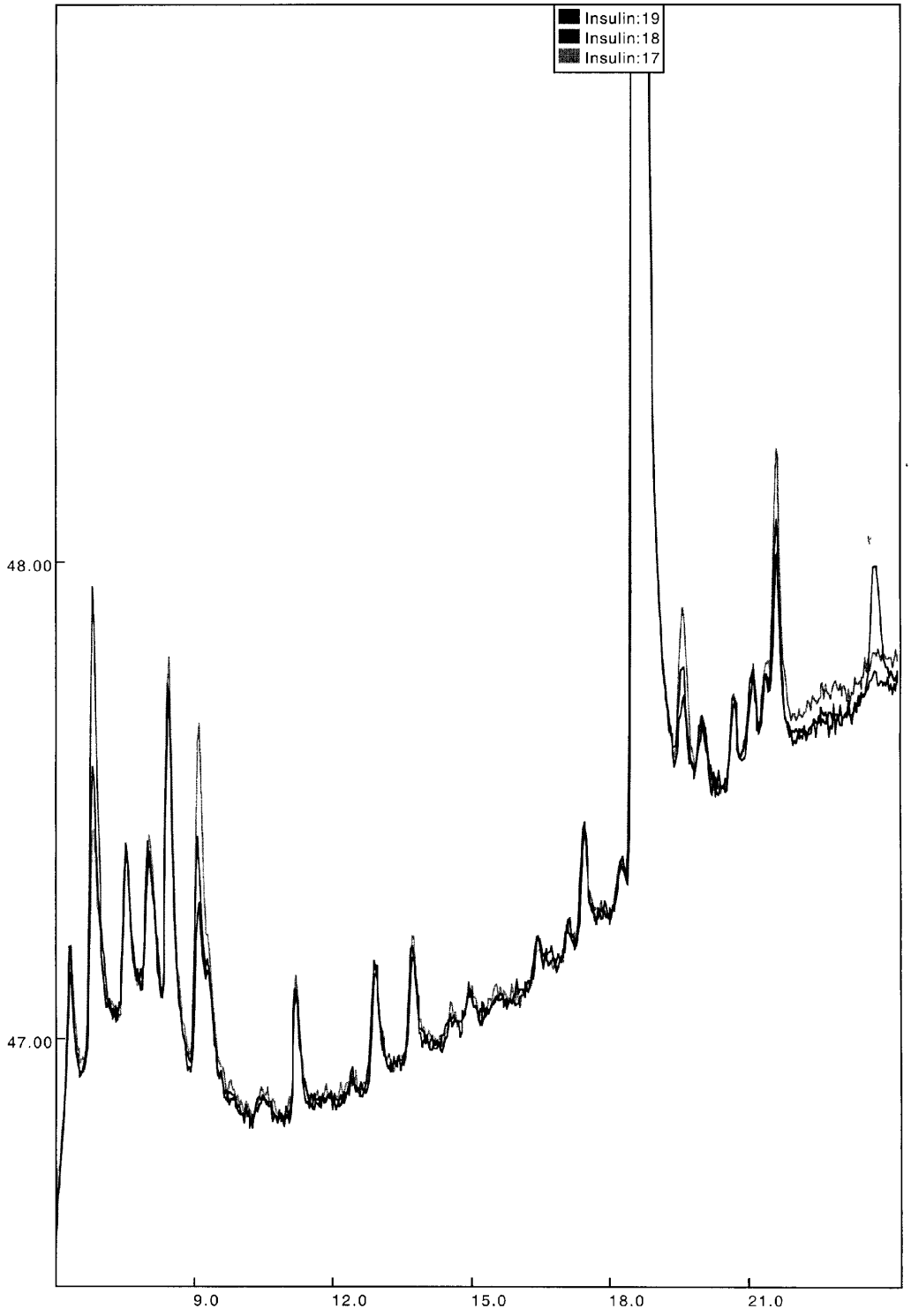
MALDI-TOF mass spectra of bovine insulin (top, m/z = 5,736.89), treated with 1.1, 2.2, 3.3, 4.4, 5.5, 6.6, 7.7 and 8.8 equivalents (bottom, m/z = 6,744.62) of IPy2BF4.




6491.64
6602.20
6743.55
6364.93
6240.79
% int.
m/z
6744.17
6618.40
6492.38
% int.
m/z
6744.62
% int.
m/z

Edman sequencing for iodinated Insulin A chain in a Procise sequencer (Applied Biosystems).

The results indicate diiodination of both Tyr14 and Tyr19.

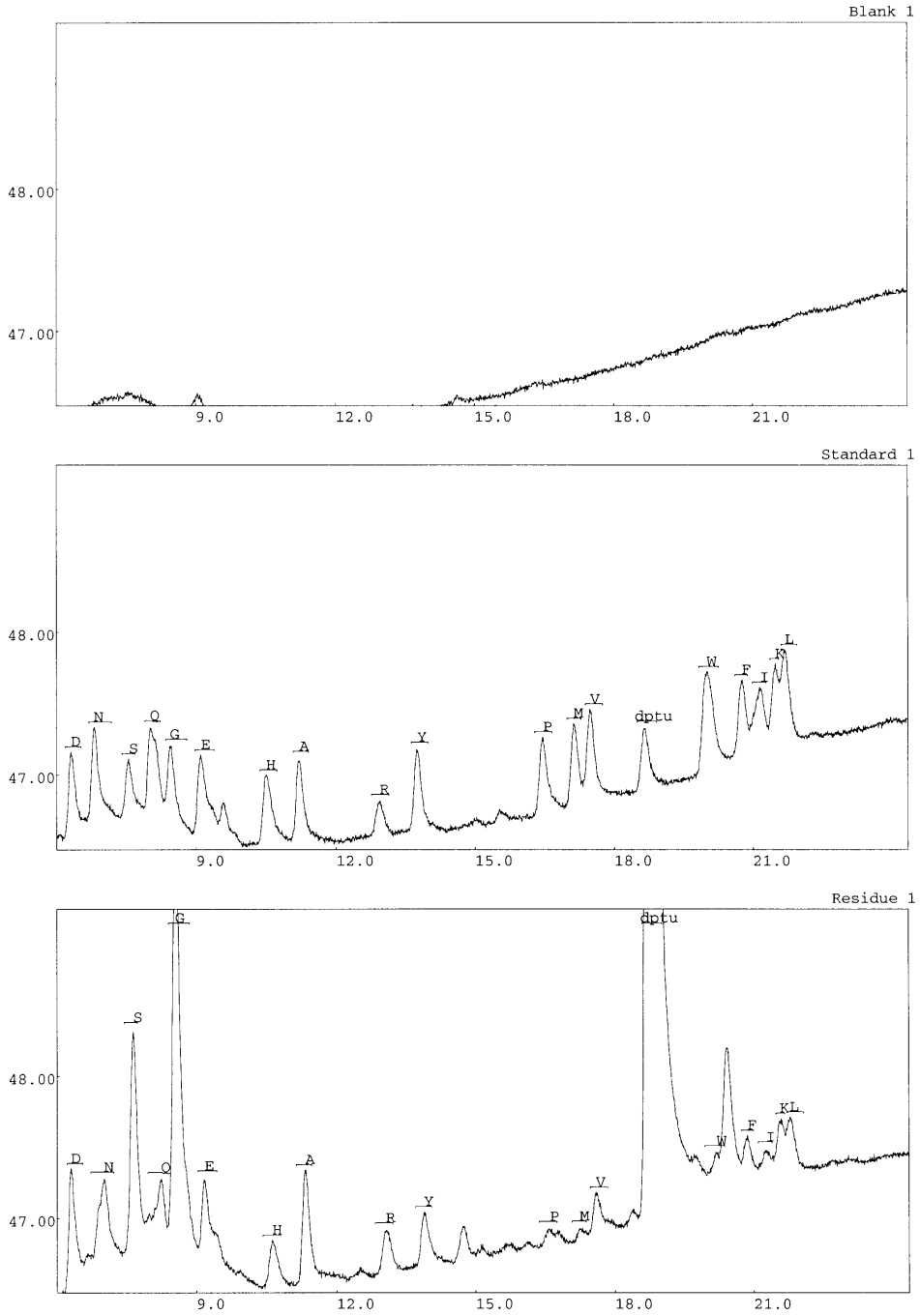

Insulin:19
Insulin:18
Insulin:17
48.00
47.00
9.0
12.0
15.0
18.0
21.0

File Name: Insulina1
In Folder: Procise HDD:Procise*f*:
21 cycles acquired on an Applied Biosystems Procise Sequencer
Sequencer Name: PROCISE
Run started on Wednesday, January 13, 1904 at 4:07 PM

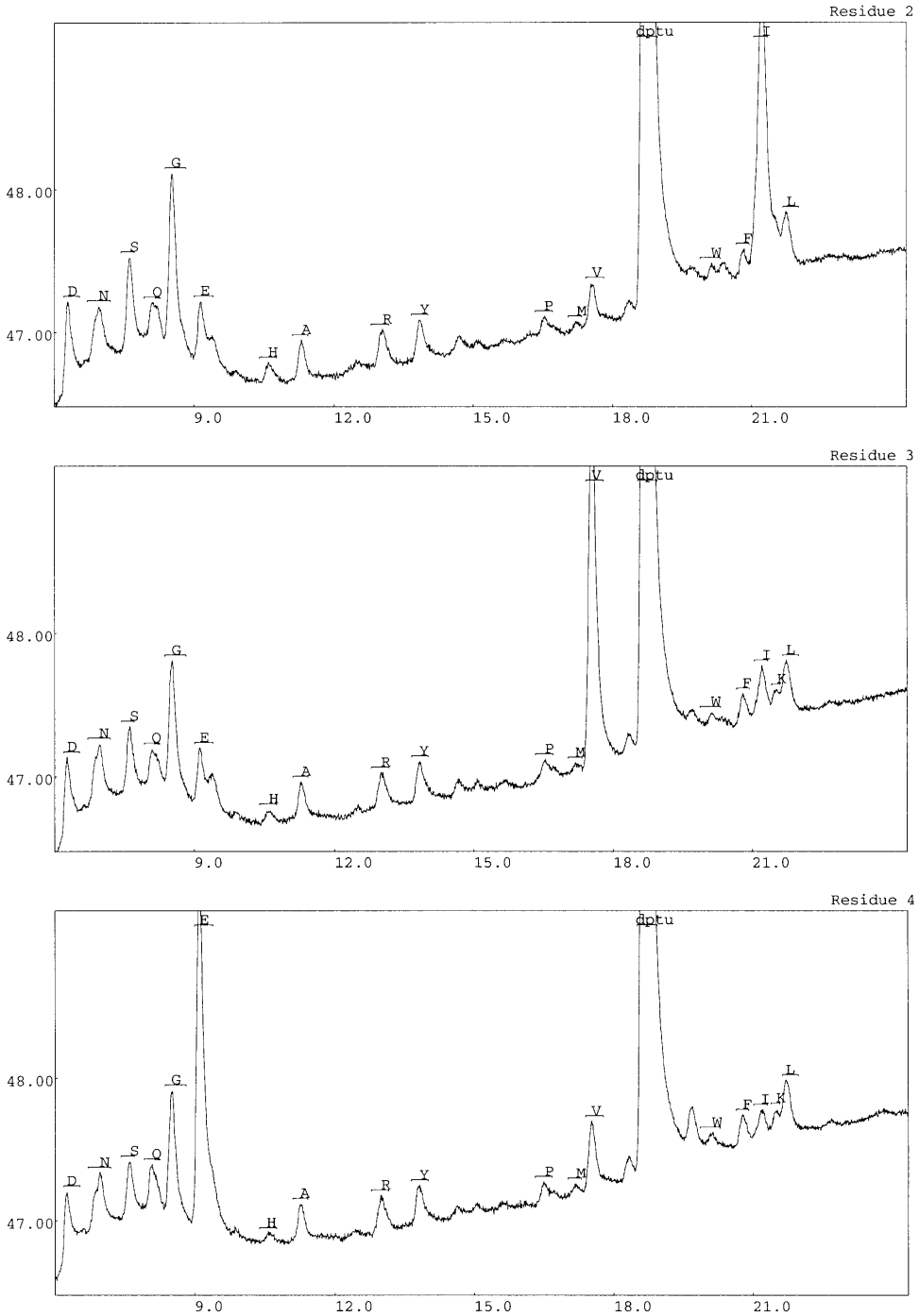
Sample Name:	Insulina1		
Sample Amount:	0.0 picomoles	Sampling Rate:	4.0 Hertz
Standard Amount:	1.0 picomoles	Detector Scale:	1.000 AUFS

Analyzed Using Smooth Degree 9, Interpolated Baseline

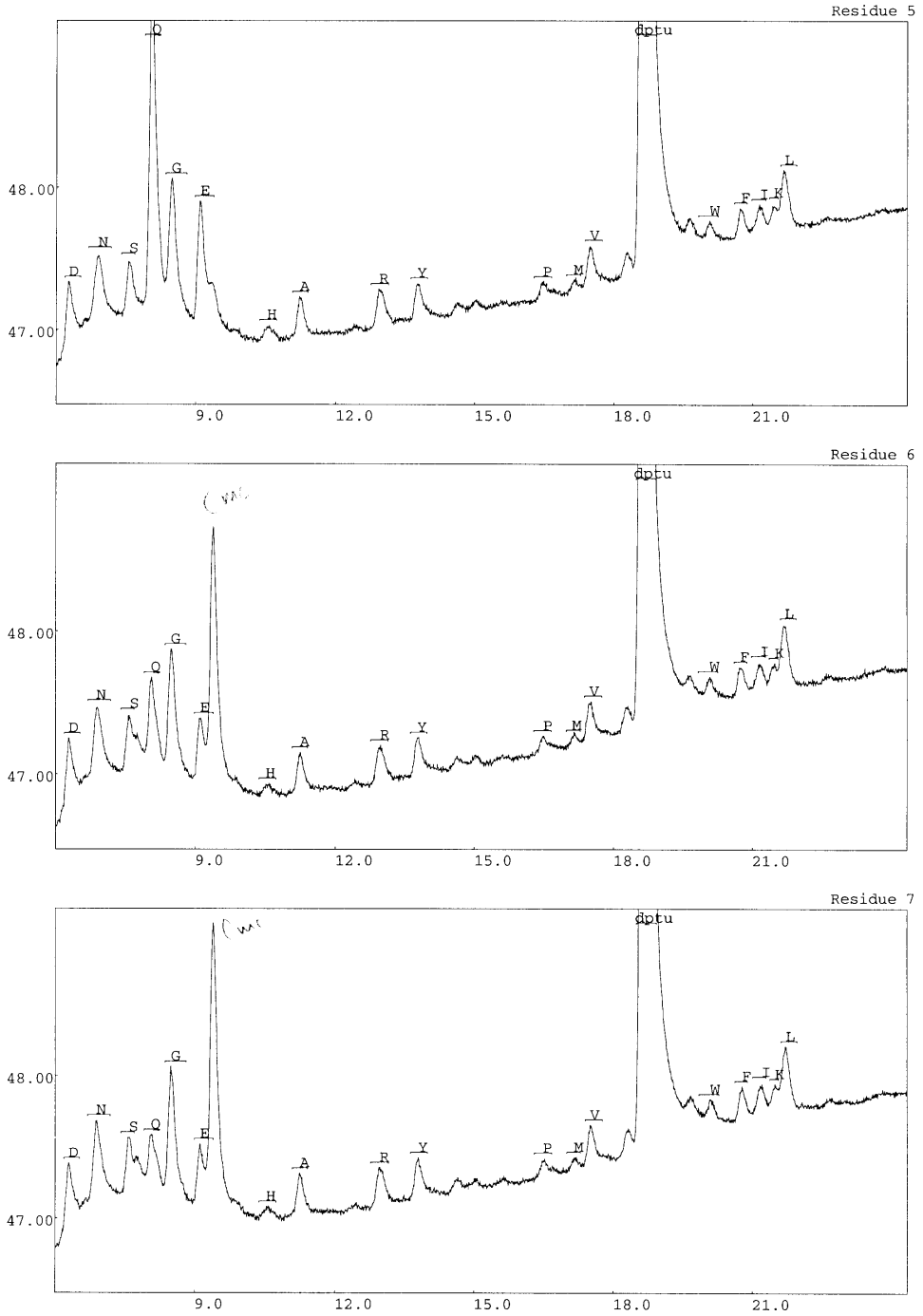
Analysis Limits:	1 - 128
Integration Limits:	0.00 - 24.00 min.
Baseline Limits:	0.10 - 1.00 min.
Peak Half Width:	0.05 min.
Peak Separation:	0.05 min.
Search Length:	0.10 min.
RT Factor:	0.75
Sensitivity:	1.00



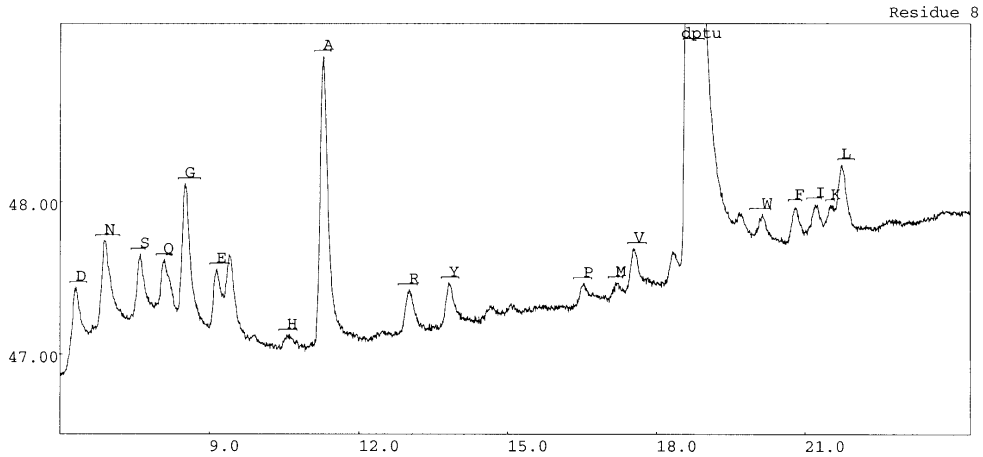
Applied Biosystems Procise - PROCISE

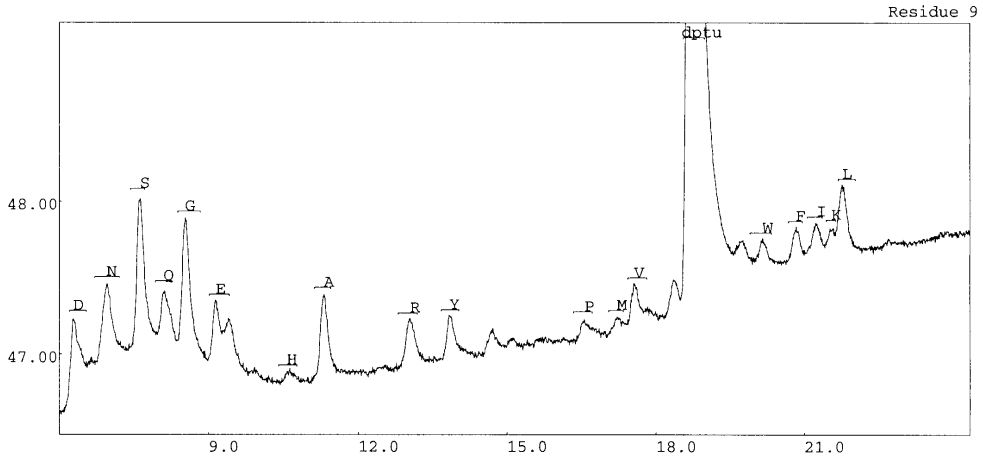


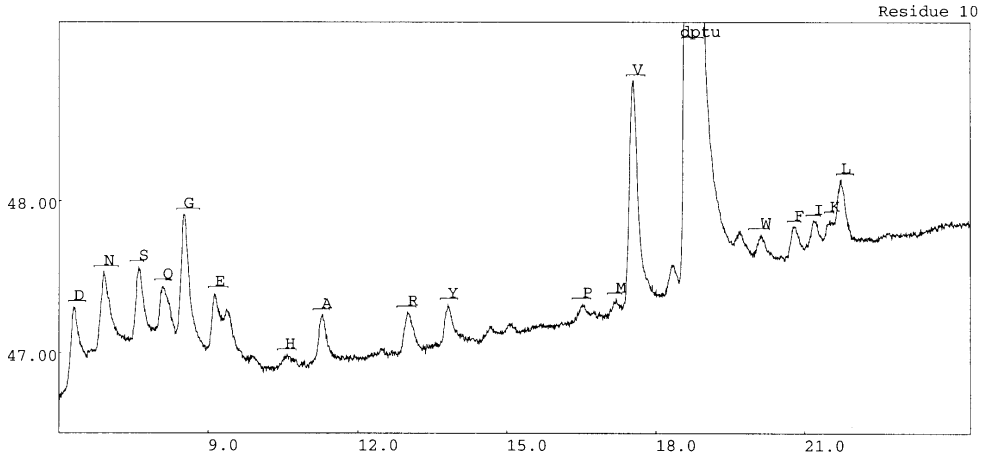
Applied Biosystems Procise - PROCISE



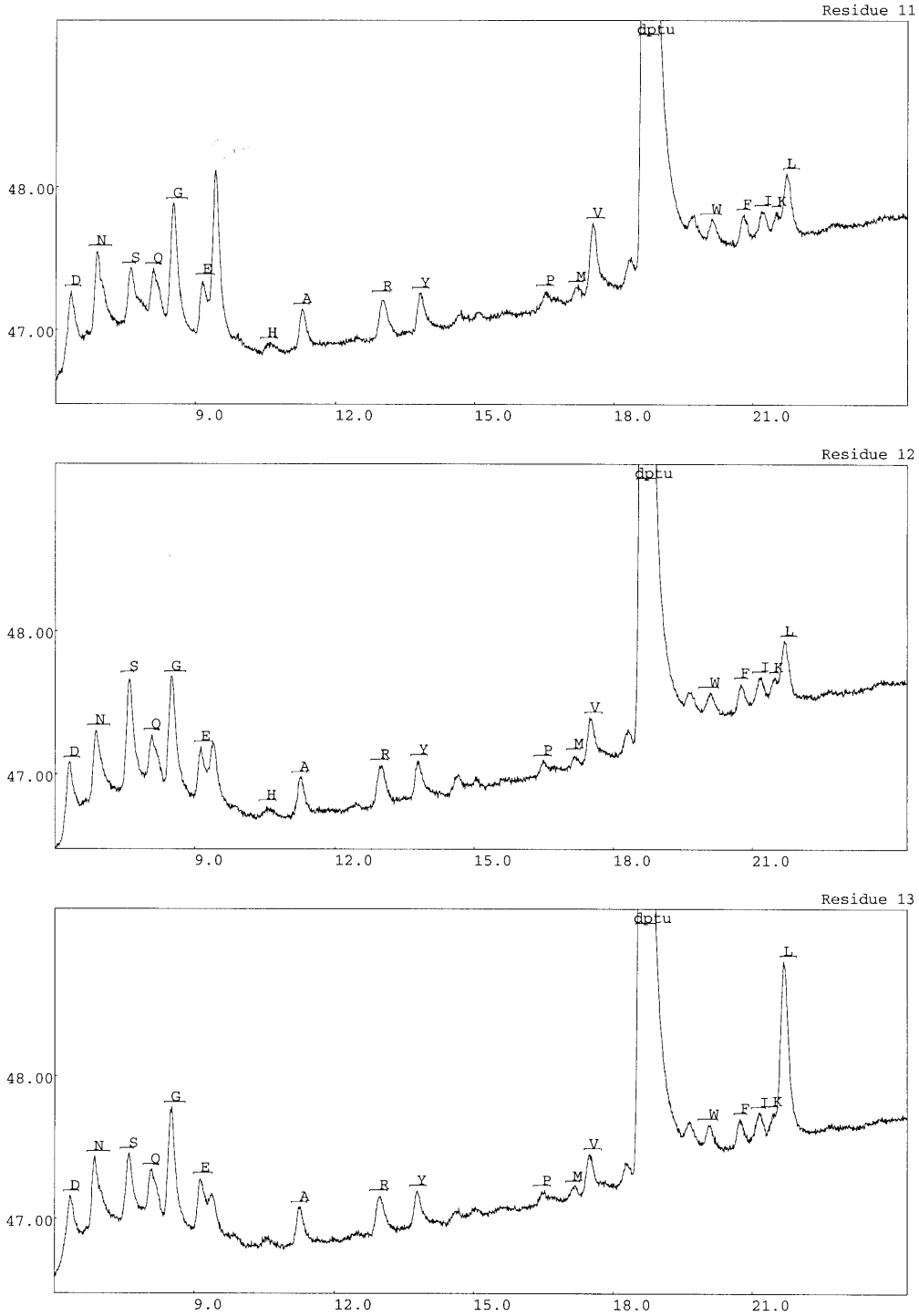
Applied Biosystems Procise - PROCISE



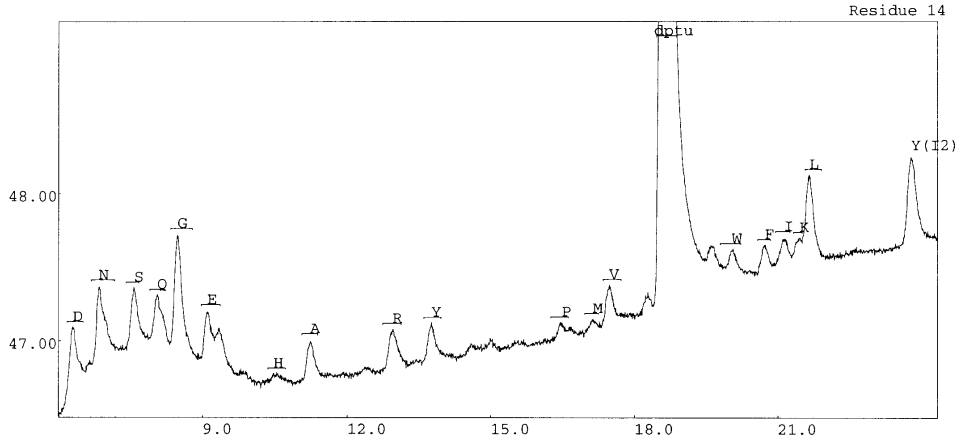


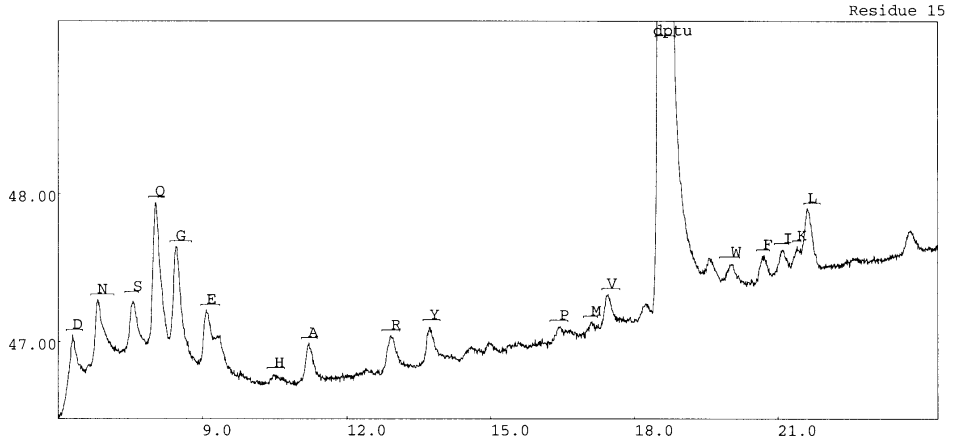


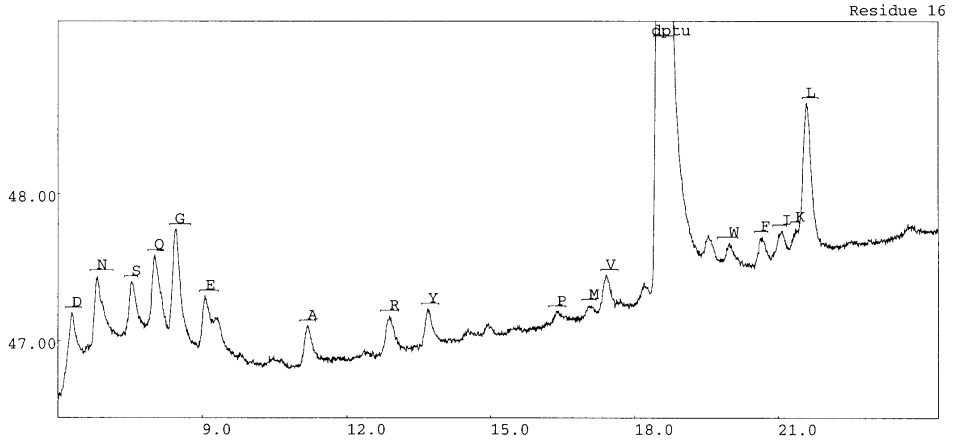
Applied Biosystems Procise - PROCISE



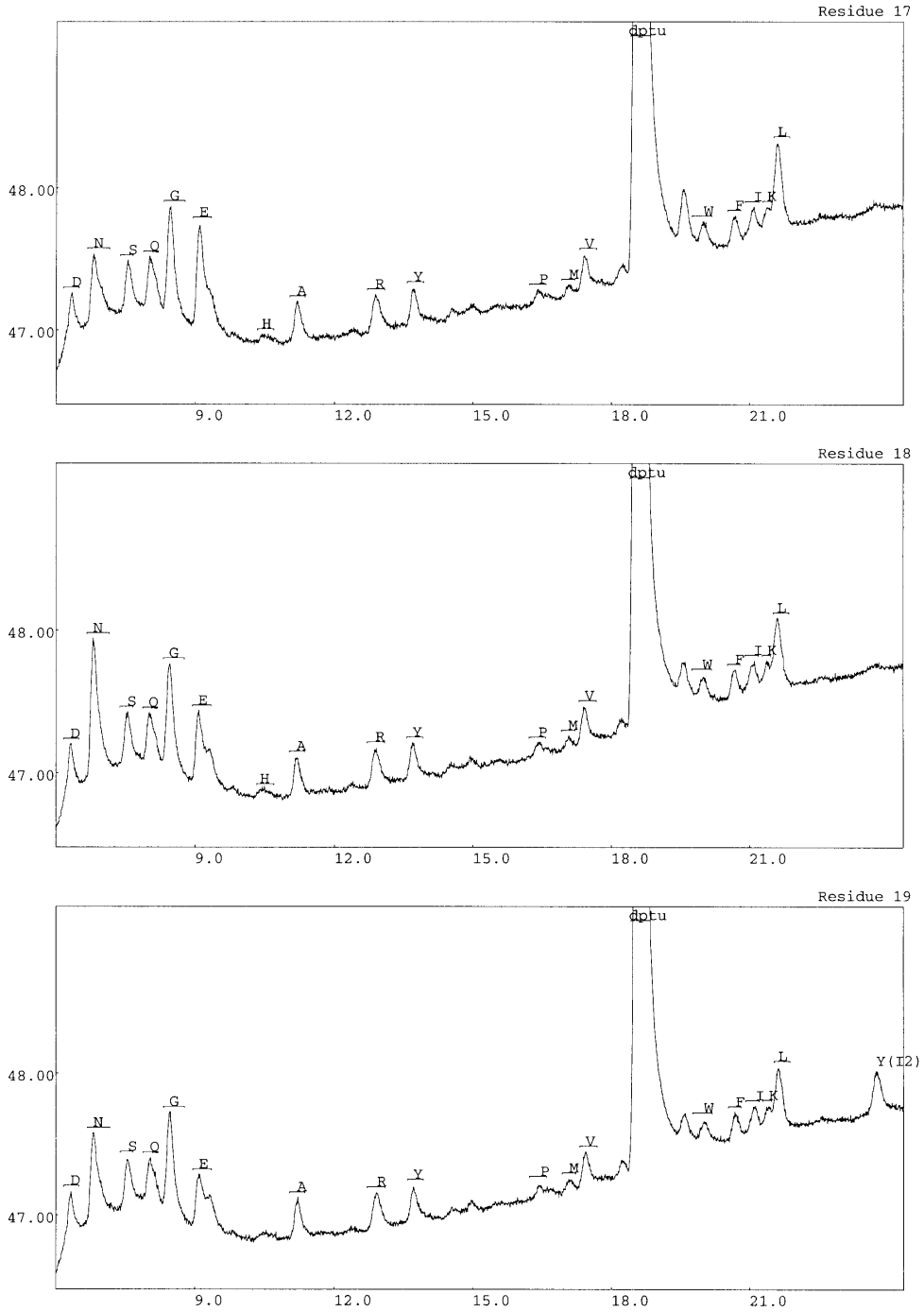
Applied Biosystems Procise - PROCISE







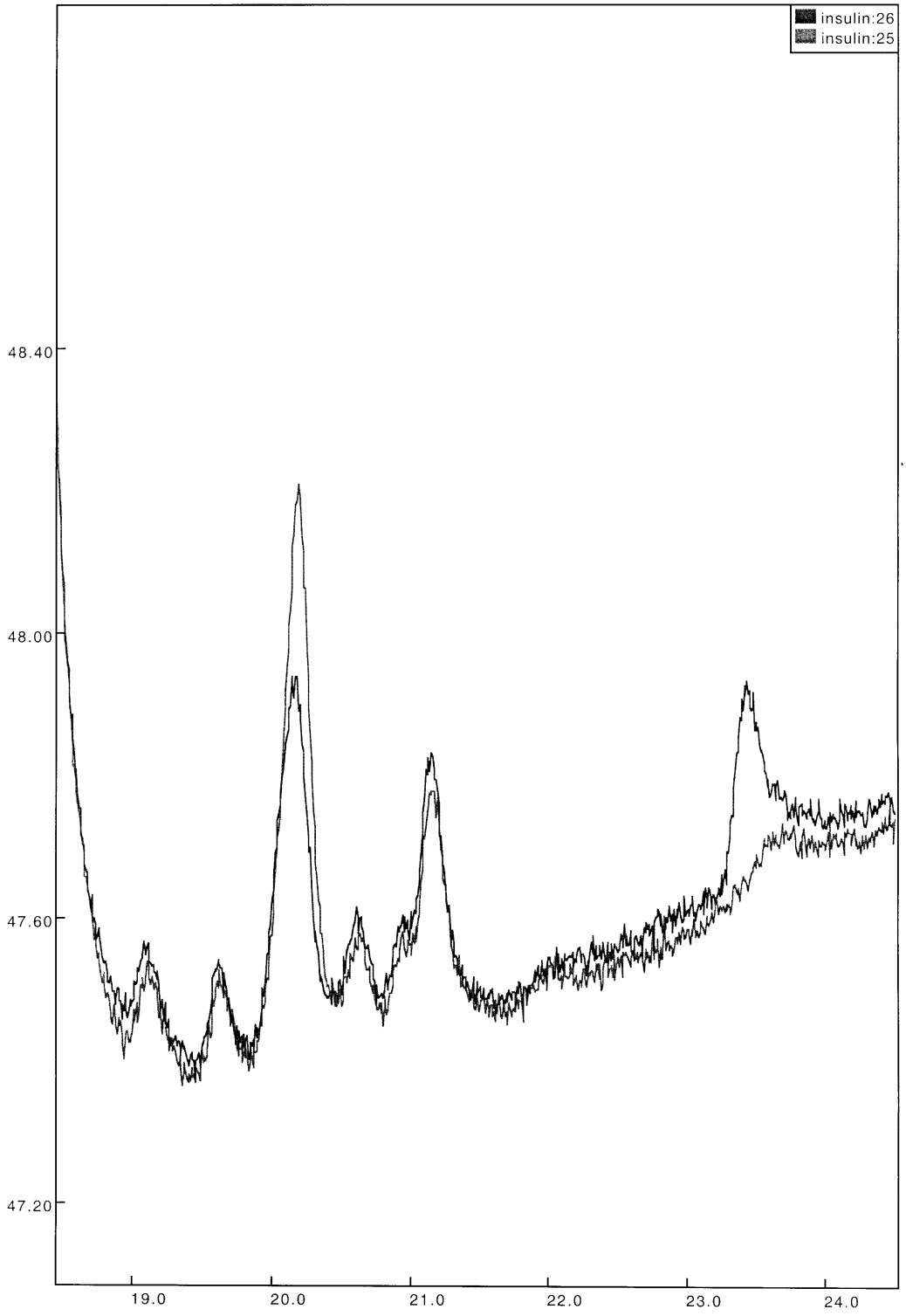
Applied Biosystems Procise - PROCISE

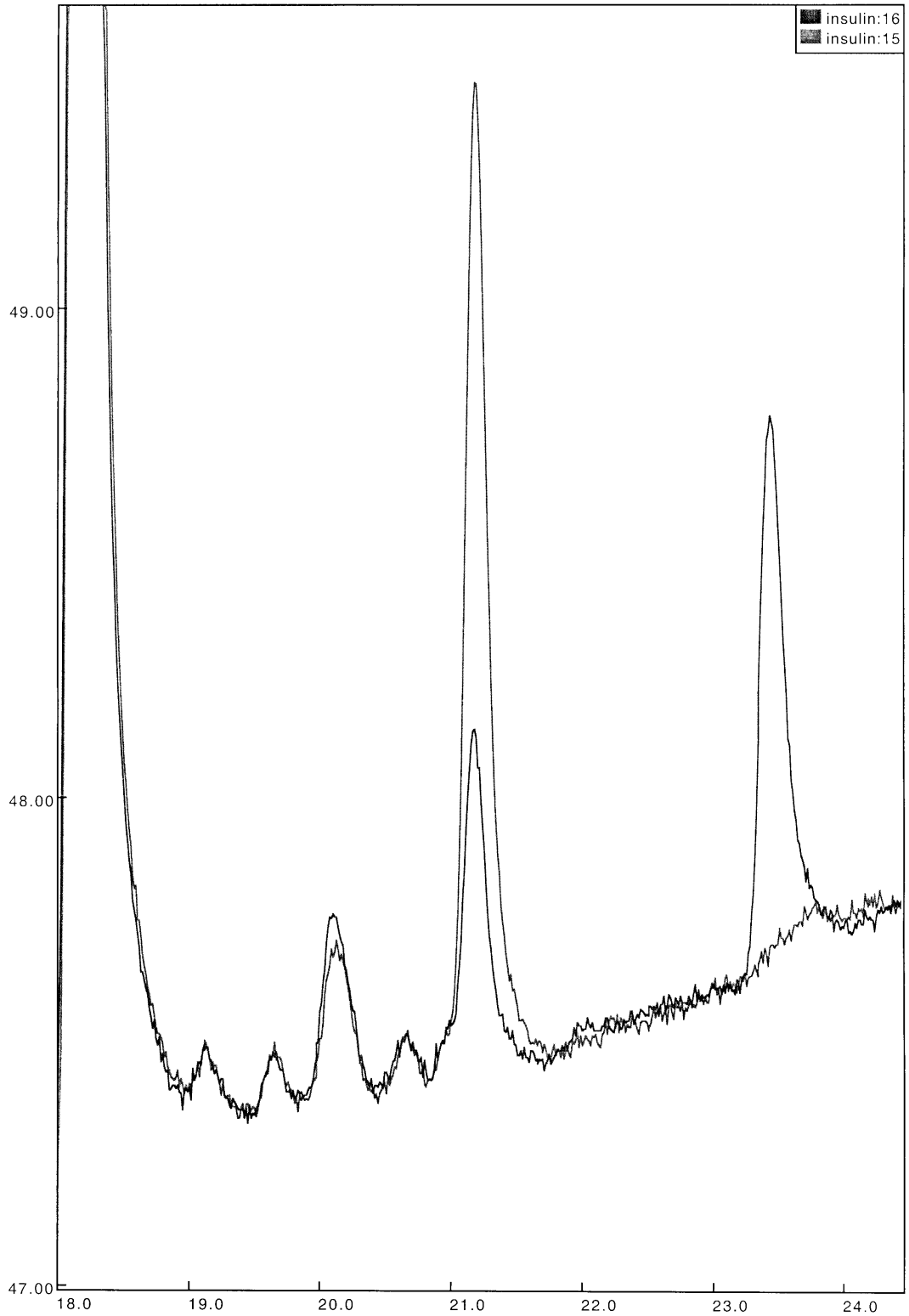


Applied Biosystems Procise - PROCISE

Edman sequencing for iodinated Insulin B chain in a Procise sequencer (Applied Biosystems).

The results indicate diiodination of both Tyr16 and Tyr26.


insulin:26
insulin:25
48.40
48.00
47.60
47.20
19.0
20.0
21.0
22.0
23.0
24.0

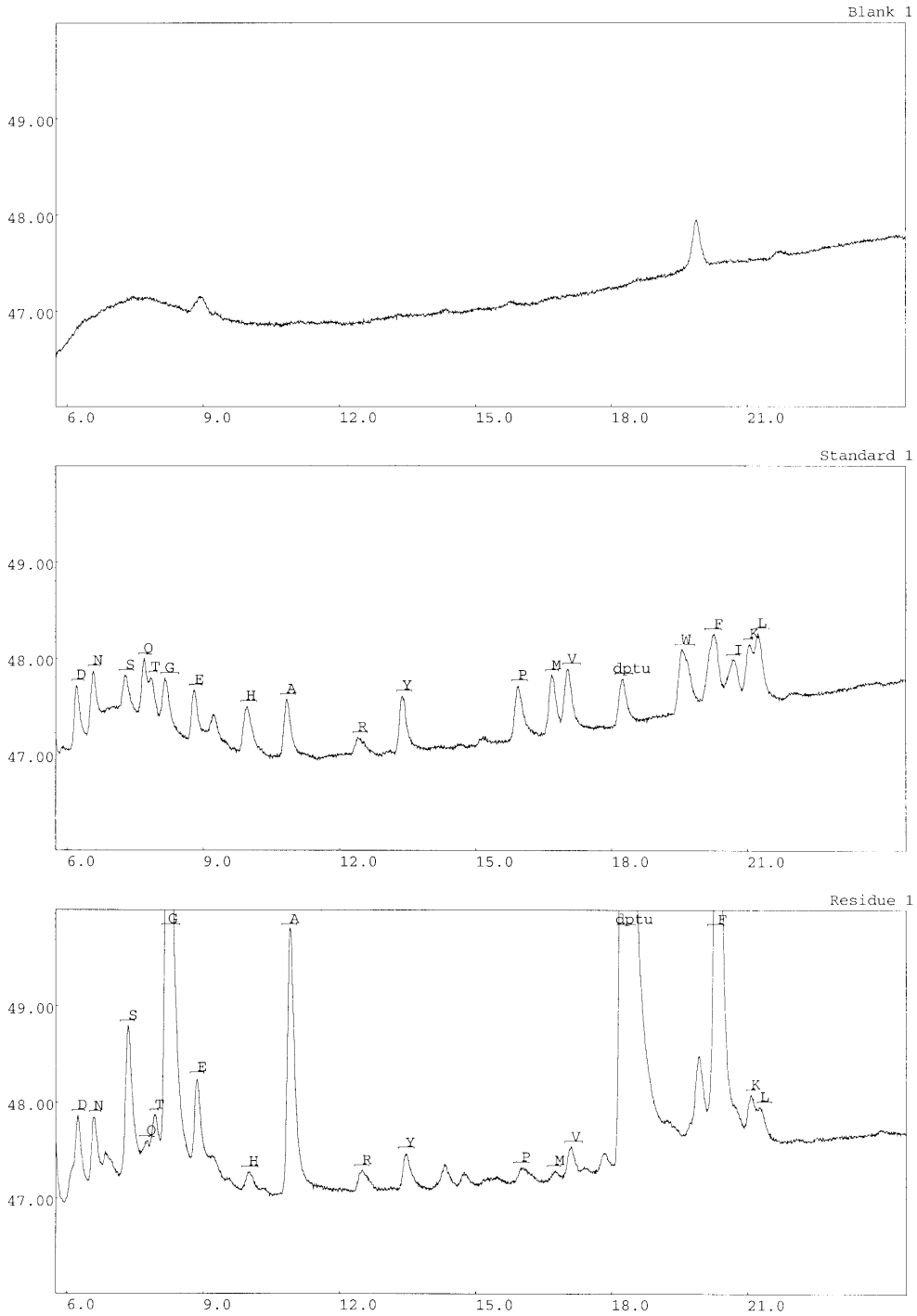

insulin:16
insulin:15
49.00
48.00
47.00
18.0
19.0
20.0
21.0
22.0
23.0
24.0

File Name: insuline2
In Folder: Procise HDD:Procise*f*:
28 cycles acquired on an Applied Biosystems Procise Sequencer
Sequencer Name: PROCISE
Run started on Friday, January 15, 1904 at 8:13 PM

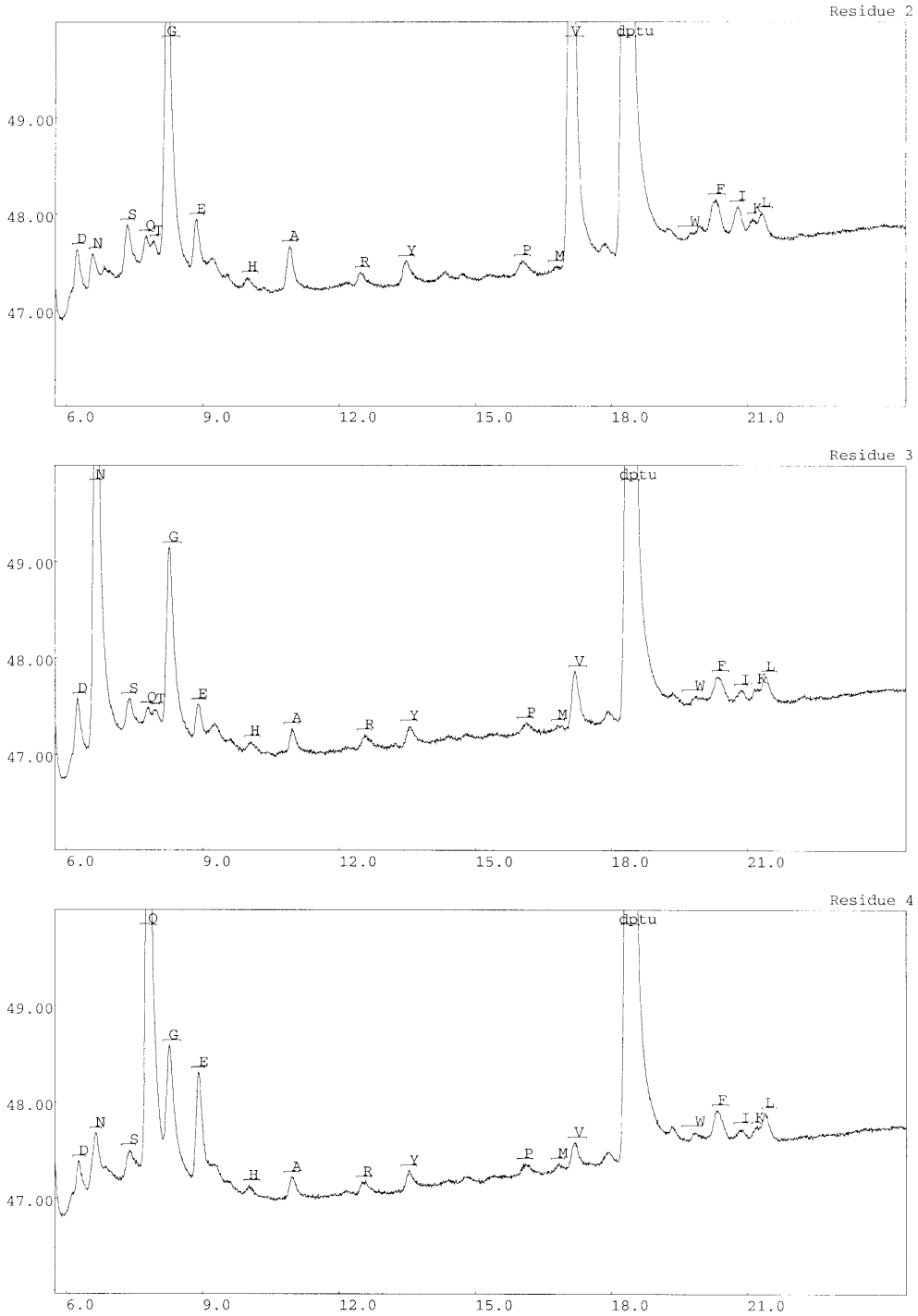
Sample Name:	insuline2		
Sample Amount:	0.0 picomoles	Sampling Rate:	4.0 Hertz
Standard Amount:	1.0 picomoles	Detector Scale:	1.000 AUFS

Analyzed Using Smooth Degree 9, Interpolated Baseline

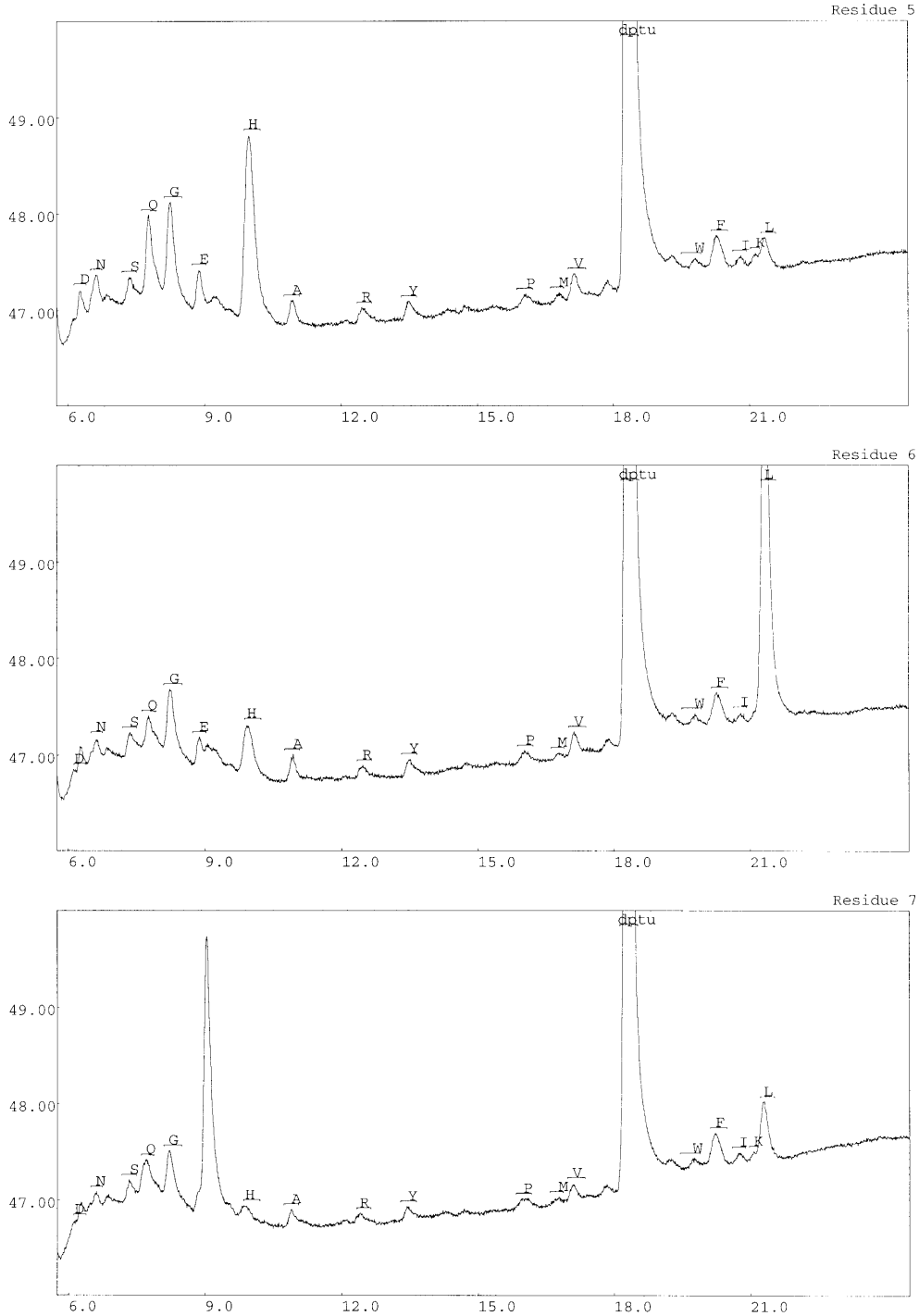
Analysis Limits:	1 - 128
Integration Limits:	0.00 - 24.00 min.
Baseline Limits:	0.10 - 1.00 min.
Peak Half Width:	0.05 min.
Peak Separation:	0.05 min.
Search Length:	0.10 min.
RT Factor:	0.75
Sensitivity:	1.00



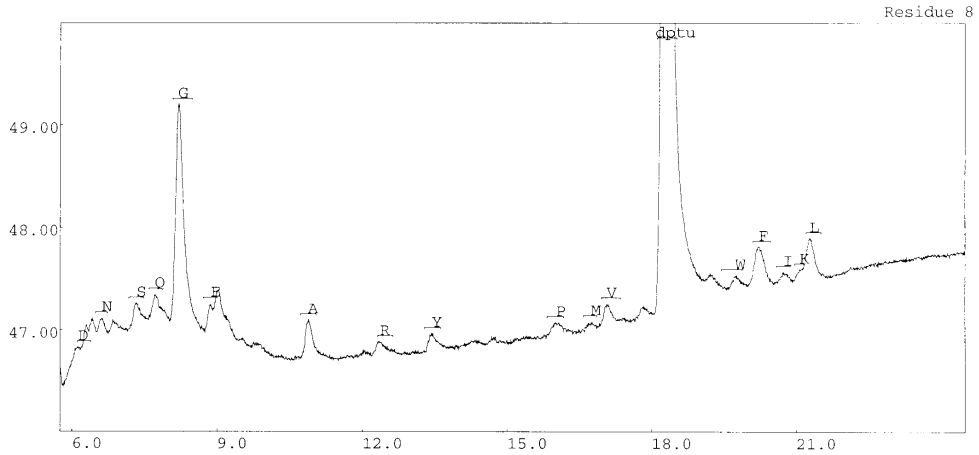
Applied Biosystems Procise - PROCISE

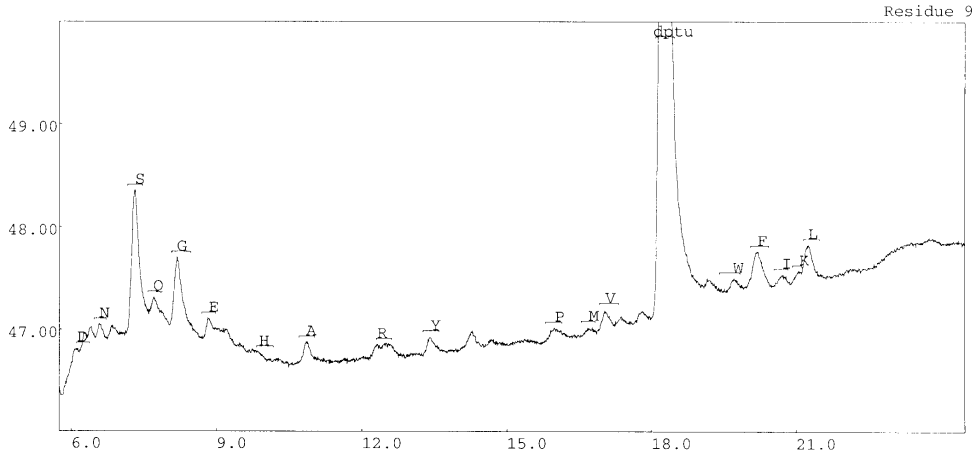


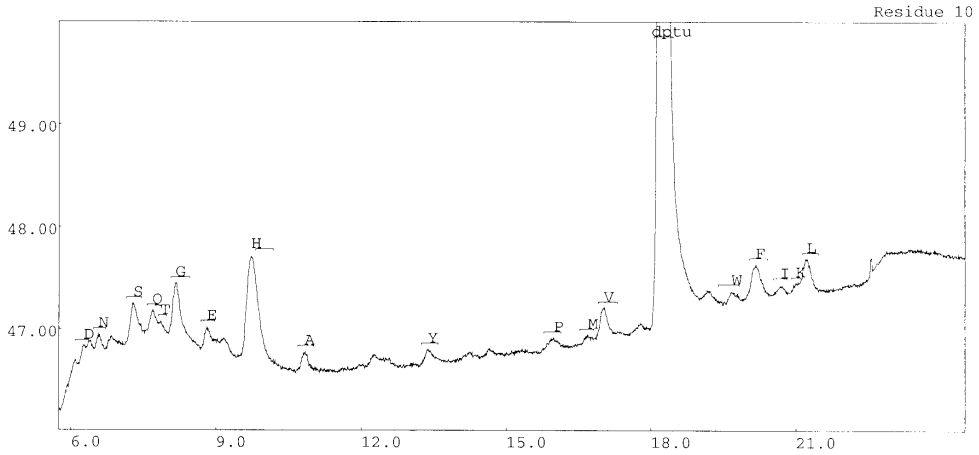
Applied Biosystems Procise - PROCISE



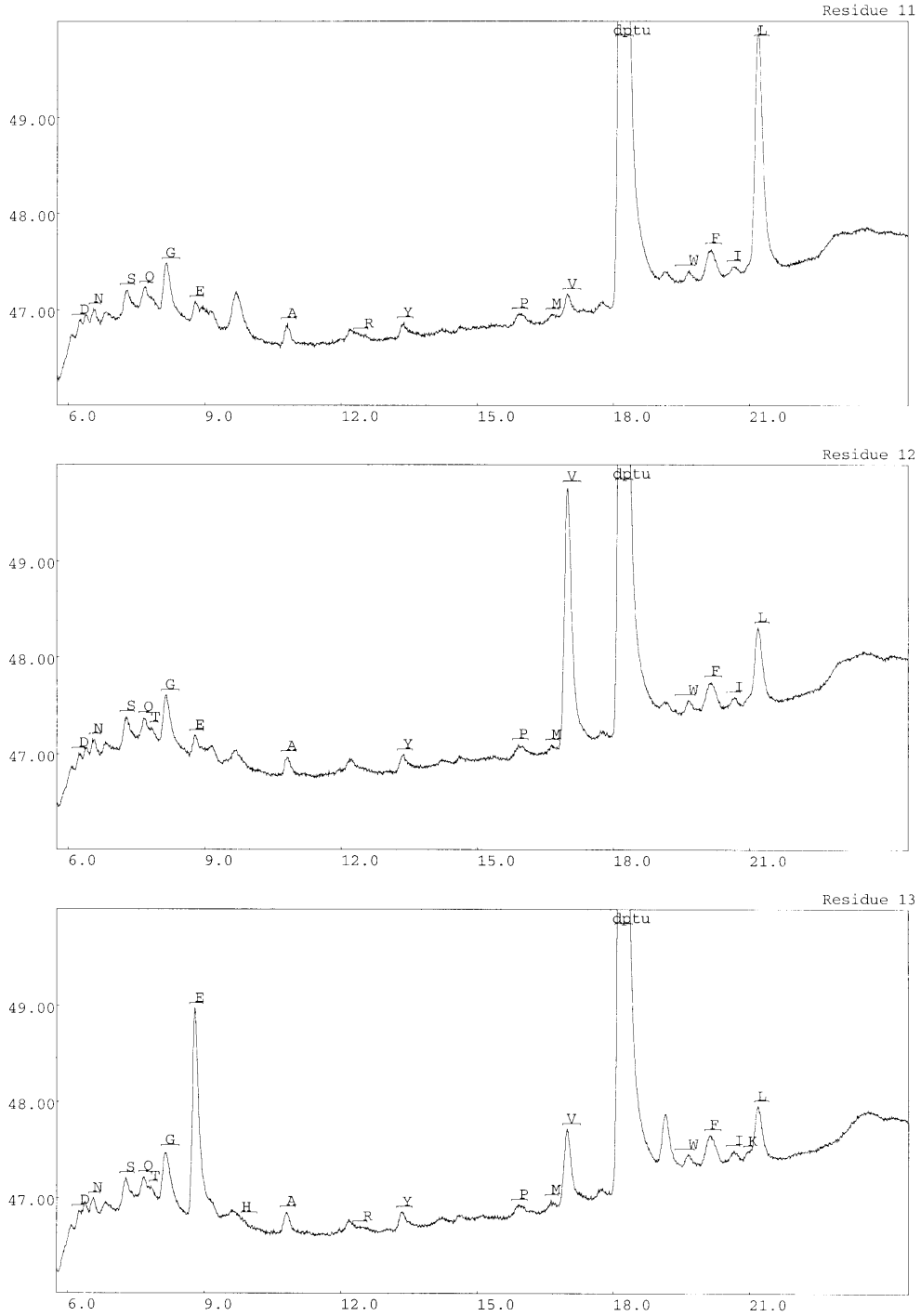
Applied Biosystems Procise - PROCISE



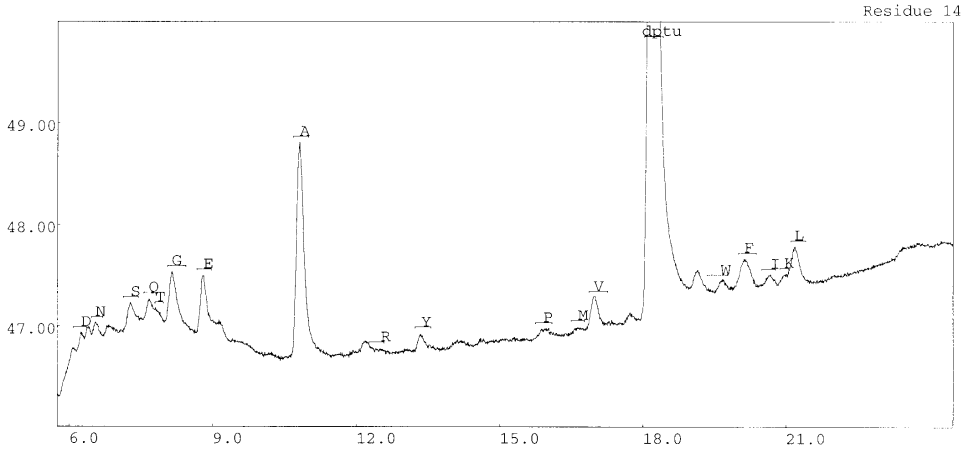


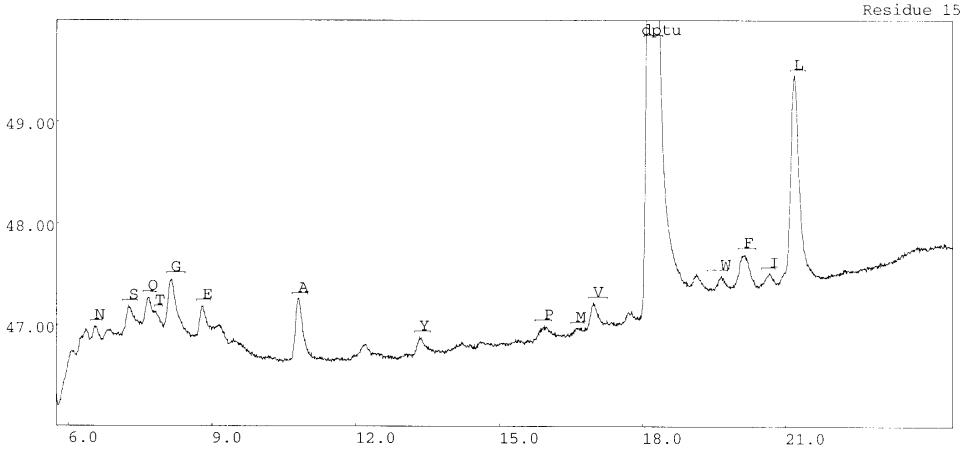


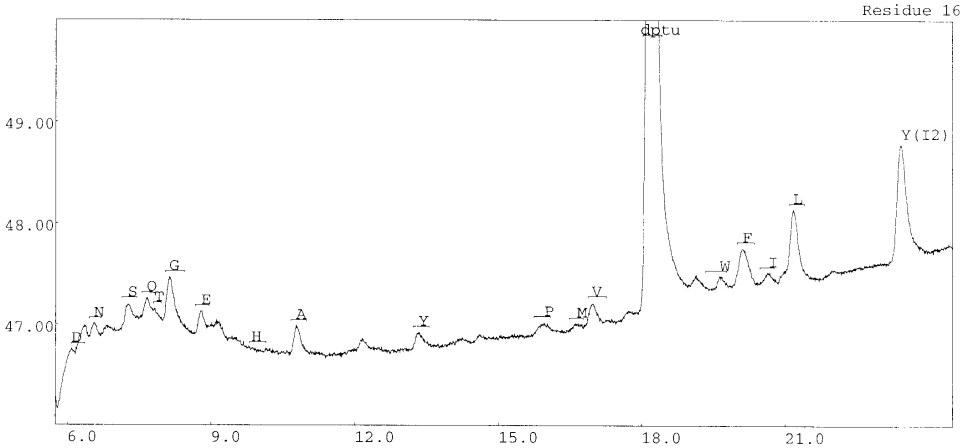
Applied Biosystems Procise - PROCISE



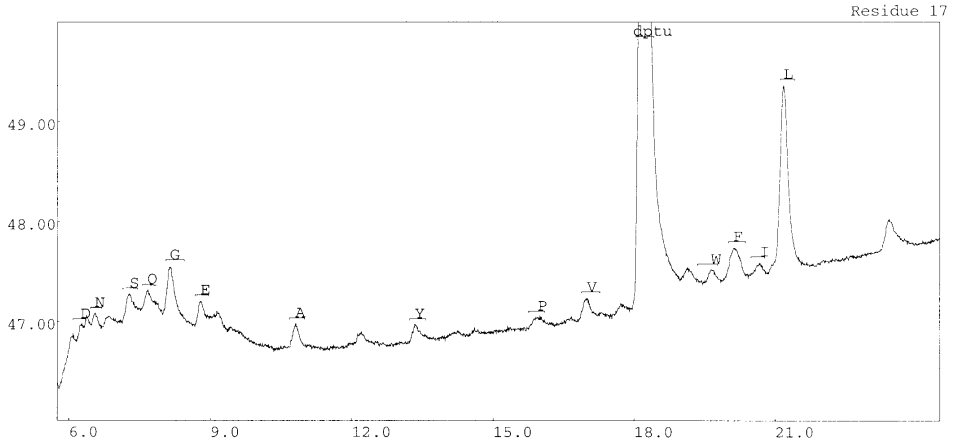
Applied Biosystems Procise - PROCISE

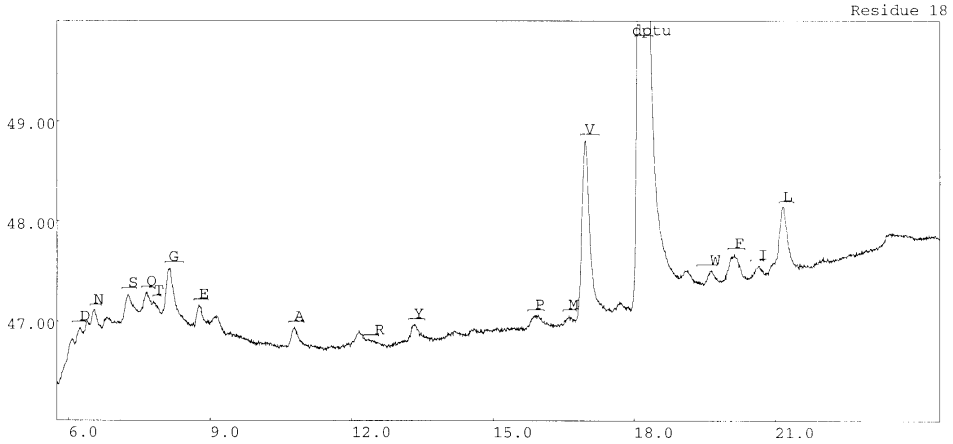


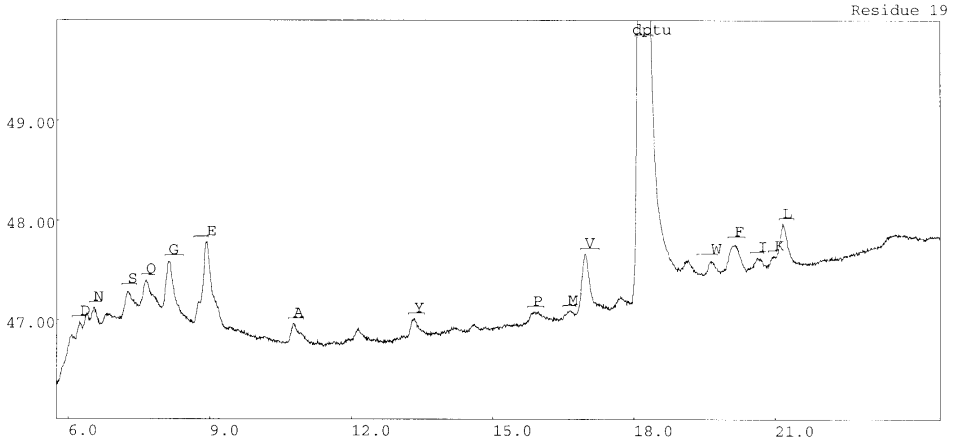




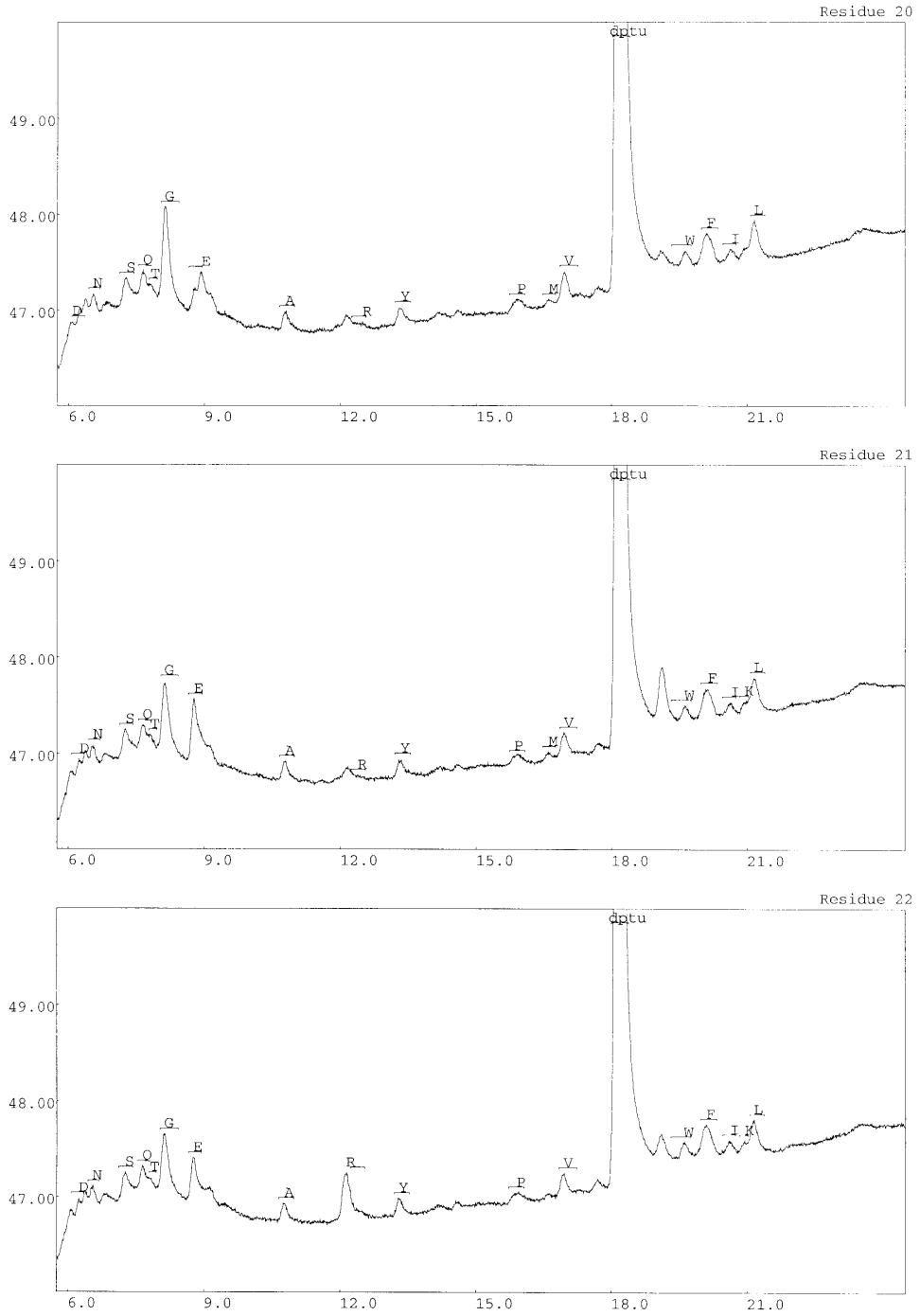
Applied Biosystems Procise - PROCISE



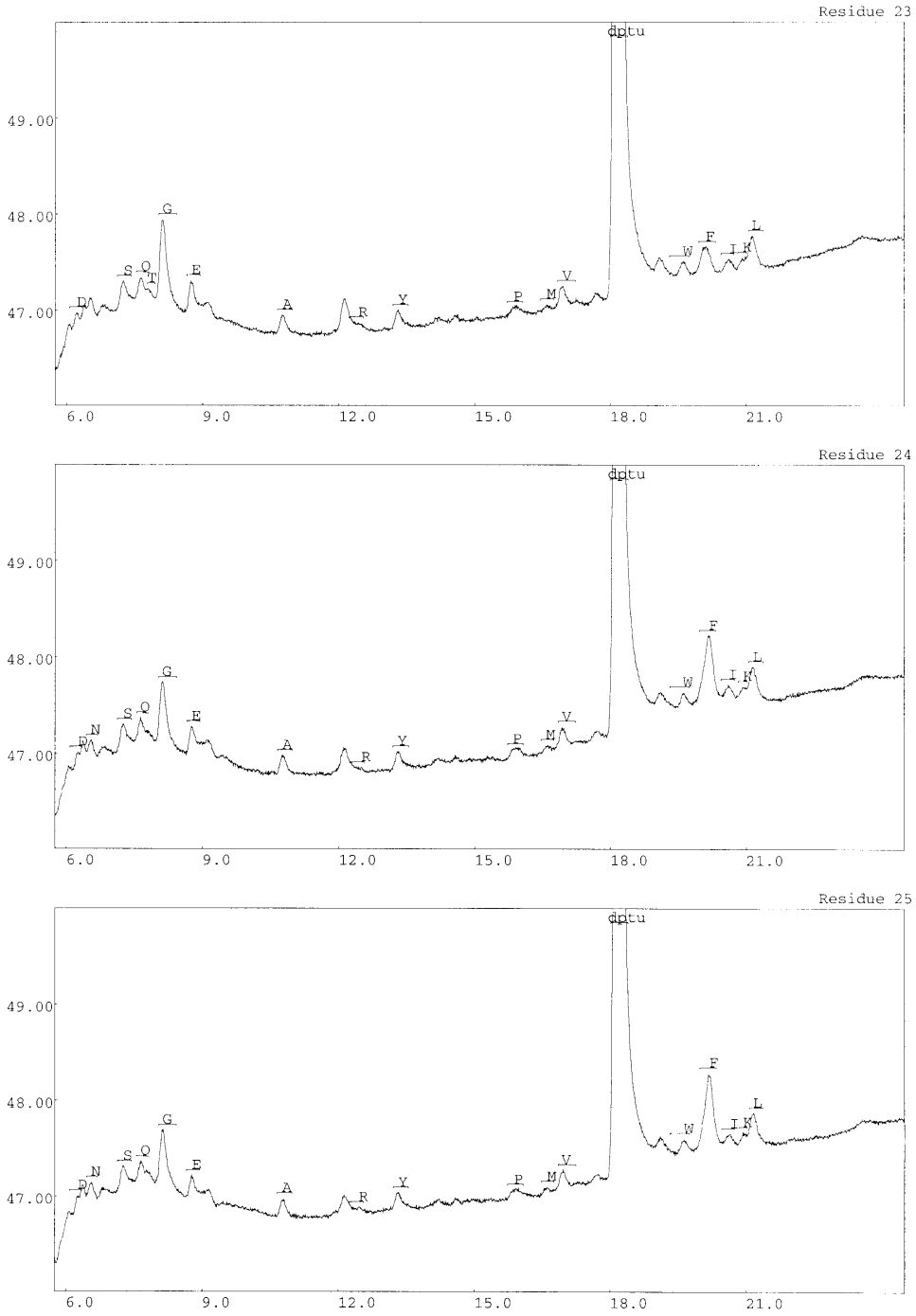




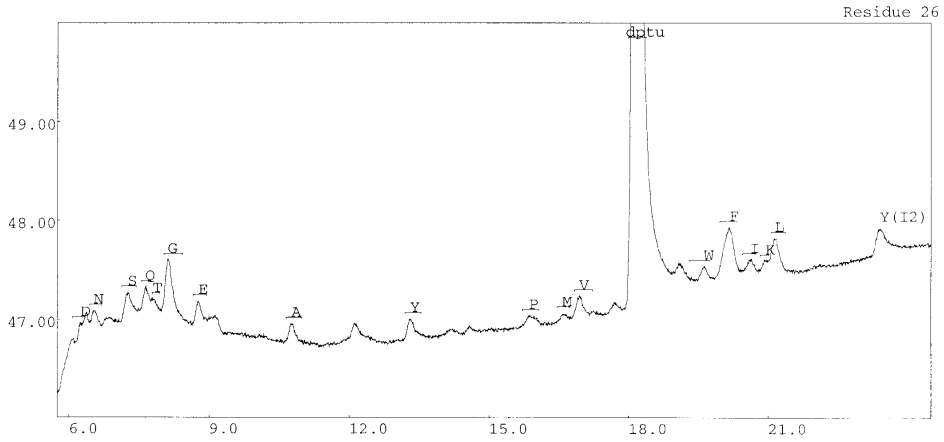
Applied Biosystems Procise - PROCISE



Applied Biosystems Procise - PROCISE



Applied Biosystems Procise - PROCISE



Applied Biosystems Procise - PROCISE

LYSOZYME: Supplementary material.

A) Tryptic digestion: calculated values

B) MS/MS sequence analysis

A) Tryptic digestion: calculated values

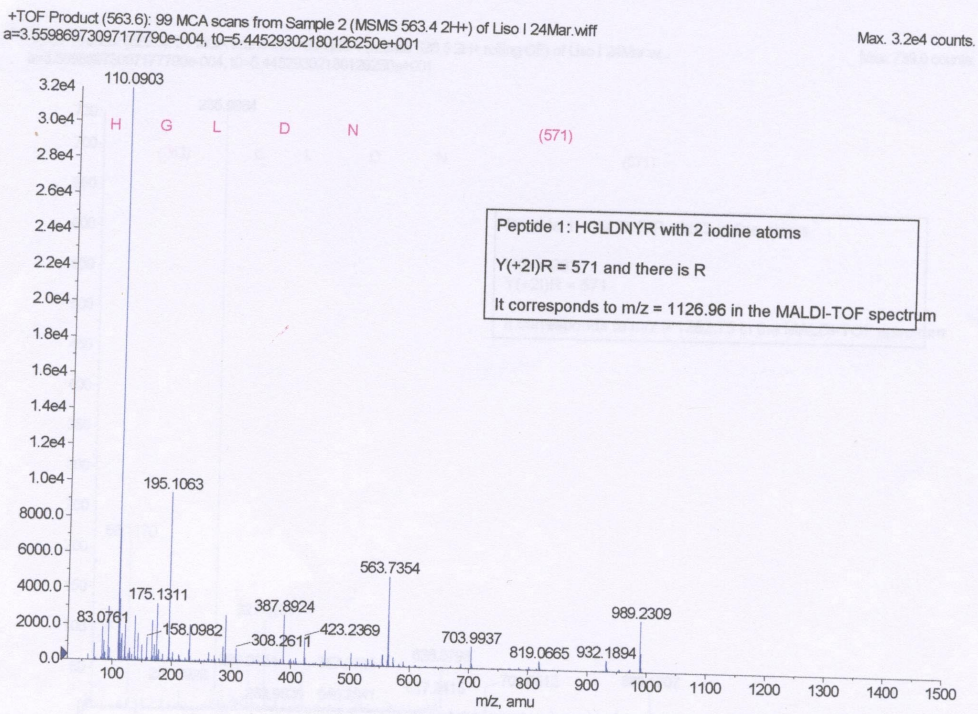
Peptide	$[M+H]^+$	$[M+I+H]^+$	$[M+2I+H]^+$	$[M+3I+H]^+$	#MC	position	artif. modifications		peptide sequence
	2337.12	-	-	-	0	74-96	Cys_CAM: 76, 80, 94	2508.19	NLCNIPCSALLSSDITASVNCAK
3	1753.83	1879.74	2005.64	-	0	46-61			NTDGSTDYGILQINSR
	1675.80	-	-	-	0	98-112	MSO: 105	1691.79	IVSDGNGMNAWVAWR
	1428.65	-	-	-	0	34-45			FESNFNTQATNR
2	1268.61	1451.53 [a]	1577.43 [a]	-	0	22-33	Cys_CAM: 30	1325.63	GYSLGNWVCAAK
	1045.54	-	-	-	0	117-125			GTDVQAWIR
	936.38	-	-	-	0	62-68	Cys_CAM: 64	993.39	WWCNDGR
1	874.42	1000.32	1126.22	1252.12	0	15-21			HGLDNYR
	836.40	-	-	-	0	6-13	Cys_CAM: 6 MSO: 12	893.42 852.39	CELAAAMK

[a] These masses include the carbamidomethylation of Cys, as they are found experimentally.

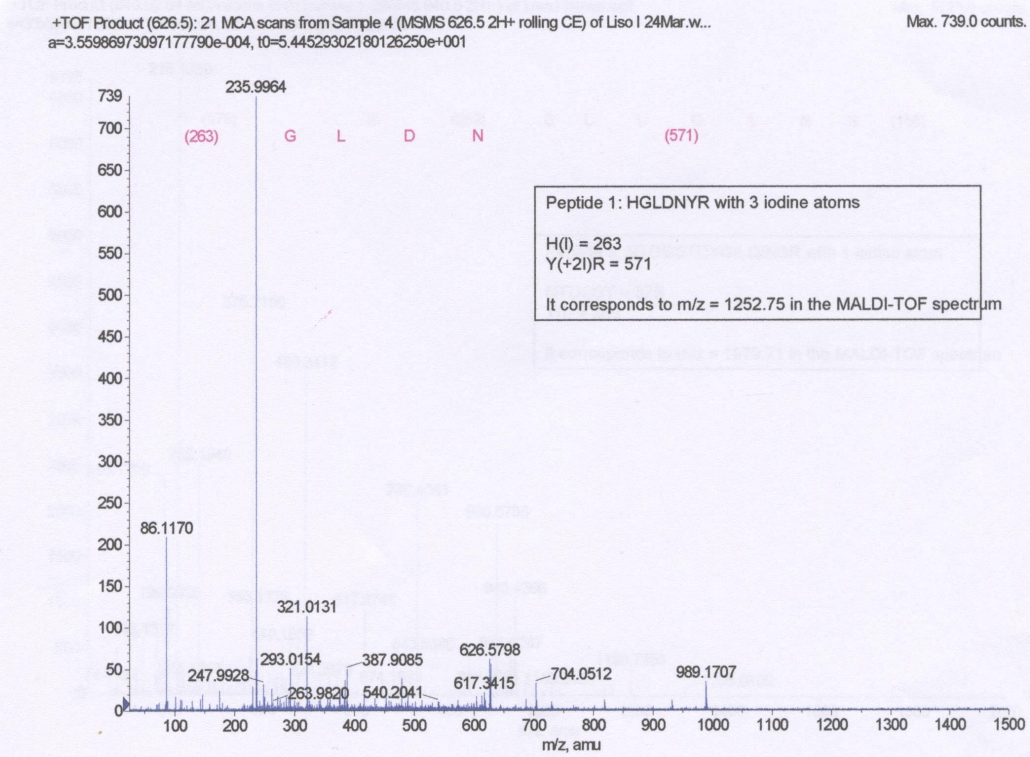
Table 2. Calculated values of hen egg white lysozyme (P00698) tryptic peptides. Maximum number of missed cleavages (MC): 0. All cysteines have been treated with iodoacetamide to form carbamidomethyl-cysteine (Cys_CAM). Methionines have been oxidized to form methionine sulfoxide (MSO). Displaying peptides with a mass bigger than 800 Dalton. Peptides 1, 2 and 3 contain Tyr residues and they are prompt to iodination. Using monoisotopic masses of the occurring amino acid residues and giving peptide masses as $[M+H]^+$, $[M+I+H]^+$, $[M+2I+H]^+$ and $[M+3I+H]^+$ for non-iodinated, mono-, di- and triiodinated peptides, respectively. For details on mass spectrometry results, see Figure 2 in the article.

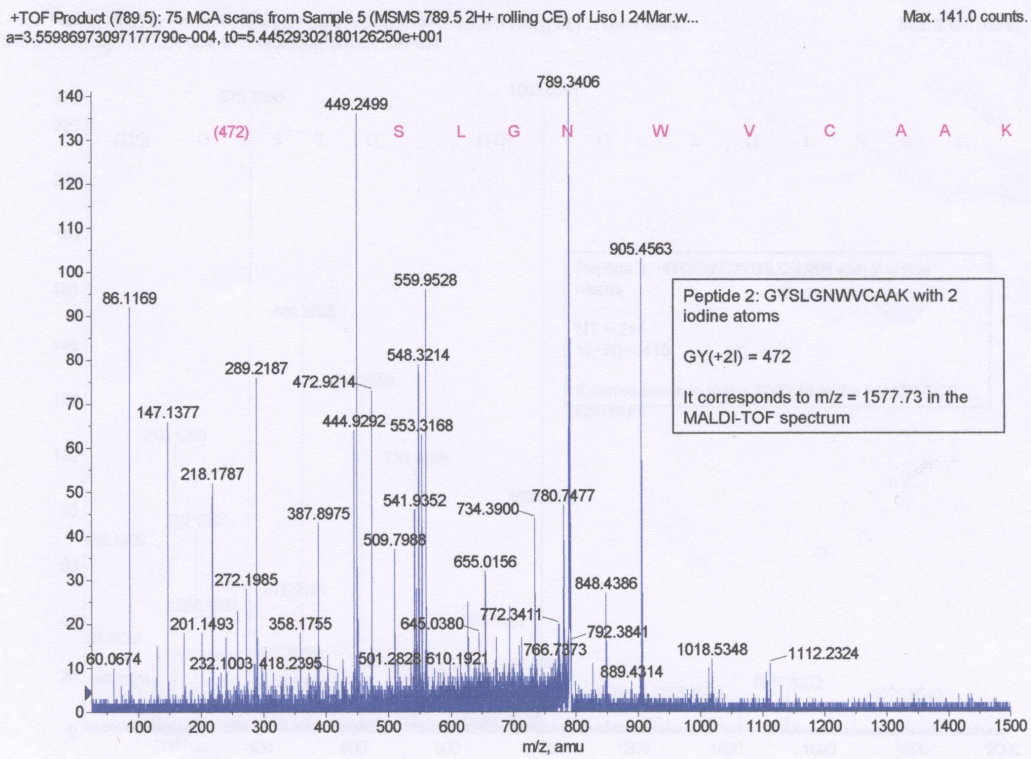
Supplementary material (continued)

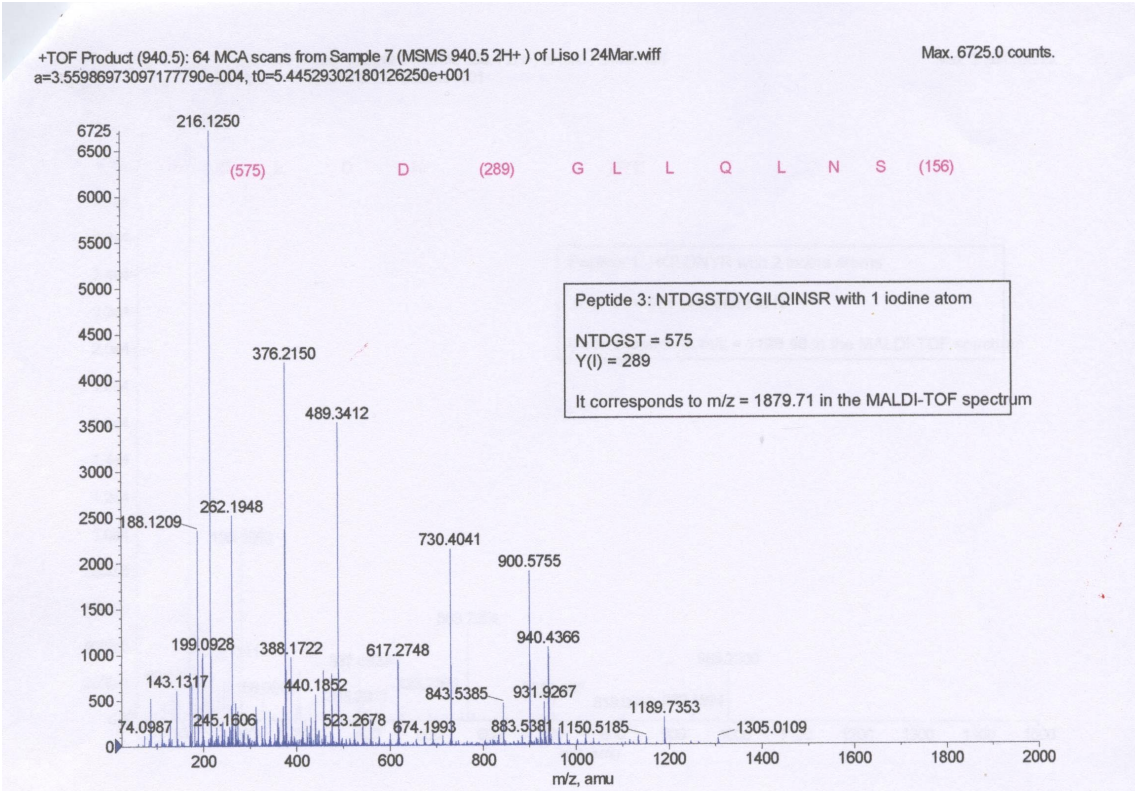
B) MS/MS sequence analysis:

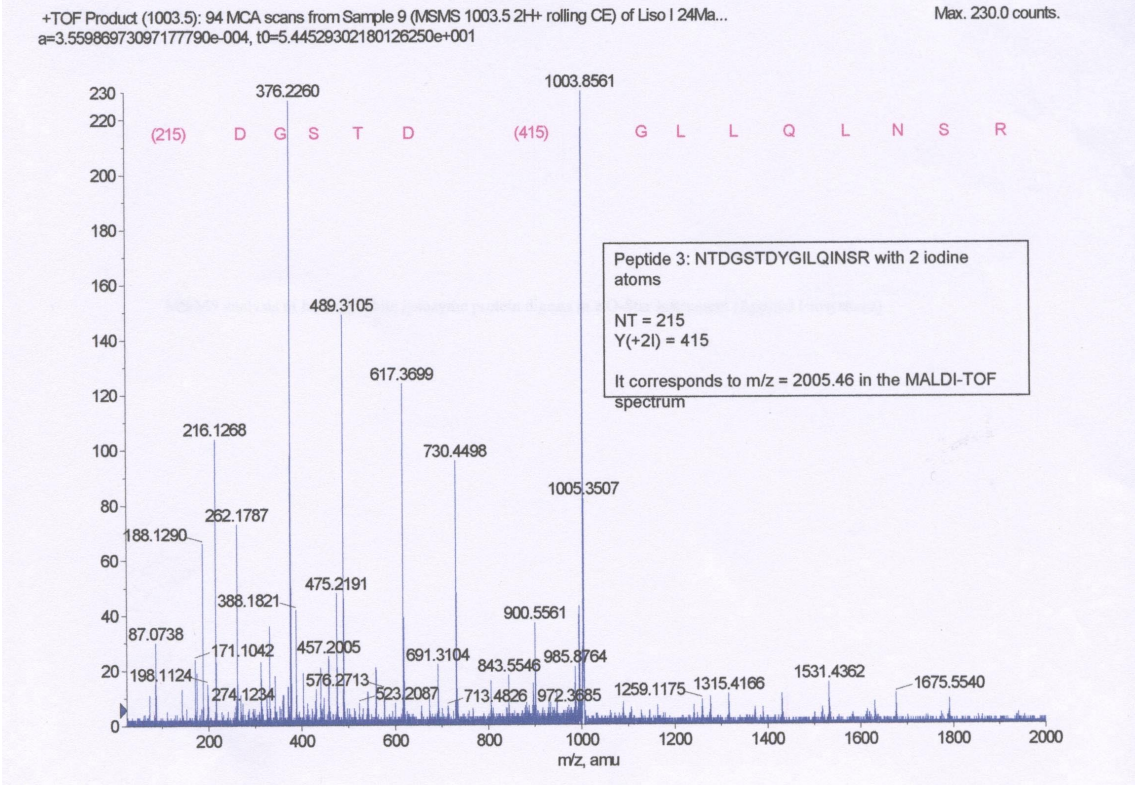


MS/MS analysis of hen egg white lysozyme protein digests in a Q-Star instrument (Applied Biosystems)


+TOF Product (626.5): 21 MCA scans from Sample 4 (MSMS 626.5 2H+ rolling CE) of Liso I 24Mar.w...
a=3.55986973097177790e-004, t0=5.44529302180126250e+001
Max. 739.0 counts.
(263) G L D N (571)
Peptide 1: HGLDNYR with 3 iodine atoms
H(I) = 263
Y(+2I)R = 571
It corresponds to m/z = 1252.75 in the MALDI-TOF spectrum
235.9964
86.1170
321.0131
293.0154
387.9085
626.5798
247.9928
263.9820
540.2041
617.3415
704.0512
989.1707
m/z, amu


+TOF Product (789.5): 75 MCA scans from Sample 5 (MSMS 789.5 2H+ rolling CE) of Liso I 24Mar.w...
a=3.55986973097177790e-004, t0=5.44529302180126250e+001
Max. 141.0 counts.
(472) S L G N W V C A A K
Peptide 2: GYSLGNWVCAAK with 2 iodine atoms
GY(+2I) = 472
It corresponds to m/z = 1577.73 in the MALDI-TOF spectrum
789.3406
449.2499
905.4563
559.9528
86.1169
548.3214
289.2187
472.9214
147.1377
444.9292
553.3168
218.1787
780.7477
541.9352
734.3900
387.8975
509.7988
655.0156
272.1985
848.4386
201.1493
358.1755
645.0380
772.3411
792.3841
60.0674
232.1003
418.2395
501.2828
610.1921
766.7873
889.4314
1018.5348
1112.2324
m/z, amu


+TOF Product (940.5): 64 MCA scans from Sample 7 (MSMS 940.5 2H+) of Liso I 24Mar.wiff
a=3.55986973097177790e-004, t0=5.44529302180126250e+001
Max. 6725.0 counts.
(575) D (289) G L L Q L N S (156)
Peptide 3: NTDGSTDYGILQINSR with 1 iodine atom
NTDGST = 575
Y(I) = 289
It corresponds to m/z = 1879.71 in the MALDI-TOF spectrum
216.1250
376.2150
489.3412
262.1948
188.1209
730.4041
900.5755
940.4366
199.0928
388.1722
617.2748
143.1317
440.1852
843.5385
931.9267
1189.7353
74.0987
245.1606
523.2678
674.1993
883.5381
1150.5185
1305.0109
m/z, amu


+TOF Product (1003.5): 94 MCA scans from Sample 9 (MSMS 1003.5 2H+ rolling CE) of Liso I 24Ma...
a=3.55986973097177790e-004, t0=5.44529302180126250e+001
Max. 230.0 counts.
(215) D G S T D (415) G L L Q L N S R
Peptide 3: NTDGSTDYGILQINSR with 2 iodine atoms
NT = 215
Y(+2I) = 415
It corresponds to m/z = 2005.46 in the MALDI-TOF spectrum
376.2260
1003.8561
489.3105
617.3699
216.1268
730.4498
1005.3507
262.1787
188.1290
475.2191
388.1821
900.5561
87.0738
171.1042
457.2005
691.3104
843.5546
985.8764
198.1124
576.2713
1531.4362
1315.4166
1675.5540
274.1234
523.2087
713.4826
972.3685
1259.1175
m/z, amu

GLUCANASE: Supplementary material

B) Calculated values of recombinant 1,3-1,4-beta-D-glucanohydrolase TRYPTIC peptides

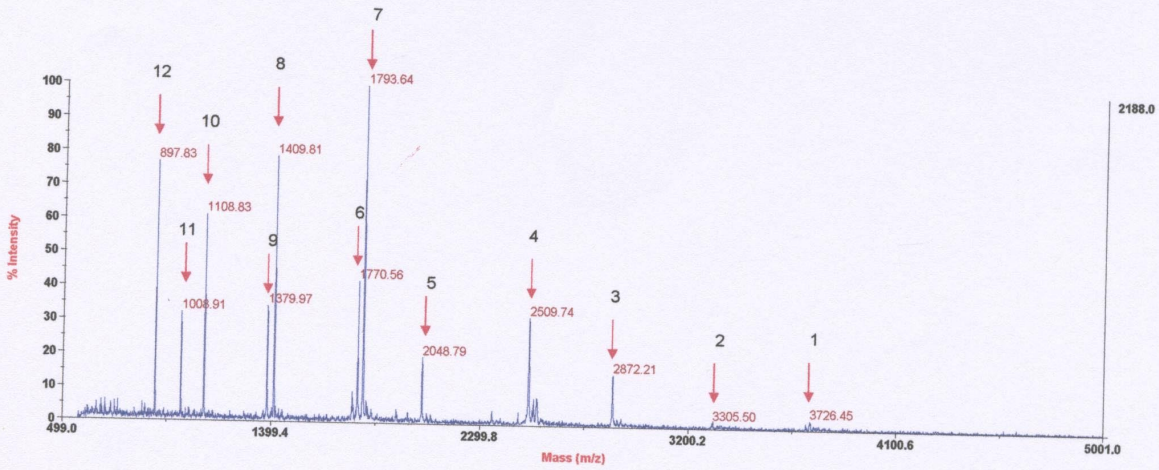
C) MALDI-TOF-MS spectrum of tryptically digested recombinant glucanase

entry	$[M+H]^+$	$[M+I+H]^+$	$[M+2I+H]^+$	#MC	position	artif. modifications		peptide sequence
1	3726.76	3852,66	-	0	181-212	MSO: 182, 183	3758.75	IMMNLWNGAGVDEWLGSYNGVTPLYAHYNWVR
2	3305.59	3431,49	-	0	86-115			NVGIVSSFFTYTGPTDGTPWDEIDIEFLGK
3	2871.37	-	-	0	135-159			IVNLGFDAANSYHTYAFDWQPNSIK
4	2509.14	2635,04	2760,94	0	1-22			SAQTGGSFYEPFNNYNTGLWQK
5	2048.01	2173,91	-	0	68-85	MSO: 81	2064.01	SVQTYGYGLYEVNMKPAK
6	1769.82	1895,72	-	0	120-134			VQFNYYTNGVGNHEK
7	1735.69	1918,62[a]	-	0	23-37	Cys_CAM: 34 MSO: 31	1792.72 1751.69	ADGYSNGNMFNCTWR
8	1409.65	-	-	0	38-50	MSO: 43, 49	1441.64	ANNVSMTSLGEMR
9	1379.73	-	-	0	168-180			HTATTQIPQTPGK
10	1109.58	1235,48	1361,38	0	51-60			LSLTSPSYNK
11	1008.51	1134,41	-	0	160-167			WYVDGQLK
12	840.33	-	-	0	61-67	Cys_CAM: 63	897.35	FDCGENR

[a] This mass includes the carbamidomethylation of Cys, as it is found experimentally.

Table 3. Calculated values of recombinant 1,3-1,4-β-D-glucanohydrolase tryptic peptides. Maximum number of missed cleavages (MC): 0. All cysteines have been treated with iodoacetamide to form carbamidomethyl-cysteine (Cys_CAM). Methionines have been oxidized to form methionine sulfoxide (MSO). Displaying peptides with a mass bigger than 500 Dalton. Using monoisotopic masses of the occurring amino acid residues and giving peptide masses as $[M+H]^+$, $[M+I+H]^+$ $[M+2I+H]^+$ for non-iodinated, mono- and diiodinated peptides, respectively. In the case of iodinated peptides only masses that have been detected are indicated in the table. For details on experimental results, see the following MALDI-TOF spectra.

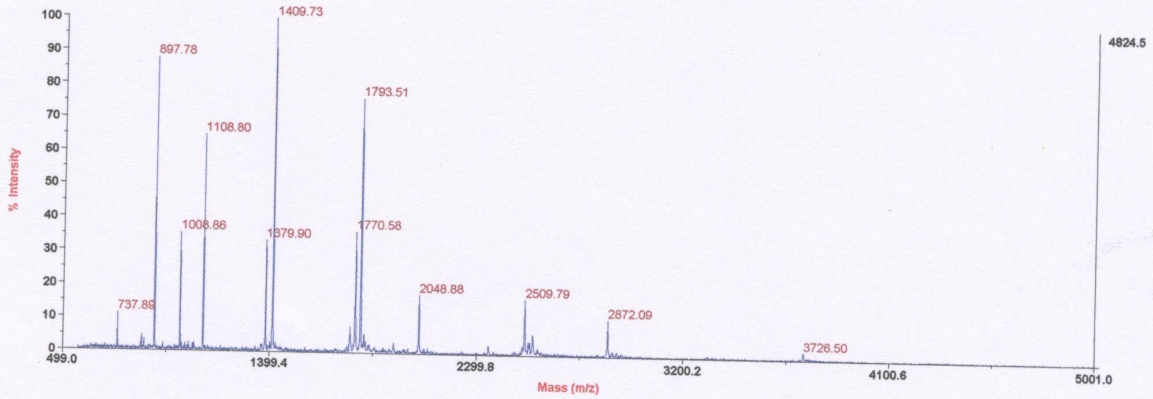
GLUCANASE: MALDI-TOF-MS spectrum of tryptically digested recombinant glucanase



MALDI-TOF spectrum of recombinant 1,3-1,4-β-D-glucanohydrolase tryptic peptides.
Numbers on top of the arrows correspond to the entry number in Table 3.

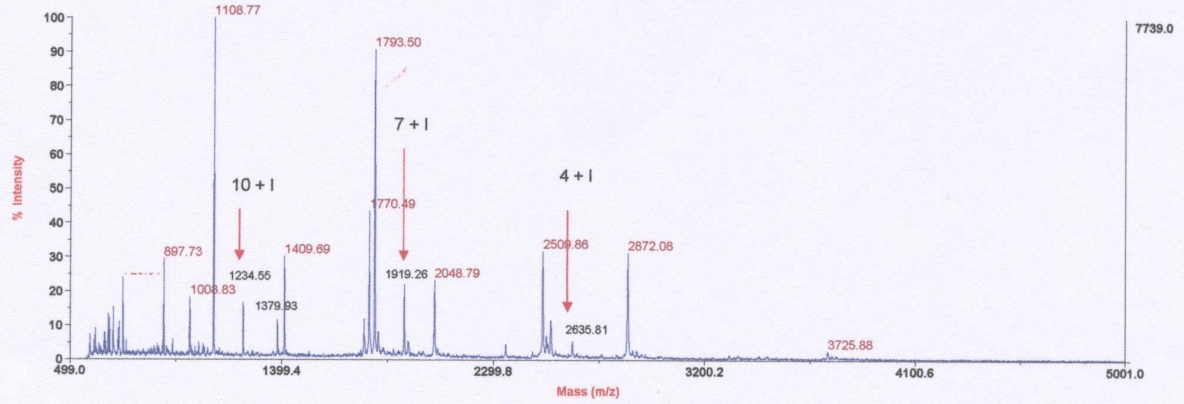
GLUCANASE: MALDI-TOF-MS spectrum of tryptically digested recombinant glucanase treated with IPy_2BF_4

1,1 equiv



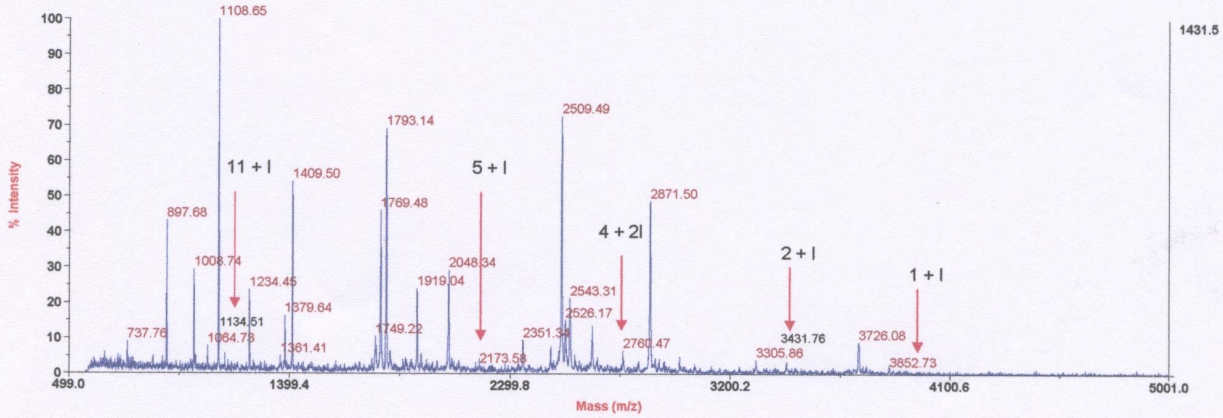
MALDI-TOF spectrum of tryptically digested recombinant 1,3-1,4-β-D-glucanohydrolase treated with 1.1 equiv. IPy_2BF_4. No changes have been observed vs. the previous spectrum with respect to the masses in the Table 3.

2.2 equiv



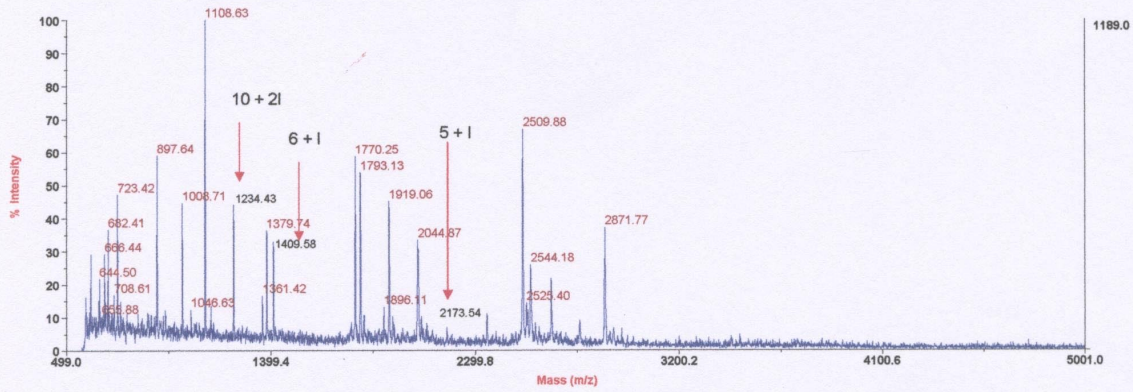
MALDI-TOF spectrum of tryptically digested recombinant 1,3-1,4-β-D-glucanohydrolase treated with 2.2 equiv. IPy_2BF_4. Masses corresponding to iodination are indicated by arrows and the entry number from Table 3.

3.3 equiv



MALDI-TOF spectrum of tryptically digested recombinant 1,3-1,4-β-D-glucanohydrolase treated with 3.3 equiv. IPy_2BF_4. Masses corresponding to iodination are indicated by arrows and the entry number from Table 3. Only new iodination peaks with respect to the protein treated with 2.2 equiv. IPy_2BF_4 are shown.

7.7 equiv



MALDI-TOF spectrum of tryptically digested recombinant 1,3-1,4-β-D-glucanohydrolase treated with 7.7 equiv. IPy_2BF_4. Masses corresponding to iodination are indicated by arrows and the entry number from Table 3. Only new iodination peaks with respect to the protein treated with 3.3 equiv. IPy_2BF_4 are shown.