



(86) Date de dépôt PCT/PCT Filing Date: 2015/08/11
(87) Date publication PCT/PCT Publication Date: 2016/02/18
(85) Entrée phase nationale/National Entry: 2017/02/13
(86) N° demande PCT/PCT Application No.: US 2015/044697
(87) N° publication PCT/PCT Publication No.: 2016/025506
(30) Priorités/Priorities: 2014/08/11 (US62/036,032);
2014/08/11 (US62/036,024); 2014/08/11 (US62/035,968);
2014/08/11 (US62/036,016); 2014/08/11 (US62/036,007);
2014/08/11 (US62/035,997); 2014/08/11 (US62/035,981)

(51) Cl.Int./Int.Cl. *G01N 30/72* (2006.01),
G01N 27/62 (2006.01), *G01N 30/02* (2006.01),
G01N 33/68 (2006.01)

(71) Demandeur/Applicant:
DOW AGROSCIENCES LLC, US

(72) Inventeurs/Inventors:
OMAN, TRENT JAMES, US;
SCHAFER, BARRY WILLIAM, US;
HILL, RYAN CHRISTOPHER, US;
SHAN, GUOMIN, US

(74) Agent: SMART & BIGGAR

(54) Titre : SYSTEMES ET PROCEDES DE QUANTIFICATION ET DE DETECTION SELECTIVE D'ALLERGENES
(54) Title: SYSTEMS AND METHODS FOR SELECTIVE QUANTITATION AND DETECTION OF ALLERGENS

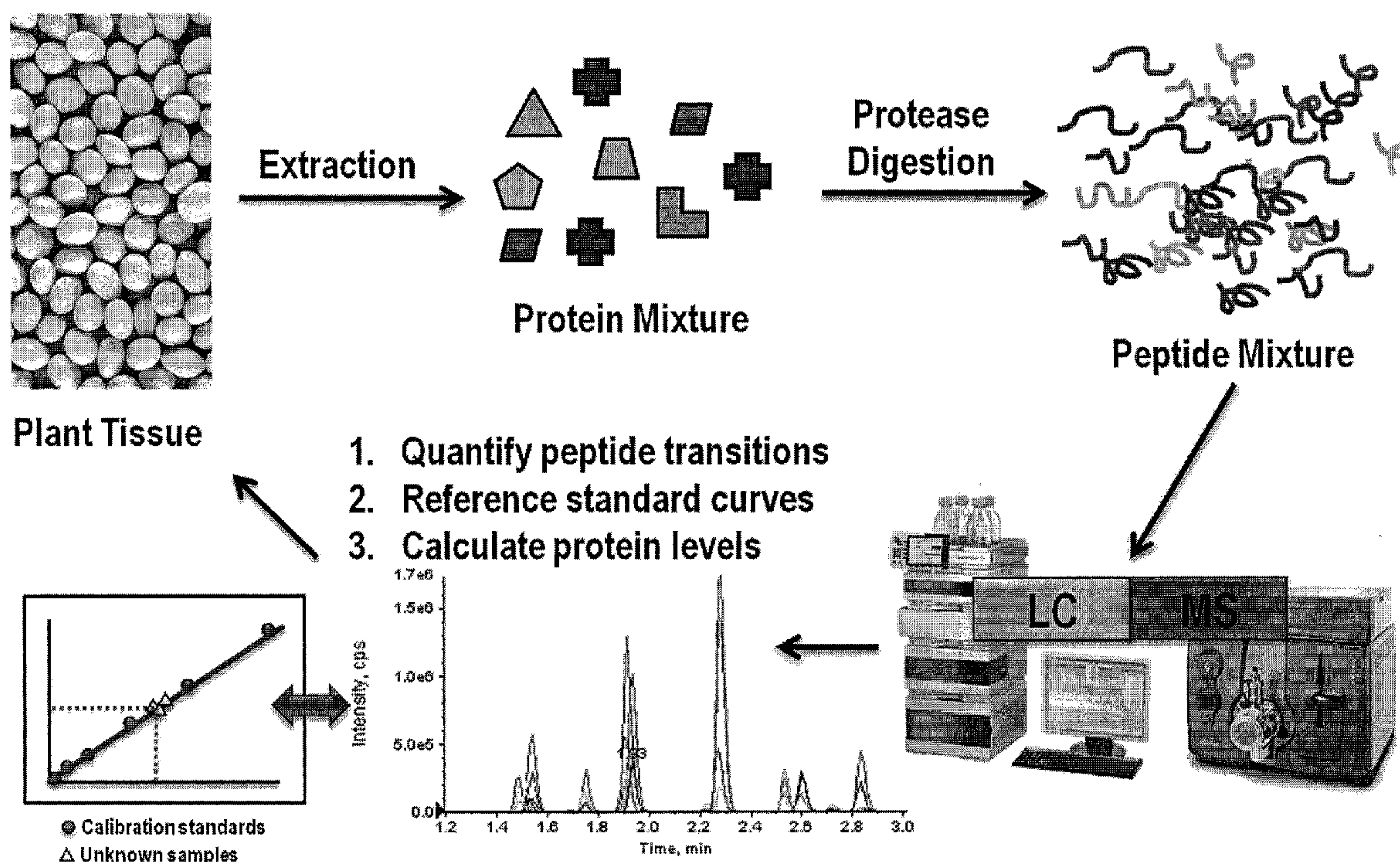


FIG. 1

(57) **Abrégé/Abstract:**

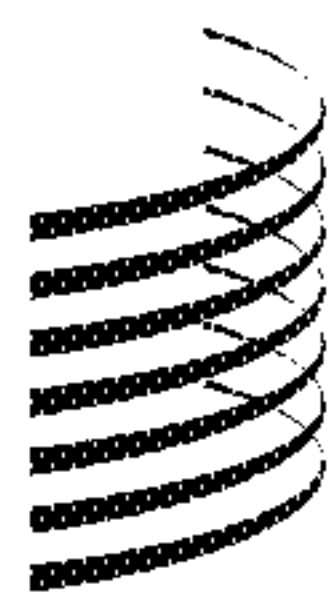
The invention relates to methods and systems taking advantage of bioinformatic investigations to identify candidate signature peptides for quantitative multiplex analysis of complex protein samples from plants, plant parts, and/or food products using mass spectroscopy. Provided are use and methods for selecting candidate signature peptides for quantitation using a bioinformatic approach. Also provided are systems comprising a chromatography and mass spectrometry for using selected signature peptides.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau



(10) International Publication Number

WO 2016/025506 A1

(43) International Publication Date

18 February 2016 (18.02.2016)

WIPO | PCT

(51) International Patent Classification:

G01N 30/72 (2006.01) G01N 27/62 (2006.01)

G01N 30/02 (2006.01) G01N 33/68 (2006.01)

(21) International Application Number:

PCT/US2015/044697

(22) International Filing Date:

11 August 2015 (11.08.2015)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/035,968 11 August 2014 (11.08.2014) US

62/035,981 11 August 2014 (11.08.2014) US

62/035,997 11 August 2014 (11.08.2014) US

62/036,007 11 August 2014 (11.08.2014) US

62/036,016 11 August 2014 (11.08.2014) US

62/036,024 11 August 2014 (11.08.2014) US

62/036,032 11 August 2014 (11.08.2014) US

(71) Applicant: DOW AGROSCIENCES LLC [US/US];

9330 Zionsville Road, Indianapolis, Indiana 46268 (US).

(72) Inventors: OMAN, Trent James; 10835 Independence

Way, Carmel, Indiana 46032 (US). SCHAFER, Barry

William; 23229 Marin Drive, Cicero, Indiana 46034 (US).

HILL, Ryan Christopher; 6045 Woodmill Drive, Fishers,

Indiana 46038 (US). SHAN, Guomin; 413 Ira Way, Carmel, Indiana 46032 (US).

(74) Agent: CATAXINOS, Edgar R.; PO Box 2550, Salt Lake City, Utah 84110 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: SYSTEMS AND METHODS FOR SELECTIVE QUANTITATION AND DETECTION OF ALLERGENS

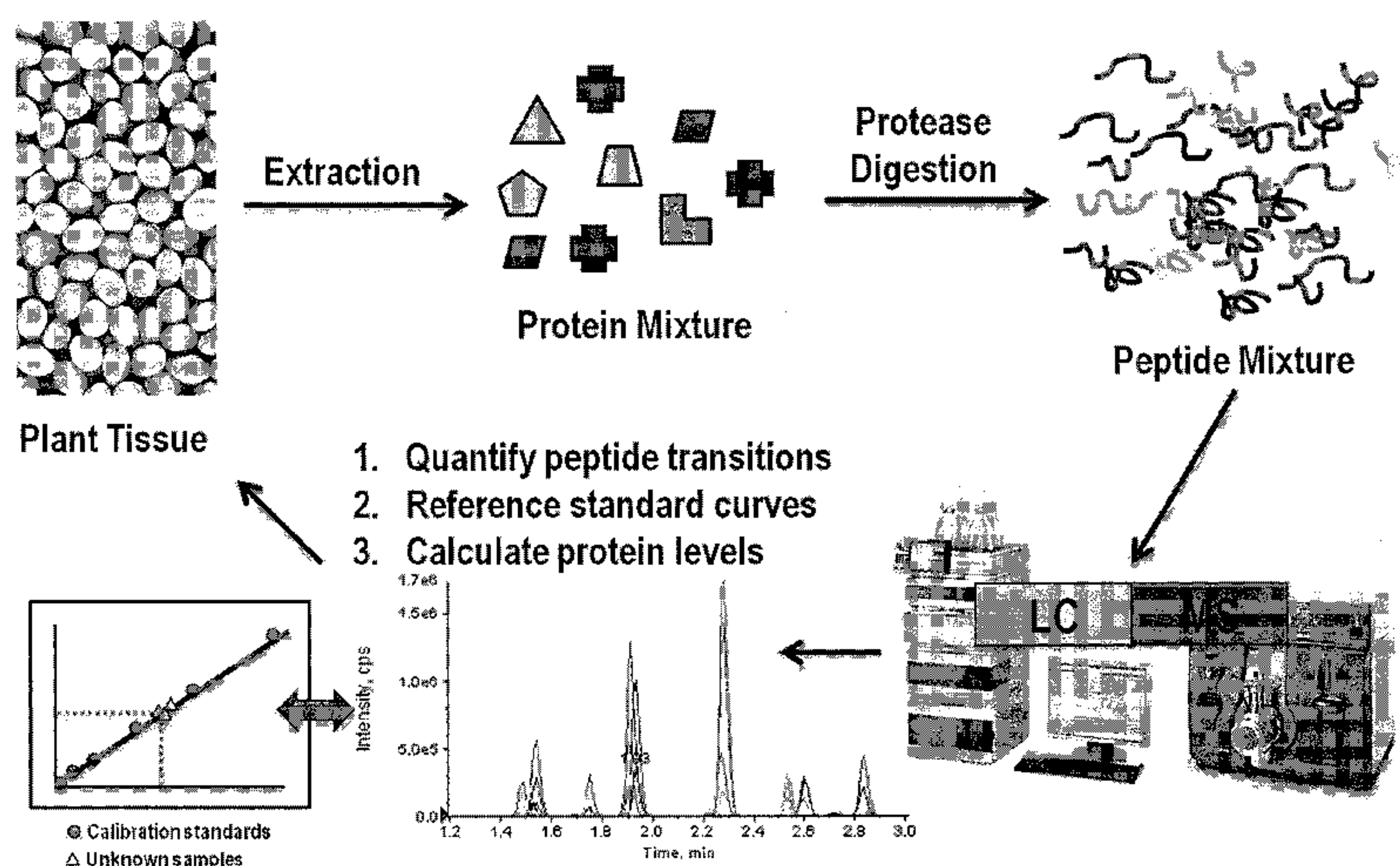


FIG. 1

(57) Abstract: The invention relates to methods and systems taking advantage of bioinformatic investigations to identify candidate signature peptides for quantitative multiplex analysis of complex protein samples from plants, plant parts, and/or food products using mass spectroscopy. Provided are use and methods for selecting candidate signature peptides for quantitation using a bioinformatic approach. Also provided are systems comprising a chromatography and mass spectrometry for using selected signature peptides.

WO 2016/025506 A1

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 29

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 29

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

SYSTEMS AND METHODS FOR SELECTIVE QUANTITATION AND DETECTION OF ALLERGENS

PRIORITY CLAIM

5 This application claims the benefit of the filing date of United States Provisional
Patent Applications: Serial No. 62/035,968, filed August 11, 2014, for “Methods and
Systems for Selective Quantitation and Detection of Allergens Including GLY M BD 28
K”; Serial No. 62/035,981, filed August 11, 2014, for “Methods and Systems for Selective
Quantitation and Detection of Allergens Including GLY M BD 30 K”; Serial No.
10 62/035,997, filed August 11, 2014, for “Methods and Systems for Selective Quantitation
and Detection of Allergens Including Kunitz Trypsin Inhibitor 1”; Serial No. 62/036,007,
filed August 11, 2014, for “Methods and Systems for Selective Quantitation and Detection
of Allergens Including Kunitz Trypsin Inhibitor 3”; Serial No. 62/036,016, filed August 11,
2014, for “Methods and Systems for Selective Quantitation and Detection of Allergens
15 Including GLY M 8 (2S Albumin)”; Serial No. 62/036,024, filed August 11, 2014, for
“Methods and Systems for Selective Quantitation and Detection of Allergens Including
Lectin”; and Serial No. 62/036,032, filed August 11, 2014, for “Methods and Systems for
Selective Quantitation and Detection of Allergens Including Lipoxxygenase,” the disclosure
of each of which is hereby incorporated herein in its entirety by this reference.

20

BACKGROUND

 The current methods for analysis of gene expression in plants that are preferred in
the art include DNA-based techniques (for example PCR and/or RT-PCR); the use of
reporter genes; Southern blotting; and immunochemistry. All of these methodologies
25 suffer from various shortcomings. Detection of known and potential allergens in plants,
plant parts, and/or food products is an important subject for public safety.

 Although mass spectrometry has been disclosed previously, existing approaches are
limited without selected and sensitive quantitation. There remains a need for a high-
throughput method for selected and sensitive quantitation of known and/or potential
30 allergens in plant, plant parts, and/or food products.

DISCLOSURE

The invention relates to methods and systems taking advantage of bioinformatic investigations to identify candidate signature peptides for quantitative multiplex analysis of complex protein samples from plants, plant parts, and/or food products using mass spectrometry. Provided are use and methods for selecting candidate signature peptides for quantitation using a bioinformatic approach. Also provided are systems comprising a chromatography and mass spectrometry for using selected signature peptides.

In one aspect, provided is a method of selecting candidate signature peptide for quantitation of known allergen and potential allergens from a plant-based sample. The method comprises:

- (a) identifying potential allergens based on homology to at least one known allergen protein sequence;
- (b) performing sequence alignment of the at least one known allergen and potential allergens identified in step (a);
- (c) selecting a consensus sequence or representative sequence based on the sequence alignment;
- (d) determining a plural of candidate signature peptides based on conservative regions or domains from the sequence alignment and *in silico* digestion data of the consensus sequence or representative sequence selected in Step (c); and
- (e) quantitating the amount of the at least one known allergen and potential allergens in the plant-based sample based on measurements of the signature peptides.

In one embodiment, the quantitating step uses a column chromatography and mass spectrometry. In another embodiment, the quantitating step comprises measuring the plural of candidate signature peptides using high resolution accurate mass spectrometry (HRAM MS). In another embodiment, the quantitating step comprises calculating corresponding peak heights or peak areas of the candidate signature peptides from mass spectrometry. In another embodiment, the quantitating step comprises comparing data from high fragmentation mode and low fragmentation mode from mass spectrometry.

In one embodiment, the at least one known allergen comprises at least one allergen selected from the group consisting of Gly m 1, Gly m 3, Gly m 4, Gly m 5 (beta-conglycinin), Gly m 6 (Glycinin) G1, Gly m 6 (Glycinin) G2, Gly m 6 (Glycinin) G3, Gly

m 6 (Glycinin) G4, Gly m 6 (Glycinin) precursor, Kunitz trypsin inhibitor 1, Kunitz trypsin inhibitor 3, Gly m Bd 28 K, Gly m Bd 30 K, Gly m 8 (2S albumin), Lectin, and lipoxygenase. In another embodiment, the at least one known allergen comprises Gly m Bd 28 K, Gly m Bd 30 K, Kunitz trypsin inhibitor 1, Kunitz trypsin inhibitor 3, Gly m 8
 5 (2S albumin), Lectin, or lipoxygenase.

In another embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 21, 28, 29, 30, 31, 32, or 33 for Gly m Bd 28 K. In a further embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 21 or 28 for Gly m Bd 28 K. In another embodiment, the consensus sequence or
 10 representative sequence of step (c) comprises SEQ ID NO: 22, 43, or 44 for Gly m Bd 30 K. In a further embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 22 or 43 for Gly m Bd 30 K. In another embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 23, 52, 53, or 54 for Kunitz trypsin inhibitor 1. In a further embodiment, the consensus sequence
 15 or representative sequence of step (c) comprises SEQ ID NO: 23 or 52 for Kunitz trypsin inhibitor 1. In another embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 24, 61, 62, or 63 for Kunitz trypsin inhibitor 3. In a further embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 24 or 61 for Kunitz trypsin inhibitor 3. In another embodiment, the consensus
 20 sequence or representative sequence of step (c) comprises SEQ ID NO: 25, 72, 73, or 74 for Gly m 8 (2S albumin). In a further embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 25 for Gly m 8 (2S albumin). In another embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 26, 82, 83, 84, 85, 86, 87, 88, 89, or 90 for Lectin. In a further
 25 embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 26 or 82 for Lectin. In another embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 27, 95, 96, 97, 98, 99, 100, 101, 102, 103, or 104 for lipoxygenase. In a further embodiment, the consensus sequence or representative sequence of step (c) comprises SEQ ID NO: 27 or 95 for lipoxygenase.

30 In one embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 21 and 29-33 for Gly m Bd 28 K. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 22 and 43-44 for Gly m Bd 30 K. In another embodiment, the potential allergens comprise at least one sequence

selected from SEQ ID NOs: 23 and 53-54 for Kunitz trypsin inhibitor 1. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 24 and 62-63 for Kunitz trypsin inhibitor 3. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 25 and 72-74 for Gly m 8 (2S albumin). In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 26 and 83-90 for Lectin. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 27 and 96-104 for lipoxygenase.

In another embodiment, the candidate signature peptides comprise at least one sequence selected from SEQ ID NOs: 9 and 34-42 for Gly m Bd 28 K. In another embodiment, the candidate signature peptides comprise SEQ ID NOs: 34-42 for Gly m Bd 28 K. In another embodiment, the candidate signature peptides comprise at least one sequence selected from SEQ ID NOs: 10 and 45-51 for Gly m Bd 30 K. In another embodiment, the candidate signature peptides comprise SEQ ID NOs: 45-51 for Gly m Bd 30 K. In another embodiment, the candidate signature peptides comprise at least one sequence selected from SEQ ID NOs: 7 and 55-60 for Kunitz trypsin inhibitor 1. In another embodiment, the candidate signature peptides comprise SEQ ID NOs: 55-60 for Kunitz trypsin inhibitor 1. In another embodiment, the candidate signature peptides comprise at least one sequence selected from SEQ ID NOs: 8 and 64-71 for Kunitz trypsin inhibitor 3. In another embodiment, the candidate signature peptides comprise SEQ ID NOs: 64-71 for Kunitz trypsin inhibitor 3. In another embodiment, the candidate signature peptides comprise at least one sequence selected from SEQ ID NOs: 11 and 75-81 for Gly m 8 (2S albumin). In another embodiment, the candidate signature peptides comprise SEQ ID NOs: 75-81 for Gly m 8 (2S albumin). In another embodiment, the candidate signature peptides comprise at least one sequence selected from SEQ ID NOs: 91-94 for Lectin. In another embodiment, the candidate signature peptides comprise SEQ ID NOs: 91-94 for Lectin. In another embodiment, the candidate signature peptides comprise at least one sequence selected from SEQ ID NOs: 105-120 for lipoxygenase. In another embodiment, the candidate signature peptides comprise SEQ ID NOs: 105-120 for lipoxygenase. In another embodiment, the plant-based sample comprises a soybean seed or part of a soybean seed.

In another aspect, provided is a system for quantitating one or more protein of interest with known amino acid sequence in a plant-based sample. The system comprises:

- (a) a high-throughput means for extracting proteins from a plant-based sample;
- (b) a process module for digesting extracted proteins with at least one protease;
- 5 (c) a separation module for separating peptides in a single step;
- (d) a selection module for selecting a plural of signature peptides for at least one known allergen and potential allergens; and
- (e) a mass spectrometry for measuring the plural of signature peptides.

In one embodiment, the separation module comprises a column chromatography.

10 In a further embodiment, the column chromatography comprises a liquid column chromatography. In another embodiment, the mass spectrometry comprises a high resolution accurate mass spectrometry (HRAM MS). In another embodiment, the selection module uses a method provided herein.

In one embodiment, the one or more protein of interest with known amino acid

15 sequence in a plant-based sample comprises potential allergens. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 21 and 29-33 for Gly m Bd 28 K. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 22 and 43-44 for Gly m Bd 30 K. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID

20 NOs: 23 and 53-54 for Kunitz trypsin inhibitor 1. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 24 and 62-63 for Kunitz trypsin inhibitor 3. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 25 and 72-74 for Gly m 8 (2S albumin). In another embodiment, the potential allergens comprise at least one sequence selected from

25 SEQ ID NOs: 26 and 83-90 for Lectin. In another embodiment, the potential allergens comprise at least one sequence selected from SEQ ID NOs: 27 and 96-104 for lipoxygenase.

In another embodiment, the signature peptides comprise at least one sequence selected from SEQ ID NOs: 9 and 34-42 for Gly m Bd 28 K. In another embodiment, the

30 signature peptides comprise SEQ ID NOs: 34-42 for Gly m Bd 28 K. In another embodiment, the signature peptides comprise at least one sequence selected from SEQ ID NOs: 10 and 45-51 for Gly m Bd 30 K. In another embodiment, the signature peptides comprise SEQ ID NOs: 45-51 for Gly m Bd 30 K. In another embodiment, the signature

peptides comprise at least one sequence selected from SEQ ID NOs: 7 and 55-60 for Kunitz trypsin inhibitor 1. In another embodiment, the signature peptides comprise SEQ ID NOs: 55-60 for Kunitz trypsin inhibitor 1. In another embodiment, the signature peptides comprise at least one sequence selected from SEQ ID NOs: 8 and 64-71 for Kunitz trypsin inhibitor 3. In another embodiment, the signature peptides comprise SEQ ID NOs: 64-71 for Kunitz trypsin inhibitor 3. In another embodiment, the signature peptides comprise at least one sequence selected from SEQ ID NOs: 11 and 75-81 for Gly m 8 (2S albumin). In another embodiment, the signature peptides comprise SEQ ID NOs: 75-81 for Gly m 8 (2S albumin). In another embodiment, the signature peptides comprise at least one sequence selected from SEQ ID NOs: 91-94 for Lectin. In another embodiment, the signature peptides comprise SEQ ID NOs: 91-94 for Lectin. In another embodiment, the signature peptides comprise at least one sequence selected from SEQ ID NOs: 105-120 for lipoxygenase. In another embodiment, the signature peptides comprise SEQ ID NOs: 105-120 for lipoxygenase. In another embodiment, the plant-based sample comprises a soybean seed or part of a soybean seed.

In another aspect, provided is a high-throughput method of quantitating at least one allergen with known amino acid sequence and homologous potential allergens in a plant-based sample. The method comprises using the system provided herein.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1 shows a representative analysis work flow for the methods and systems disclosed herein.

FIGs. 2A-2J show representative SRM LC-MS/MS for selected signature peptides SEQ ID NO: 9 NKPQFLAGAASLLR; SEQ ID NO: 34 LGFIYDDELAER; SEQ ID NO: 35 TVVEEIFSK; SEQ ID NO: 36 MMQDQEEDDEEEK; SEQ ID NO: 37 NAYGWSK; SEQ ID NO: 38 ALHGGEYPPLSEPDIGVLLVK; SEQ ID NO: 39 QGDVFVVP; SEQ ID NO: 40 YFPFCQVASR; SEQ ID NO: 41 TLMGPELSAAGVSEDTLR; SEQ ID NO: 42 SFANDVVMDF from trypsin digested soybean sample chromatogram for Gly m Bd 28 K.

FIG. 2K shows representative SRM LC-MS/MS Standard Chromatogram - 500 ng/mL Synthetic Peptide SEQ ID NO: 9 NKPQFLAGAASLLR natural abundance peptide and heavy isotope labeled peptide transitions.

FIG. 2L shows sequences alignments among Gly m Bd 28 K and potential homologs of Gly m Bd 28 K.

FIGs. 3A-3H show representative SRM LC-MS/MS for selected signature peptides
 SEQ ID NO: 10 GVITQVK; SEQ ID NO: 45 SILDLDTK; SEQ ID NO: 46 FTTQK;
 5 SEQ ID NO: 47>NNLNYIR; SEQ ID NO: 48 FADITPQEFSK; SEQ ID NO: 49
 EQYSCDHPPASWDWR; SEQ ID NO: 50 VTIDGYETLIMSDESTESETEQAFL
 SAILEQPISVSIDAK; and SEQ ID NO: 51 NTGNLLGVCGMNYFASYPTK; from
 trypsin digested soybean sample chromatogram for Gly m Bd 30 K.

FIG. 3I shows representative SRM LC-MS/MS Standard Chromatogram - 500
 10 ng/mL Synthetic Peptide SEQ ID NO: 10 GVITQVK natural abundance peptide and heavy
 isotope labeled peptide transitions.

FIG. 3J shows sequences alignments among Gly m Bd 30 K and potential homologs of Gly m Bd 30 K.

FIGs. 4A-4G show representative SRM LC-MS/MS for selected signature peptides
 15 SEQ ID NO: 7 GGGIEVDSTGK; SEQ ID NO: 55 EICPLTVVQSPNELDK; SEQ ID NO:
 56 EGLQAVK; SEQ ID NO: 57 LVFCPQQAEDNK; SEQ ID NO: 58
 CEDIGIQIDDDGIR; SEQ ID NO: 59 LVLSK; and SEQ ID NO: 60 NKPLVVQFQK from
 trypsin digested soybean sample chromatogram for Kunitz trypsin inhibitor 1.

FIG. 4H shows representative SRM LC-MS/MS Standard Chromatogram - 500
 20 ng/mL Synthetic Peptide SEQ ID NO: 7 GGGIEVDSTGK natural abundance peptide and
 heavy isotope labeled peptide transitions.

FIG. 4I shows sequences alignments among Kunitz trypsin inhibitor 1 and potential homologs of Kunitz trypsin inhibitor 1.

FIGs. 5A-5I show representative SRM LC-MS/MS for selected signature peptides
 25 SEQ ID NO: 8 GIGTHISSPYR; SEQ ID NO: 64 CPLTVVQSR; SEQ ID NO: 65 NELDK;
 SEQ ID NO: 66 IGENK; SEQ ID NO: 67 DAMDGWFR; SEQ ID 68 LVFCPQQAEDDK;
 SEQ ID NO: 69 CGDIGISIDHDDGTR; SEQ ID NO: 70 LVVSK; and SEQ ID NO: 71
 NKPLVVQFQK from trypsin digested soybean sample chromatogram for Kunitz trypsin
 inhibitor 3.

30 FIG. 5J shows representative SRM LC-MS/MS Standard Chromatogram - 500
 ng/mL Synthetic Peptide SEQ ID NO: 8 GIGTHISSPYR natural abundance peptide and
 heavy isotope labeled peptide transitions.

FIG. 5K shows sequences alignments among Kunitz trypsin inhibitor 3 and potential homologs of Kunitz trypsin inhibitor 3.

FIGs. 6A-6H show representative SRM LC-MS/MS for selected signature peptides
 SEQ ID NO: 11 IMENQSEELEEK; SEQ ID NO: 75 WQHQQDSCR; SEQ ID NO: 76
 5 QLQGVNLTPEK; SEQ ID NO: 77 HIMEK; SEQ ID NO: 78 DEDEEEEGHMQK; SEQ
 ID NO: 79 CCTEMSELR; SEQ ID NO: 80 ELINLATMCR; and SEQ ID NO: 81
 FGPMIQCDLSSDD from trypsin digested soybean sample chromatogram for Gly m 8 (2S
 albumin).

FIG. 6I shows representative SRM LC-MS/MS Standard Chromatogram - 500
 10 ng/mL Synthetic Peptide SEQ ID NO: 11 IMENQSEELEEK natural abundance peptide
 and heavy isotope labeled peptide transitions.

FIG. 6J shows sequences alignments among Gly m 8 (2S albumin) and potential
 homologs of Gly m 8.

FIGs. 7A-7D show representative SRM LC-MS/MS for selected signature peptides
 15 SEQ ID NO: 91 VFSPNK; SEQ ID NO: 92 ANSTNTVSFTVSK; SEQ ID NO: 93
 QQNLIFQGDAISPGLR; and SEQ ID NO: 94 TADGLAFLAPVGSKPQSK from
 trypsin digested soybean sample chromatogram for Lectin.

FIG. 7E shows representative SRM LC-MS/MS Standard Chromatogram - 500
 ng/mL Synthetic Peptide SEQ ID NO: 91 VFSPNK natural abundance peptide and heavy
 20 isotope labeled peptide transitions.

FIG 7F shows sequences alignments among lectin and potential homologs of
 Lectin.

FIGs. 8A-8P show representative SRM LC-MS/MS for selected signature peptides
 SEQ ID NO: 105 SSDFLTYGIK; SEQ ID NO: 106 GTVVLMPK; SEQ ID NO: 107
 25 NVLDFNAITSIGK; SEQ ID NO: 108 GGVIDTATGILGQGVSLVGGVIDTATSFLGR;
 SEQ ID NO: 109 IFFVNDTYLPSATPAPLLK; SEQ ID NO: 110 DENFGHLK; SEQ ID
 NO: 111 SLSHDVIPLFK; SEQ ID NO: 112 SLYEGGIK; SEQ ID NO: 113
 TDGENVLQFPPHVAK; SEQ ID NO: 114 INSLPTAK; SEQ ID NO: 115 TILFLK; SEQ
 ID NO: 116 HLSVLHPIYK; SEQ ID NO: 117 QSLINADGIIK; SEQ ID NO: 118
 30 FIPAEGTPEYDEMVK; SEQ ID NO: 119 ALEAFK; and SEQ ID NO: 120 GIPNSISI
 from trypsin digested soybean sample chromatogram.

FIG. 8Q shows representative SRM LC-MS/MS Standard Chromatogram - 500 ng/mL Synthetic Peptide SEQ ID NO: 105 SSDFLTYGIK natural abundance peptide and heavy isotope labeled peptide transitions.

FIG. 8R shows sequences alignments among lipoxygenase and potential homologs
5 of lipoxygenase.

MODE(S) FOR CARRYING OUT THE INVENTION

It is of significance to enable a sensitive multiplex assay that is capable of selectively detecting and measuring levels of proteins of interest. Currently, relevant
10 technologies for protein expression detection rely heavily on traditional immunochemistry technologies which present a challenge to accommodate the volume of data required to generate per sample.

Soybean is a multi-billion dollar commodity due to its balanced composition of 2:2:1 protein, starch, and oil by weight. Many seeds, including soybeans, contain proteins
15 that are allergens and anti-nutritional factors. As such, there are concerns regarding the potential of altering allergen levels in genetically-modified soybean varieties when compared to varieties developed through traditional breeding. The measurement of allergen levels in crops has been achieved almost exclusively by immunoassays, such as enzyme-linked immunosorbent assays (ELISA) or IgE-immunoblotting; however, these
20 methods suffer from limited sensitivity and specificity and high variability.

There has been recent interest in developing LC-MS/MS based methods to quantify several plant-expressed proteins in a single analysis. Analysis using these “signature peptides” involves tracking protein expression levels by quantifying several highly specific digest fragments of the proteins of interest. This can be typically accomplished using
25 liquid chromatography coupled with selected reaction monitoring (SRM) tandem mass spectrometry. Improved multiplexed LC-MS/MS methods and systems are provided herein to enable simultaneous quantitation(s) of several allergen proteins in transgenic and non-transgenic soybean. Methods and systems provided herein are validated for analytical figures of merit including accuracy, precision, linearity, limits of detection and
30 quantitation; and for other considerations including sample throughput, transferability, and ease of use. The allergens can be quantified using a multiplexing format and samples can be harvested from the field, processed, and analyzed/quantitated for example within a day (twenty-four hours) window (from field to measured numerical value). In addition, sample

preparations of the methods and systems provided can be fully scalable for high-throughput, thus enabling hundreds of samples to be analyzed in a single batch.

Representative soybean allergens include, for example, Gly m 1, Gly m 3, Gly m 4, Gly m 5 (beta-conglycinin), Gly m 6 (Glycinin) G1, Gly m 6 (Glycinin) G2, Gly m 6 (Glycinin) G3, Gly m 6 (Glycinin) G4, Gly m 6 (Glycinin) precursor, Gly m 6 (Glycinin) G4 precursor, Kunitz trypsin inhibitor 1, Kunitz trypsin inhibitor 3, Gly m Bd 28 K, Gly m Bd 30 K, Gly m 8 (2S albumin), Lectin, and lipoxygenase.

Representative wheat allergens include, for example, profilin (Tri a12), wheat lipid transfer protein 1 (Tri a14), agglutinin isolectin 1 (Tri a18), omega-5 gliadin – seed storage protein (Tri a19), gliadin (Tri a20; NCBI Accession Nos. M10092, M11073, M11074, M11075, M11076, K03074, and K03075), thioredoxin (Tri a25), high molecular weight glutenin (Tri a26), low molecular weight glutenin (Tri a36), and alpha purothionin (Tri a37).

Representative corn allergens include, for example, maize lipid transfer protein (LTP) (Zea m14) and thioredoxin (Zea m25).

Representative corn allergens include, for example, rice profilin A (Ory s12).

In some embodiments, the methods and systems provided use liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) to detect protein expression levels of sixteen different allergens from soybean. In some embodiments, the methods and systems enable analysis of each allergen by itself or combined with additional proteins for a multiplexing assay for qualitative and quantitative analysis in plant matrices.

In some embodiments, the mass spectrometry detection for quantitative studies may be accomplished using selected reaction monitoring, performed on a triple quadrupole mass spectrometer. Using this type of instrumentation, initial mass-selection of ion (peptide) of interest formed in the source, followed by, dissociation of this precursor ion in the collision region of the MS, then mass-selection, and counting, of a specific product (daughter) ion. In some embodiments, the mass spectrometry detection for quantitative studies may be accomplished using selected reaction monitoring (SRM). Using particular type of instrumentation, initial mass-selection of ion of interest formed in the source, followed by, dissociation of this precursor (protein) ion in the collision region of the mass spectrometer (MS), then mass-selection, and counting, of a specific product (peptide) ion. In some embodiment, counts per unit time may provide an integratable peak area from which amounts or concentration of analytes can be determined. In some embodiment, the use of

high resolution accurate mass (HRAM) monitoring for quantitation, performed on a HRAM capable mass spectrometer, may include, but is not limited to, hybrid quadrupole-time-of-flight, quadrupole-orbitrap, ion trap-orbitrap, or quadrupole-ion-trap-orbitrap (tribrid) mass spectrometers. Using particular type of instrumentation, peptides are not subject to fragmentation conditions, but rather are measured as intact peptides using full scan or targeted scan modes (for example selective ion monitoring mode or SIM). Integratable peak area can be determined by generating an extracted ion chromatogram for each specific analyte and amounts or concentration of analytes can be calculated. The high resolution and accurate mass nature of the data enable highly specific and sensitive ion signals for the analyte (protein and/or peptide) of interest.

Unless otherwise stated, the following terms used in this application, including the specification and claims, have the definitions given below. It must be noted that, as used in the specification and the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise.

As used herein, the term “bioconfinement” refers to restriction of the movement of genetically modified plants or their genetic material to designated areas. The term includes physical, physicochemical, biological confinement, as well as other forms of confinement that prevent the survival, spread or reproduction of a genetically modified plants in the natural environment or in artificial growth conditions.

As used herein, the term “complex protein sample” is used to distinguish a sample from a purified protein sample. A complex protein sample contains multiple proteins, and may additionally contain other contaminants.

As used herein, the general term “mass spectrometry” or “MS” refers to any suitable mass spectrometry method, device or configuration including, e.g., electrospray ionization (ESI), matrix-assisted laser desorption/ionization (MALDI) MS, MALDI-time of flight (TOF) MS, atmospheric pressure (AP) MALDI MS, vacuum MALDI MS, or combinations thereof. Mass spectrometry devices measure the molecular mass of a molecule (as a function of the molecule's mass-to-charge ratio) by measuring the molecule's flight path through a set of magnetic and electric fields. The mass-to-charge ratio is a physical quantity that is widely used in the electrodynamics of charged particles. The mass-to-charge ratio of a particular peptide can be calculated, a priori, by one of skill in the art. Two particles with different mass-to-charge ratio will not move in the same path in a vacuum when subjected to the same electric and magnetic fields.

Mass spectrometry instruments consist of three modules: an ion source, which splits the sample molecules into ions; a mass analyzer, which sorts the ions by their masses by applying electromagnetic fields; and a detector, which measures the value of an indicator quantity and thus provides data for calculating the abundances of each ion present. The technique has both qualitative and quantitative applications. These include identifying unknown compounds, determining the isotopic composition of elements in a molecule, determining the structure of a compound by observing its fragmentation, and quantifying the amount of a compound in a sample.

A detailed overview of mass spectrometry methodologies and devices can be found in the following references which are hereby incorporated by reference: Can and Annan (1997) Overview of peptide and protein analysis by mass spectrometry. In: *Current Protocols in Molecular Biology*, edited by Ausubel, et al. New York: Wiley, p. 10.21.1-10.21.27; Paterson and Aebersold (1995) *Electrophoresis* 16: 1791-1814; Patterson (1998) Protein identification and characterization by mass spectrometry. In: *Current Protocols in Molecular Biology*, edited by Ausubel, et al. New York: Wiley, p. 10.22.1-10.22.24; and Domon and Aebersold (2006) *Science* 312(5771):212-17.

As the term is used herein, proteins and/or peptides are “multiplexed” when two or more proteins and/or peptides of interest are present in the same sample.

As used herein, a “plant trait” may refer to any single feature or quantifiable measurement of a plant.

As used herein, the phrase “peptide” or peptides” may refer to short polymers formed from the linking, in a defined order, of α -amino acids. Peptides may also be generated by the digestion of polypeptides, for example proteins, with a protease.

As used herein, the phrase “protein” or proteins” may refer to organic compounds made of amino acids arranged in a linear chain and joined together by peptide bonds between the carboxyl and amino groups of adjacent amino acid residues. The sequence of amino acids in a protein is defined by the sequence of a gene, which is encoded in the genetic code. In general, the genetic code specifies 20 standard amino acids, however in certain organisms the genetic code can include selenocysteine—and in certain archaea-pyrrolysine. The residues in a protein are often observed to be chemically modified by post-translational modification, which can happen either before the protein is used in the cell, or as part of control mechanisms. Protein residues may also be modified by design, according to techniques familiar to those of skill in the art. As used herein, the term

“protein” encompasses linear chains comprising naturally occurring amino acids, synthetic amino acids, modified amino acids, or combinations of any or all of the above.

As used herein, the term “single injection” refers to the initial step in the operation of a MS or LC-MS device. When a protein sample is introduced into the device in a single injection, the entire sample is introduced in a single step.

As used herein, the phrase “signature peptide” refers an identifier (short peptide) sequence of a specific protein. Any protein may contain an average of between 10 and 100 signature peptides. Typically signature peptides have at least one of the following criteria: easily detected by mass spectroscopy, predictably and stably eluted from a liquid chromatography (LC) column, enriched by reversed phase high performance liquid chromatography (RP-HPLC), good ionization, good fragmentation, or combinations thereof. A peptide that is readily quantified by mass spectrometry typically has at least one of the following criteria: readily synthesized, ability to be highly purified (>97%), soluble in $\leq 20\%$ acetonitrile, low non-specific binding, oxidation resistant, post-synthesis modification resistant, and a hydrophobicity or hydrophobicity index ≥ 10 and ≤ 40 . The hydrophobicity index can be calculated according to Krokhin, Molecular and Cellular Proteomics 3 (2004) 908, which is incorporated by reference. It’s known that a peptide having a hydrophobicity index less than 10 or greater than 40 may not be reproducibly resolved or eluted by a RP-HPLC column.

As used herein, the term “stacked” refers to the presence of multiple heterologous polynucleotides incorporated in the genome of a plant.

Tandem mass spectrometry: In tandem mass spectrometry, a parent ion generated from a molecule of interest may be filtered in a mass spectrometry instrument, and the parent ion subsequently fragmented to yield one or more daughter ions that are then analyzed (detected and/or quantified) in a second mass spectrometry procedure. In some embodiments, the use of tandem mass spectrometry is excluded. In these embodiments, tandem mass spectrometry is not used in the methods and systems provided. Thus, neither parent ions nor daughter ions are generated in these embodiments.

As used herein, the term “transgenic plant” includes reference to a plant which comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant expression cassette. “Transgenic” is used

herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acid including those transgenic plants initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic plant.

5 Any plants that provide useful plant parts may be treated in the practice of the present invention. Examples include plants that provide flowers, fruits, vegetables, and grains.

As used herein, the phrase "plant" includes dicotyledonous plants and monocotyledonous plants. Examples of dicotyledonous plants include tobacco,
10 Arabidopsis, soybean, tomato, papaya, canola, sunflower, cotton, alfalfa, potato, grapevine, pigeon pea, pea, Brassica, chickpea, sugar beet, rapeseed, watermelon, melon, pepper, peanut, pumpkin, radish, spinach, squash, broccoli, cabbage, carrot, cauliflower, celery, Chinese cabbage, cucumber, eggplant, and lettuce. Examples of monocotyledonous plants include corn, rice, wheat, sugarcane, barley, rye, sorghum, orchids, bamboo, banana,
15 cattails, lilies, oat, onion, millet, and triticale. Examples of fruit include banana, pineapple, oranges, grapes, grapefruit, watermelon, melon, apples, peaches, pears, kiwifruit, mango, nectarines, guava, persimmon, avocado, lemon, fig, and berries. Examples of flowers include baby's breath, carnation, dahlia, daffodil, geranium, gerbera, lily, orchid, peony, Queen Anne's lace, rose, snapdragon, or other cut-flowers or ornamental flowers, potted-
20 flowers, and flower bulbs.

The specificity allowed in a mass spectrometry approach for identifying a single protein from a complex sample is unique in that only the sequence of the protein of interest is required in order to identify the protein of interest. Compared to other formats of multiplexing, mass spectrometry is unique in being able to exploit the full length of a
25 protein's primary amino acid sequence to target unique identifier-type portions of a protein's primary amino acid sequence to virtually eliminate non-specific detection. In some embodiments of the present invention, a proteolytic fragment or set of proteolytic fragments that uniquely identifies a protein(s) of interest is used to detect the protein(s) of interest in a complex protein sample.

30 In some embodiments, disclosed methods enable the quantification or determination of ratios of multiple proteins in a complex protein sample by a single mass spectrometry analysis, as opposed to measuring each protein of interest individually multiple times and compiling the individual results into one sample result.

In some embodiments, the present disclosure also provides methods useful for the development and use of transgenic plant technology. Specifically, disclosed methods may be used to maintain the genotype of transgenic plants through successive generations. Also, some embodiments of the methods disclosed herein may be used to provide high-throughput analysis of non-transgenic plants that are at risk of being contaminated with transgenes from neighboring plants, for example, by cross-pollination. By these embodiments, bioconfinement of transgenes may be facilitated and/or accomplished. In other embodiments, methods disclosed herein may be used to screen the results of a plant transformation procedure in a high-throughput manner to identify transformants that exhibit desirable expression characteristics

The mass-to-charge ratio may be determined using a quadrupole analyzer. For example, in a “quadrupole” or “quadrupole ion trap” instrument, ions in an oscillating radio frequency field experience a force proportional to the DC potential applied between electrodes, the amplitude of the RF signal, and m/z . The voltage and amplitude can be selected so that only ions having a particular m/z travel the length of the quadrupole, while all other ions are deflected. Thus, quadrupole instruments can act as a “mass filter” and “mass detector” for the ions injected into the instrument.

Collision-induced dissociation (“CID”) is often used to generate the daughter ions for further detection. In CID, parent ions gain energy through collisions with an inert gas, such as argon, and subsequently fragmented by a process referred to as “unimolecular decomposition.” Sufficient energy must be deposited in the parent ion so that certain bonds within the ion can be broken due to increased energy.

The mass spectrometer typically provides the user with an ion scan; that is, the relative abundance of each m/z over a given range (for example 10 to 1200 amu). The results of an analyte assay, that is, a mass spectrum, can be related to the amount of the analyte in the original sample by numerous methods known in the art. For example, given that sampling and analysis parameters are carefully controlled, the relative abundance of a given ion can be compared to a table that converts that relative abundance to an absolute amount of the original molecule. Alternatively, molecular standards (e.g., internal standards and external standards) can be run with the samples and a standard curve constructed based on ions generated from those standards. Using such a standard curve, the relative abundance of a given ion can be converted into an absolute amount of the original molecule. Numerous other methods for relating the presence or amount of an ion to the

presence or amount of the original molecule are well known to those of ordinary skill in the art.

The choice of ionization method can be determined based on the analyte to be measured, type of sample, the type of detector, the choice of positive versus negative mode, etc. Ions can be produced using a variety of methods including, but not limited to, electron ionization, chemical ionization, fast atom bombardment, field desorption, and matrix-assisted laser desorption ionization (MALDI), surface enhanced laser desorption ionization (SELDI), desorption electrospray ionization (DESI), photon ionization, electrospray ionization, and inductively coupled plasma. Electrospray ionization refers to methods in which a solution is passed along a short length of capillary tube, to the end of which is applied a high positive or negative electric potential. Solution reaching the end of the tube, is vaporized (nebulized) into a jet or spray of very small droplets of solution in solvent vapor. This mist of droplets flows through an evaporation chamber which is heated to prevent condensation and to evaporate solvent. As the droplets get smaller the electrical surface charge density increases until such time that the natural repulsion between like charges causes ions as well as neutral molecules to be released.

The effluent of an LC may be injected directly and automatically (*i.e.*, “in-line”) into the electrospray device. In some embodiments, proteins contained in an LC effluent are first ionized by electrospray into a parent ion.

Various different mass analyzers can be used in liquid chromatography – mass spectrometry combination (LC-MS). Exemplary mass analyzers include, but not limited to, single quadrupole, triple quadrupole, ion trap, TOF (time of flight), and quadrupole-time of flight (Q-TOF).

The quadrupole mass analyzer may consist of 4 circular rods, set parallel to each other. In a quadrupole mass spectrometer (QMS), the quadrupole is the component of the instrument responsible for filtering sample ions, based on their mass-to-charge ratio (m/z). Ions are separated in a quadrupole based on the stability of their trajectories in the oscillating electric fields that are applied to the rods.

An ion trap is a combination of electric or magnetic fields that captures ions in a region of a vacuum system or tube. Ion traps can be used in mass spectrometry while the ion’s quantum state is manipulated.

Time-of-flight mass spectrometry (TOFMS) is a method of mass spectrometry in which an ion’s mass-to-charge ratio is determined via a time measurement. Ions are

accelerated by an electric field of known strength. This acceleration results in an ion having the same kinetic energy as any other ion that has the same charge. The velocity of the ion depends on the mass-to-charge ratio. The time that it subsequently takes for the particle to reach a detector at a known distance is measured. This time will depend on the mass-to-charge ratio of the particle (heavier particles reach lower speeds). From this time and the known experimental parameters one can find the mass-to-charge ratio of the ion.

In some embodiments, the particular instrument used by the methods and/or systems provided may comprise a high fragmentation mode and a low fragmentation mode (or alternatively a non-fragmentation mode). Such different modes may include alternating scan high and low energy acquisition methodology to generate high resolution mass data. In some embodiments, the high resolution mass data may comprise a product data set (for example data derived from product ion (fragmented ions) under the high fragmentation mode) and a precursor data set (for example data derived from precursor ions (unfragmented ions) under the low fragmentation or non-fragmentation mode).

In some embodiments, the methods and/or systems provided use a mass spectrometer comprising a filtering device that may be used in the selection step, a fragmentation device that may be used in the fragmentation step, and/or one or more mass analyzers that may be used in the acquisition and/or mass spectrum creation step or steps.

The filtering device and/or mass analyzer may comprise a quadrupole. The selection step and/or acquisition step and/or mass spectrum creation step or steps may involve the use of a resolving quadrupole. Additionally or alternatively, the filtering device may comprise a two-dimensional or three-dimensional ion trap or time-of-flight (ToF) mass analyzer. The mass analyzer or mass analyzers may comprise or further comprise one or more of a time-of-flight mass analyzer and/or an ion cyclotron resonance mass analyzer and/or an orbitrap mass analyzer and/or a two-dimensional or three-dimensional ion trap.

Filtering by means of selection based upon mass-to-charge ratio (m/z) can be achieved by using a mass analyzer which can select ions based upon m/z , for example a quadrupole; or to transmit a wide m/z range, separate ions according to their m/z , and then select the ions of interest by means of their m/z value. An example of the latter would be a time-of-flight mass analyzer combined with a timed ion selector(s). The methods and/or systems provided may comprise isolating and/or separating the one or more proteins of interest, for example from two or more of a plurality of proteins, using a chromatographic

technique for example liquid chromatography (LC). The method may further comprise measuring an elution time for the protein of interest and/or comparing the measured elution time with an expected elution time.

5 Additionally or alternatively, the proteins of interest may be separated using an ion mobility technique, which may be carried out using an ion mobility cell. Additionally, the proteins of interest may be selected by order or time of ion mobility drift. The method may further comprise measuring a drift time for the proteins of interest and/or comparing the measured drift time with an expected drift time.

10 In some embodiments, the methods and/or systems provided are label-free, where quantitation can be achieved by comparison of the peak intensity, or area under the mass spectral peak for the precursor or product m/z values of interest between injections and across samples. In some embodiments, internal standard normalization may be used to account for any known associated analytical error. Another label-free method of quantification, spectral counting, involves summing the number of fragment ion spectra, or
15 scans, that are acquired for each given peptide, in a non-redundant or redundant fashion. The associated peptide mass spectra for each protein are then summed, providing a measure of the number of scans per protein with this being proportional to its abundance. Comparison can then be made between samples/injections.

In some embodiments, the ion source is selected from the group consisting of: (1)
20 an electrospray ionization (“ESI”) ion source; (2) an atmospheric pressure photo ionization (“APPI”) ion source; (3) an atmospheric pressure chemical ionization (“APCI”) ion source; (4) a matrix assisted laser desorption ionization (“MALDI”) ion source; (5) a laser desorption ionization (“LDI”) ion source; (6) an atmospheric pressure ionization (“API”) ion source; (7) a desorption ionization on silicon (“DIOS”) ion source; (8) an electron
25 impact (“EI”) ion source; (9) a chemical ionization (“CI”) ion source; (10) a field ionization (“FI”) ion source; (11) a field desorption (“FD”) ion source; (12) an inductively coupled plasma (“ICP”) ion source; (13) a fast atom bombardment (“FAB”) ion source; (14) a liquid secondary ion mass spectrometry (“LSIMS”) ion source; (15) a desorption electrospray ionization (“DESI”) ion source; (16) a nickel- 63 radioactive ion source; (17)
30 an atmospheric pressure matrix assisted laser desorption ionization ion source; and (18) a thermospray ion source.

In some embodiments, the methods and/or systems provided comprise an apparatus and/or control system configured to execute a computer program element comprising

computer readable program code means for causing a processor to execute a procedure to implement the methods.

In some embodiments, the methods and/or systems provided use an alternating low and elevated energy scan function in combination with liquid chromatography separation of a plant extract. A list of information for proteins of interest can be provided including, but is not limited to, m/z of precursor ion, m/z of product ions, retention time, ion mobility drift time and rate of change of mobility. During the course of the LC separation and as the target ions elute into the mass spectrometer (and as either low energy precursor ions, or elevated energy product ions are detected, or the retention time window is activated) the mass analyzer of the methods and/or systems provided may select a narrow m/z range (of a variable and changeable width) to pass ions through to the gas cell. Accordingly, the signal to noise ratio can be enhanced significantly for quantification of proteins of interest.

In some embodiments, at a chromatographic retention time when a targeted protein of interest is about to elute into the mass spectrometer ion source, the mass analyzer of the methods and/or systems provided can select a narrow m/z range (of a variable and changeable width) according to the targeted precursor ion. These selected ions are then transferred to an instrument stage capable of dissociating the ions by means of alternate and repeated switches between a high fragmentation mode where the sample precursor ions are substantially fragmented into product ions and a low fragmentation mode (or non-fragmentation mode) where the sample precursor ions are not substantially fragmented. Typically high resolution, accurate mass spectra are acquired in both modes and at the end of the experiment associated precursor and product ions are recognized by the closeness in fit of their chromatographic elution times and optionally other physicochemical properties. The signal intensity of either the precursor ion or the product ion associated with targeted proteins of interest can be used to determine the quantity of the proteins in the plant extract.

Those skilled in the art would understand certain variation can exist based on the disclosure provided. Thus, the following examples are given for the purpose of illustrating the invention and shall not be construed as being a limitation on the scope of the invention or claims.

30

EXAMPLES

Example 1

The methods and systems provided are used for determination of endogenous soybean allergen proteins in soybean seed including Gly m 1, Gly m 3, Gly m 4, Gly m 5 (beta-conglycinin), Gly m 6, Kunitz trypsin inhibitor 1, Kunitz trypsin inhibitor 3, Gly m Bd 28 K, Gly m Bd 30 K, and Gly m 8 (2S albumin).

Table 1. Preparation of signature peptide calibration standards				
Initial concentration (ng/mL)	Standard	Volume of Dilution Cocktail (μL)	Volume of Std. (μL)	Final concentration (ng/mL)
5880.00	Std 12	-	-	500.00
500.00	Std 11	200	200	250.00
250.00	Std 10	200	200	125.00
125.00	Std 9	200	200	62.50
62.50	Std 8	200	200	31.25
31.25	Std 7	200	200	15.63
15.63	Std 6	200	200	7.81
7.81	Std 5	200	200	3.91
3.91	Std 4	200	200	1.95
1.95	Std 3	200	200	0.98
0.98	Std 2	200	200	0.49
0.49	Std 1	2000	2000	0.24

A 100 ± 0.5 mg ground soybean seed sample is defatted twice with hexanes and dried before extracting with extraction buffer containing 5 M urea, 2 M thiourea, 50 mM Tris pH 8.0 and 65 mM DTT. The sample is sonicated in a water bath for thirty minutes, vortexed for one minute, sonicated for another thirty minutes and centrifuged at $> 3,000$ rpm for ten minutes at 4°C .

The aqueous supernatant is collected and diluted to bring the endogenous soybean allergen protein concentration into the calibration standard range with extraction buffer. The diluted extract is denatured at 95°C for twenty minutes with the additional 1 M Tris pH 8.0, 0.5 M DTT and deionized water followed by refrigeration at 4°C for ten minutes. The denatured extract is incubated overnight (~ 15 hours) at 37°C with 0.5 mg/mL trypsin enzyme. The digestion reaction is quenched with formic acid water (50/50 v/v) and

centrifuge at > 3,000 rpm for ten minutes at 4°C. An aliquot of digested extract is transferred to an autosampler vial and analyzed along with calibration standard by liquid chromatography with positive-ion electrospray (ESI) tandem mass spectrometry (LC-MS/MS). Calibration standards of signature peptides are prepared as listed in Table 1.

- 5 The limits of detection (LOD) and limits of quantitation (LOQ) for endogenous soybean allergens in this example are set forth in Table 2, where LOD and LOQ represent protein concentration (ng/mg).

Table 2. Limits of detection (LOD) and limits of quantitation (LOQ) for endogenous soybean allergens in Example 1 (LOD and LOQ represent protein concentration)			
Allergen	Signature peptide	LOD (ng/mg)	LOQ (ng/mg)
Gly m 1	SYPSNATCPR (SEQ ID NO: 1)	0.23	0.46
Gly m 3	YMVIQGEPGA VIR (SEQ ID NO: 2)	0.20	0.39
Gly m 5	NILEASYDTK (SEQ ID NO: 3)	1.22	2.44
Glycinin G2	VTAPAMR (SEQ ID NO: 4)	1.46	2.92
Glycinin G3	NNNPFSFLVPPK (SEQ ID NO: 5)	1.58	3.16
Glycinin precursor	ADFYNPK (SEQ ID NO: 6)	-	-
Kunitz trypsin inhibitor 1	GGGIEVDSTGK (SEQ ID NO: 7)	-	-
Kunitz trypsin inhibitor 3	GIGTLLSSPYR (SEQ ID NO: 8)	-	-
Gly m Bd 28 K	NKPQFLAGAASLLR (SEQ ID NO: 9)	5.70	11.40
Gly m Bd 30 K	GVITQVK (SEQ ID NO: 10)	1.15	2.30
Gly m 8	IMENQSEELEEK (SEQ ID NO: 11)	0.25	0.50

- 10 Concentrations of allergens are calculated from quantitation of signature peptides (for example Analyst Bioanalytical software for LC-MS/MS), and validated by other methods including enzyme-linked immunosorbent assays (ELISA). Calculated concentrations of allergens from different samples are compared using statistical analysis, and results show good consistency among samples.

15

Example 2

Several homologous protein sequences for Gly m Bd 28 K are identified from public databases including NCBI, Phytozome, and UniProt. Identified sequences (SEQ ID

NOs: 21 and 29-33) are analyzed using bioinformatics tools to identify sequence homology and shared sequence composition among the available protein sequences (see FIG. 2L). Specifically this involved the use of Vector NTI Align X alignment tool which performs a CLUSTAL W type alignment. From this analysis, a consensus sequence and/or
 5 representative sequence can be determined.

Once the consensus sequence and/or representative sequence is chosen or determined, it is digested *in silico* to generate candidate signature peptide fragments to be detected and measured by LC-MS. According to the unique approaches provided herein, signature peptides are selected based on the degree of conservation among the available
 10 protein sequences, such that the selected signature peptide can be used to quantify all or as many protein isoforms as possible among the identified protein sequences found in the public sequence databases. As a result, quantitation of selected signature peptides can not only measure Gly m Bd 28 K itself, but also measure potential allergens which are highly homologous to Gly m Bd 28 K.

15 Soybean seed samples are ground to a fine powder, defatted twice with hexane, and extracted with suitable assay buffer (for example 5 M urea, 2 M thiourea, 50 mM Tris (pH 8.0), 65 mM DTT). The samples are sonicated in buffer to extract proteins. The extracted proteins are diluted, denatured, and then proteolytically digested by adding trypsin protease and incubating at 37°C for 15-20 hours. The digestion reactions are acidified with formic
 20 acid (pH = 1-2) and are analyzed using LC-MS/MS.

The selected signature peptides can be used for both qualitative and quantitative analysis of Gly m Bd 28 K, either by itself or in combination with additional proteins in a multiplexing assay format. In this example, several signature peptides are selected from all peptide possibilities (SEQ ID NO: 9 NKPQFLAGAASLLR; SEQ ID NO: 34
 25 LGFIYDDELAER; SEQ ID NO: 35 TVVEEIFSK; SEQ ID NO: 36 MMQDQEEDDEEEK; SEQ ID NO: 37 NAYGWSK; SEQ ID NO: 38 ALHGGEYPPLSEPDIGVLLVK; SEQ ID NO: 39 QGDVFFVPR; SEQ ID NO: 40 YFPFCQVASR; SEQ ID NO: 41 TLMGPELSAAFGVSEDTLR; SEQ ID NO: 42 SFANDVVMDF), and representative quantitation of these signature peptides are shown in FIGs. 2A-2J. A peptide standard is
 30 synthesized for SEQ ID NO: 9 NKPQFLAGAASLLR for quantitative and qualitative analyses (see FIG. 2K). Synthetic peptides can directly serve as an analytical reference standard for protein quantitation.

Example 3

Several homologous protein sequences for Gly m Bd 30 K are identified from public databases including NCBI, Phytozome, and UniProt. Identified sequences (SEQ ID NOs: 22 and 43-44) are analyzed using bioinformatics tools to identify sequence homology and shared sequence composition among the available protein sequences (see FIG. 3J). Specifically, this involved the use of Vector NTI Align X alignment tool which performs a CLUSTAL W type alignment. From this analysis, a consensus sequence and/or representative sequence can be determined.

Once the consensus sequence and/or representative sequence is chosen or determined, it is digested *in silico* to generate candidate signature peptide fragments to be detected and measured by LC-MS. According to the unique approaches provided herein, signature peptides are selected based on the degree of conservation among the available protein sequences, such that the selected signature peptide can be used to quantify all or as many protein isoforms as possible among the identified protein sequences found in the public sequence databases. As a result, quantitation of selected signature peptides can not only measure Gly m Bd 30 K itself, but also measure potential allergens which are highly homologous to Gly m Bd 30 K.

Soybean seed samples are ground to a fine powder, defatted twice with hexane, and extracted with suitable assay buffer (for example 5 M urea, 2 M thiourea, 50 mM Tris (pH 8.0), 65 mM DTT). The samples are sonicated in buffer to extract proteins. The extracted proteins are diluted, denatured, and then proteolytically digested by adding trypsin protease and incubating at 37°C for 15-20 hours. The digestion reactions are acidified with formic acid (pH = 1-2) and are analyzed using LC-MS/MS.

The selected signature peptides can be used for both qualitative and quantitative analysis of Gly m Bd 30 K, either by itself or in combination with additional proteins in a multiplexing assay format. In this example, several signature peptides are selected from all peptide possibilities (SEQ ID NO: 10 GVITQVK; SEQ ID NO: 45 SILDLTLTK; SEQ ID NO: 46 FTTQK; SEQ ID NO: 47>NNLNYIR; SEQ ID NO: 48 FADITPQEFSK; SEQ ID NO: 49 EQYSCDHPPASWDWR; SEQ ID NO: 50 VTIDGYETLIMSDESTESQAFLESAILEQPISVSIDAK; and SEQ ID NO: 51 NTGNLLGVCGMNYFASYPTK), and representative quantitation of these signature peptides are shown in FIGs. 3A-3H. A peptide standard is synthesized for SEQ ID NO: 10

GVITQVK for quantitative and qualitative analyses (see FIG. 3I). Synthetic peptides can directly serve as an analytical reference standard for protein quantitation.

Example 4

5 Several homologous protein sequences for Kunitz trypsin inhibitor 1 are identified from public databases including NCBI, Phytozome, and UniProt. Identified sequences (SEQ ID NOs: 23 and 53-54) are analyzed using bioinformatics tools to identify sequence homology and shared sequence composition among the available protein sequences (see FIG. 4I). Specifically this involved the use of Vector NTI Align X alignment tool which
10 performs a CLUSTAL W type alignment. From this analysis, a consensus sequence and/or representative sequence can be determined.

Once the consensus sequence and/or representative sequence is chosen or determined, it is digested *in silico* to generate candidate signature peptide fragments to be detected and measured by LC-MS. According to the unique approaches provided herein,
15 signature peptides are selected based on the degree of conservation among the available protein sequences, such that the selected signature peptide can be used to quantify all or as many protein isoforms as possible among the identified protein sequences found in the public sequence databases. As a result, quantitation of selected signature peptides can not only measure Kunitz trypsin inhibitor 1 itself, but also measure potential allergens which
20 are highly homologous to Kunitz trypsin inhibitor 1.

Soybean seed samples are ground to a fine powder, defatted twice with hexane, and extracted with suitable assay buffer (for example 5 M urea, 2 M thiourea, 50 mM Tris (pH 8.0), 65 mM DTT). The samples are sonicated in buffer to extract proteins. The extracted proteins are diluted, denatured, and then proteolytically digested by adding trypsin protease
25 and incubating at 37°C for 15-20 hours. The digestion reactions are acidified with formic acid (pH = 1-2) and are analyzed using LC-MS/MS.

The selected signature peptides can be used for both qualitative and quantitative analysis of Kunitz trypsin inhibitor 1, either by itself or in combination with additional proteins in a multiplexing assay format. In this example, several signature peptides are
30 selected from all peptide possibilities (SEQ ID NO: 7 GGGIEVDSTGK; SEQ ID NO: 55 EICPLTVVQSPNELDK; SEQ ID NO: 56 EGLQAVK; SEQ ID NO: 57 LVFCPQQAEDNK; SEQ ID NO: 58 CEDIGIQIDDDGIR; SEQ ID NO: 59 LVLSK; and SEQ ID NO: 60 NKPLVVQFQK), and representative quantitation of these signature

peptides are shown in FIGs. 4A- 4G. A peptide standard is synthesized for SEQ ID NO: 7 GGGIEVDSTGK for quantitative and qualitative analyses (see FIG. 4H). Synthetic peptides can directly serve as an analytical reference standard for protein quantitation.

5

Example 5

Several homologous protein sequences for Kunitz trypsin inhibitor 3 are identified from public databases including NCBI, Phytozome, and UniProt. Identified sequences (SEQ ID NOs: 24 and 62-63) are analyzed using bioinformatics tools to identify sequence homology and shared sequence composition among the available protein sequences (see
10 FIG. 5K). Specifically, this involved the use of Vector NTI Align X alignment tool which performs a CLUSTAL W type alignment. From this analysis, a consensus sequence and/or representative sequence can be determined.

Once the consensus sequence and/or representative sequence is chosen or determined, it is digested *in silico* to generate candidate signature peptide fragments to be
15 detected and measured by LC-MS. According to the unique approaches provided herein, signature peptides are selected based on the degree of conservation among the available protein sequences, such that the selected signature peptide can be used to quantify all or as many protein isoforms as possible among the identified protein sequences found in the public sequence databases. As a result, quantitation of selected signature peptides can not
20 only measure Kunitz trypsin inhibitor 3 itself, but also measure potential allergens which are highly homologous to Kunitz trypsin inhibitor 3.

Soybean seed samples are ground to a fine powder, defatted twice with hexane, and extracted with suitable assay buffer (for example 5 M urea, 2 M thiourea, 50 mM Tris (pH 8.0), 65 mM DTT). The samples are sonicated in buffer to extract proteins. The extracted
25 proteins are diluted, denatured, and then proteolytically digested by adding trypsin protease and incubating at 37°C for 15-20 hours. The digestion reactions are acidified with formic acid (pH = 1-2) and are analyzed using LC-MS/MS.

The selected signature peptides can be used for both qualitative and quantitative analysis of Kunitz trypsin inhibitor 3, either by itself or in combination with additional
30 proteins in a multiplexing assay format. In this example, several signature peptides are selected from all peptide possibilities (SEQ ID NO: 8 GIGTISSPYR; SEQ ID NO: 64 CPLTVVQSR; SEQ ID NO: 65 NELDK; SEQ ID NO: 66 IGENK; SEQ ID NO: 67 DAMDGWFR; SEQ ID 68 LVFCPQQAEDDK; SEQ ID NO: 69 CGDIGISIDHDDGTR;

SEQ ID NO: 70 LVVSK; and SEQ ID NO: 71 NKPLVVQFQK), and representative quantitation of these signature peptides are shown in FIGs. 5A- 5I. A peptide standard is synthesized for SEQ ID NO: 8 GIGTISSPYR for quantitative and qualitative analyses (see FIG. 5J). Synthetic peptides can directly serve as an analytical reference standard for
 5 protein quantitation.

Example 6

Several homologous protein sequences for Gly m 8 (2S albumin) are identified from public databases including NCBI, Phytozome, and UniProt. Identified sequences
 10 (SEQ ID NOs: 25 and 72-74) are analyzed using bioinformatics tools to identify sequence homology and shared sequence composition among the available protein sequences (see FIG. 6J). Specifically this involved the use of Vector NTI Align X alignment tool which performs a CLUSTAL W type alignment. From this analysis, a consensus sequence and/or representative sequence can be determined.

15 Once the consensus sequence and/or representative sequence is chosen or determined, it is digested *in silico* to generate candidate signature peptide fragments to be detected and measured by LC-MS. According to the unique approaches provided herein, signature peptides are selected based on the degree of conservation among the available protein sequences, such that the selected signature peptide can be used to quantify all or as
 20 many protein isoforms as possible among the identified protein sequences found in the public sequence databases. As a result, quantitation of selected signature peptides can not only measure Gly m 8 (2S albumin) itself, but also measure potential allergens which are highly homologous to Gly m 8 (2S albumin).

Soybean seed samples are ground to a fine powder, defatted twice with hexane, and
 25 extracted with suitable assay buffer (for example 5 M urea, 2 M thiourea, 50 mM Tris (pH 8.0), 65 mM DTT). The samples are sonicated in buffer to extract proteins. The extracted proteins are diluted, denatured, and then proteolytically digested by adding trypsin protease and incubating at 37°C for 15-20 hours. The digestion reactions are acidified with formic acid (pH = 1-2) and are analyzed using LC-MS/MS.

30 The selected signature peptides can be used for both qualitative and quantitative analysis of Gly m 8 (2S albumin), either by itself or in combination with additional proteins in a multiplexing assay format. In this example, several signature peptides are selected from all peptide possibilities (SEQ ID NO: 11 IMENQSEELEEK; SEQ ID NO: 75

WQHQQDSCR; SEQ ID NO: 76 QLQGVNLTPEK; SEQ ID NO: 77 HIMEK; SEQ ID NO: 78 DEDEEEEGHMQK; SEQ ID NO: 79 CCTEMSELR; SEQ ID NO: 80 ELINLATMCR; and SEQ ID NO: 81 FGPMIQCDLSSDD), and representative quantitation of these signature peptides are shown in FIGs. 6A-6H. A peptide standard is
 5 synthesized for SEQ ID NO: 11 IMENQSEEELEEK for quantitative and qualitative analyses (see FIG. 6I). Synthetic peptides can directly serve as an analytical reference standard for protein quantitation.

Example 7

10 Several homologous protein sequences for Lectin are identified from public databases including NCBI, Phytozome, and UniProt. Identified sequences (SEQ ID NOs: 26 and 83-90) are analyzed using bioinformatics tools to identify sequence homology and shared sequence composition among the available protein sequences (see FIG. 7F). Specifically this involved the use of Vector NTI Align X alignment tool which performs a
 15 CLUSTAL W type alignment. From this analysis, a consensus sequence and/or representative sequence can be determined.

Once the consensus sequence and/or representative sequence is chosen or determined, it is digested *in silico* to generate candidate signature peptide fragments to be detected and measured by LC-MS. According to the unique approaches provided herein,
 20 signature peptides are selected based on the degree of conservation among the available protein sequences, such that the selected signature peptide can be used to quantify all or as many protein isoforms as possible among the identified protein sequences found in the public sequence databases. As a result, quantitation of selected signature peptides can not only measure lectin itself, but also measure potential allergens which are highly
 25 homologous to Lectin.

Soybean seed samples are ground to a fine powder, defatted twice with hexane, and extracted with suitable assay buffer (for example 5 M urea, 2 M thiourea, 50 mM Tris (pH 8.0), 65 mM DTT). The samples are sonicated in buffer to extract proteins. The extracted proteins are diluted, denatured, and then proteolytically digested by adding trypsin protease
 30 and incubating at 37°C for 15-20 hours. The digestion reactions are acidified with formic acid (pH = 1-2) and are analyzed using LC-MS/MS.

The selected signature peptides can be used for both qualitative and quantitative analysis of lectin, either by itself or in combination with additional proteins in a

multiplexing assay format. In this example, several signature peptides are selected from all peptide possibilities (SEQ ID NO: 91 VFSPNK; SEQ ID NO: 92 ANSTNTVSFTVSK; SEQ ID NO: 93 QQNLIFQGDAAISPSGVLR; and SEQ ID NO: 94 TADGLAFFLAPVGSKPQSK), and representative quantitation of these signature peptides are shown in FIGs. 7A-7D. A peptide standard is synthesized for SEQ ID NO: 91 VFSPNK for quantitative and qualitative analyses (see FIG. 7E). Synthetic peptides can directly serve as an analytical reference standard for protein quantitation.

Example 8

Several homologous protein sequences for lipoxygenase are identified from public databases including NCBI, Phytozome, and UniProt. Identified sequences (SEQ ID NOs: 27 and 96-104) are analyzed using bioinformatics tools to identify sequence homology and shared sequence composition among the available protein sequences (see FIG. 8R). Specifically this involved the use of Vector NTI Align X alignment tool which performs a CLUSTAL W type alignment. From this analysis, a consensus sequence and/or representative sequence can be determined.

Once the consensus sequence and/or representative sequence is chosen or determined, it is digested *in silico* to generate candidate signature peptide fragments to be detected and measured by LC-MS. According to the unique approaches provided herein, signature peptides are selected based on the degree of conservation among the available protein sequences, such that the selected signature peptide can be used to quantify all or as many protein isoforms as possible among the identified protein sequences found in the public sequence databases. As a result, quantitation of selected signature peptides can not only measure lipoxygenase itself, but also measure potential allergens which are highly homologous to lipoxygenase.

Soybean seed samples are ground to a fine powder, defatted twice with hexane, and extracted with suitable assay buffer (for example 5 M urea, 2 M thiourea, 50 mM Tris (pH 8.0), 65 mM DTT). The samples are sonicated in buffer to extract proteins. The extracted proteins are diluted, denatured, and then proteolytically digested by adding trypsin protease and incubating at 37°C for 15-20 hours. The digestion reactions are acidified with formic acid (pH = 1-2) and are analyzed using LC-MS/MS.

The selected signature peptides can be used for both qualitative and quantitative analysis of lipoxygenase, either by itself or in combination with additional proteins in a

multiplexing assay format. In this example, several signature peptides are selected from all peptide possibilities (SEQ ID NO: 105 SSDFLTYGIK; SEQ ID NO: 106 GTVVLMPK; SEQ ID NO: 107 NVLDFNAITSIGK; SEQ ID NO: 108 GGVIDTATGILGQGVSLVGGVIDTATSFLGR; SEQ ID NO: 109 IFFVNDTYLPSATPAPLLK; SEQ ID NO: 110 DENFGHLK; SEQ ID NO: 111 SLSHDVIPLFK; SEQ ID NO: 112 SLYEGGIK; SEQ ID NO: 113 TDGENVLQFPPPHVAK; SEQ ID NO: 114 INSLPTAK; SEQ ID NO: 115 TILFLK; SEQ ID NO: 116 HLSVLHPIYK; SEQ ID NO: 117 QSLINADGIIK; SEQ ID NO: 118 FIPAEGTPEYDEMVK; SEQ ID NO: 119 ALEAFK; and SEQ ID NO: 120 GIPNSISI), and representative quantitation of these signature peptides are shown in FIGs. 8A-8P. A peptide standard is synthesized for SEQ ID NO: 105 SSDFLTYGIK for quantitative and qualitative analyses (see FIG. 8Q). Synthetic peptides can directly serve as an analytical reference standard for protein quantitation.

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 29

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 29

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

CLAIMS

We claim:

- 5 1. A method of selecting candidate signature peptide for quantitation of known allergen and potential allergens from a plant-based sample, comprising:
- (a) identifying potential allergens based on homology to at least one known allergen protein sequence;
- (b) performing sequence alignment of the at least one known allergen and
10 potential allergens identified in step (a);
- (c) selecting a consensus sequence or representative sequence based on the sequence alignment;
- (d) determining a plural of candidate signature peptides based on conservative regions or domains from the sequence alignment and *in silico* digestion data
15 of the consensus sequence or representative sequence selected in Step (c); and
- (e) quantitating the amount of the at least one known allergen and potential allergens in the plant-based sample based on measurements of the signature peptides.
- 20 2. The method of claim 1, wherein the quantitating step uses a column chromatography and mass spectrometry.
3. The method of claim 1, wherein the quantitating step comprises measuring
25 the plural of candidate signature peptides using high resolution accurate mass spectrometry (HRAM MS).
4. The method of claim 1, wherein the quantitating step comprises calculating corresponding peak heights or peak areas of the candidate signature peptides from mass
30 spectrometry.
5. The method of claim 1, wherein the quantitating step comprises comparing data from high fragmentation mode and low fragmentation mode from mass spectrometry.

6. The method of claim 1, wherein the at least one known allergen comprises Gly m Bd 28 K, Gly m Bd 30 K, Kunitz trypsin inhibitor 1, Kunitz trypsin inhibitor 3, Gly m 8 (2S albumin), Lectin, or lipoxygenase.

5

7. The method of claim 1, wherein the potential allergens comprise at least one sequence selected from the group consisting of:

- (a) SEQ ID NOs: 21 and 29-33 for Gly m Bd 28 K;
- (b) SEQ ID NOs: 22 and 43-44 for Gly m Bd 30 K;
- 10 (c) SEQ ID NOs: 23 and 53-54 for Kunitz trypsin inhibitor 1;
- (d) SEQ ID NOs: 24 and 62-63 for Kunitz trypsin inhibitor 3;
- (e) SEQ ID NOs: 25 and 72-74 for Gly m 8 (2S albumin);
- (f) SEQ ID NOs: 26 and 83-90 for Lectin; and
- (g) SEQ ID NOs: 27 and 96-104 for lipoxygenase.

15

8. The method of claim 1, wherein the candidate signature peptides comprise at least one sequence selected from the group consisting of:

- (a) SEQ ID NOs: 9 and 34-42 for Gly m Bd 28 K;
- (b) SEQ ID NOs: 10 and 45-51 for Gly m Bd 30 K;
- 20 (c) SEQ ID NOs: 7 and 55-60 for Kunitz trypsin inhibitor 1;
- (d) SEQ ID NOs: 8 and 64-71 for Kunitz trypsin inhibitor 3;
- (e) SEQ ID NOs: 11 and 75-81 for Gly m 8 (2S albumin);
- (f) SEQ ID NOs: 91-94 for Lectin; and
- (g) SEQ ID NOs: 105-120 for lipoxygenase.

25

9. The method of claim 1, wherein the plant-based sample comprises a soybean seed or part of a soybean seed.

10. A system for quantitating one or more protein of interest with known amino acid sequence in a plant-based sample, the system comprising:
- (a) a high-throughput means for extracting proteins from a plant-based sample;
 - (b) a process module for digesting extracted proteins with at least one protease;
 - 5 (c) a separation module for separating peptides in a single step;
 - (d) a selection module for selecting a plural of signature peptides for at least one known allergen and potential allergens; and
 - (e) a mass spectrometry for measuring the plural of signature peptides.
- 10 11. The system of claim 10, wherein the separation module comprises a column chromatography.
12. The system of claim 11, wherein the column chromatography comprises a liquid column chromatography.
- 15 13. The system of claim 10, wherein the mass spectrometry comprises a high resolution accurate mass spectrometry (HRAM MS).
14. The system of claim 10, wherein the selection module uses a method
20 according to claim 1.
15. The system of claim 10, wherein the at least one known allergen comprises Gly m Bd 28 K, Gly m Bd 30 K, Kunitz trypsin inhibitor 1, Kunitz trypsin inhibitor 3, Gly m 8 (2S albumin), Lectin, or lipoxygenase.
- 25 16. The system of claim 10, wherein the potential allergens comprise at least one sequence selected from the group consisting of:
- (a) SEQ ID NOs: 21 and 29-33 for Gly m Bd 28 K;
 - (b) SEQ ID NOs: 22 and 43-44 for Gly m Bd 30 K;
 - 30 (c) SEQ ID NOs: 23 and 53-54 for Kunitz trypsin inhibitor 1;
 - (d) SEQ ID NOs: 24 and 62-63 for Kunitz trypsin inhibitor 3;

- (e) SEQ ID NOs: 25 and 72-74 for Gly m 8 (2S albumin);
- (f) SEQ ID NOs: 26 and 83-90 for Lectin; and
- (g) SEQ ID NOs: 27 and 96-104 for lipoxygenase.

5 17. The system of claim 10, wherein the signature peptides comprise at least one sequence selected from the group consisting of:

- (a) SEQ ID NOs: 9 and 34-42 for Gly m Bd 28 K;
- (b) SEQ ID NOs: 10 and 45-51 for Gly m Bd 30 K;
- (c) SEQ ID NOs: 7 and 55-60 for Kunitz trypsin inhibitor 1;
- 10 (d) SEQ ID NOs: 8 and 64-71 for Kunitz trypsin inhibitor 3;
- (e) SEQ ID NOs: 11 and 75-81 for Gly m 8 (2S albumin);
- (f) SEQ ID NOs: 91-94 for Lectin; and
- (e) SEQ ID NOs: 105-120 for lipoxygenase.

15 18. The system of claim 10, wherein the plant-based sample comprises a soybean seed or part of a soybean seed.

 19. A high-throughput method of quantitating at least one allergen with known amino acid sequence and homologous potential allergens in a plant-based sample,
20 comprising using the system of claim 10.

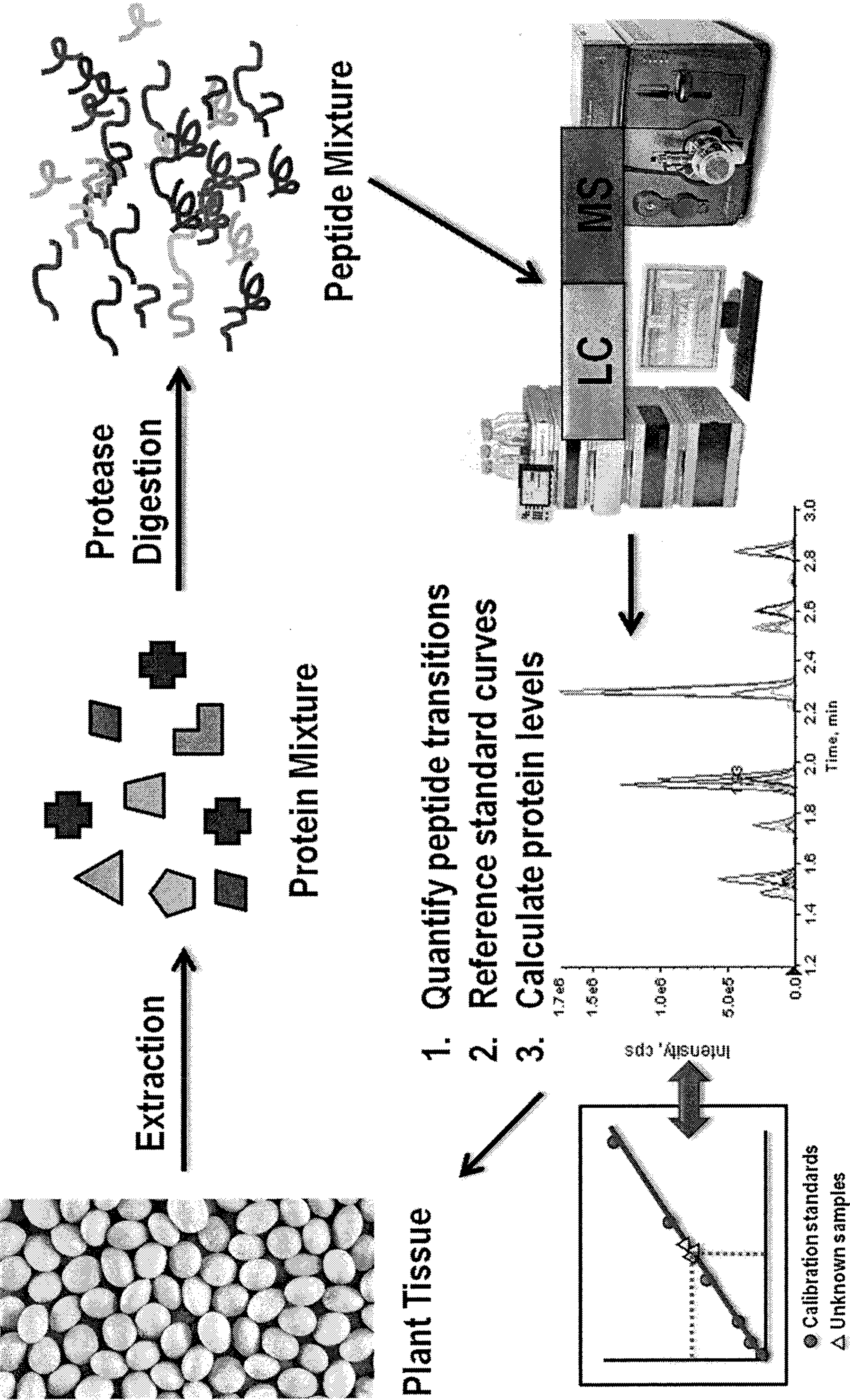
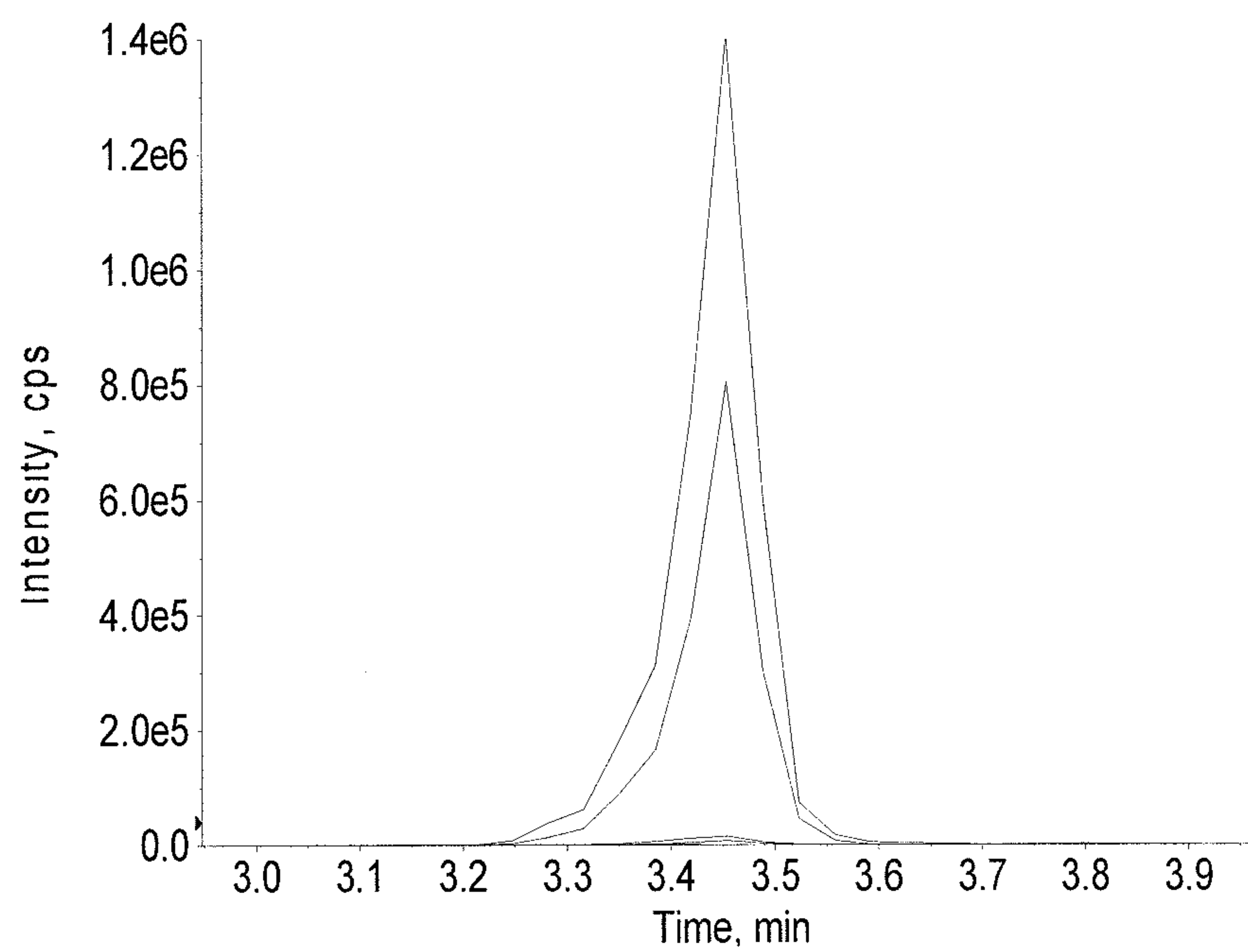
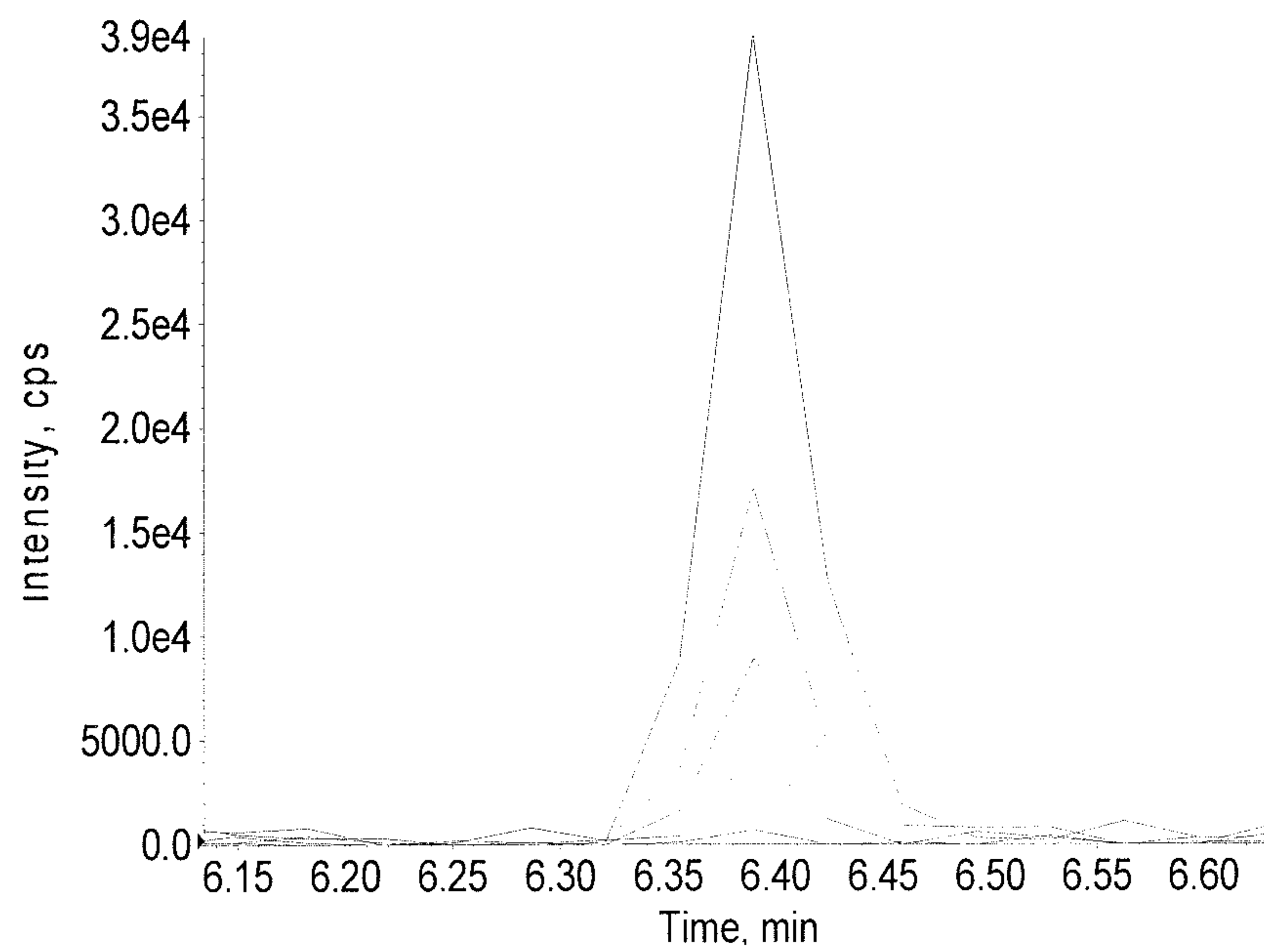


FIG. 1

2/52



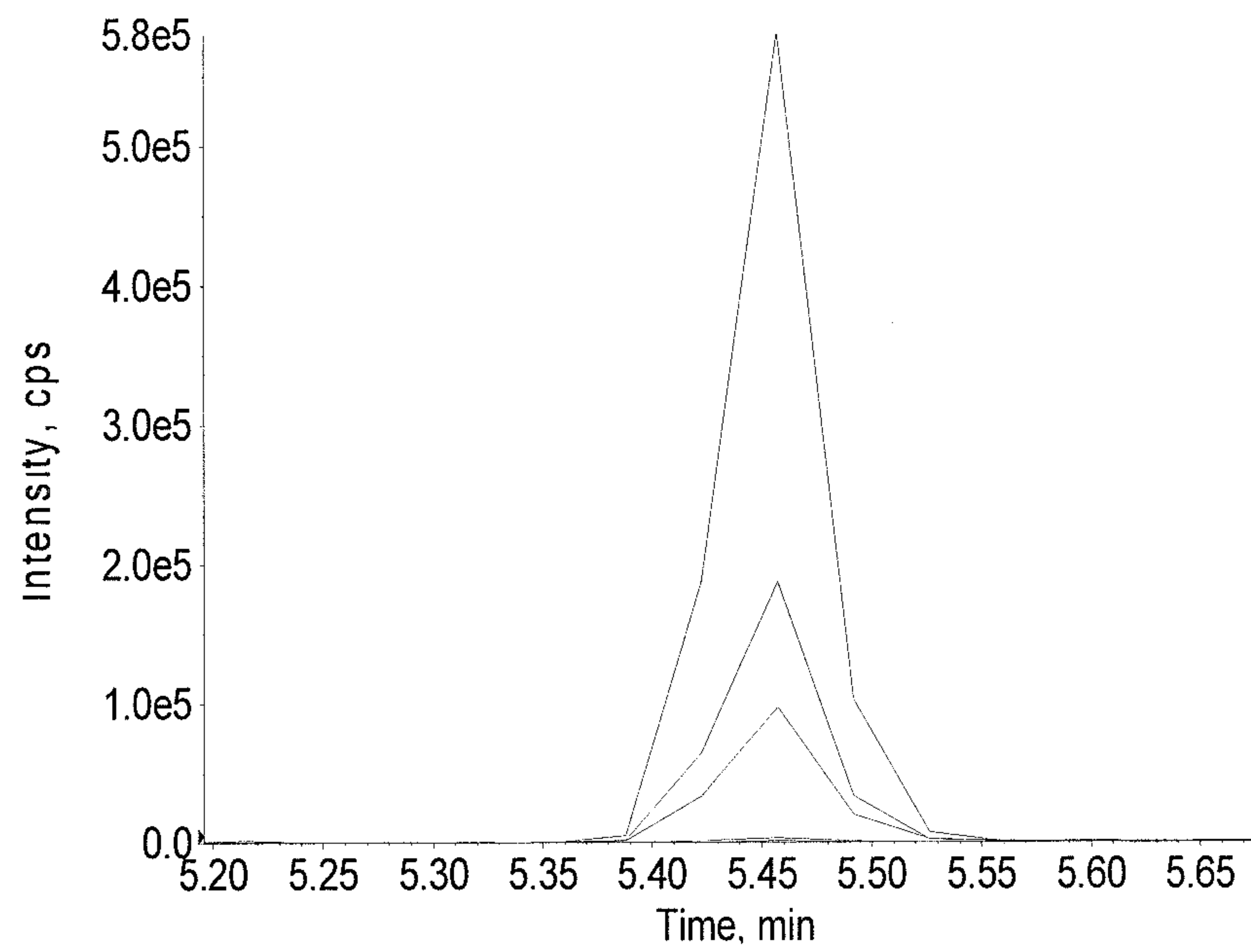
NKPQFLAGAASLLR

FIG. 2A

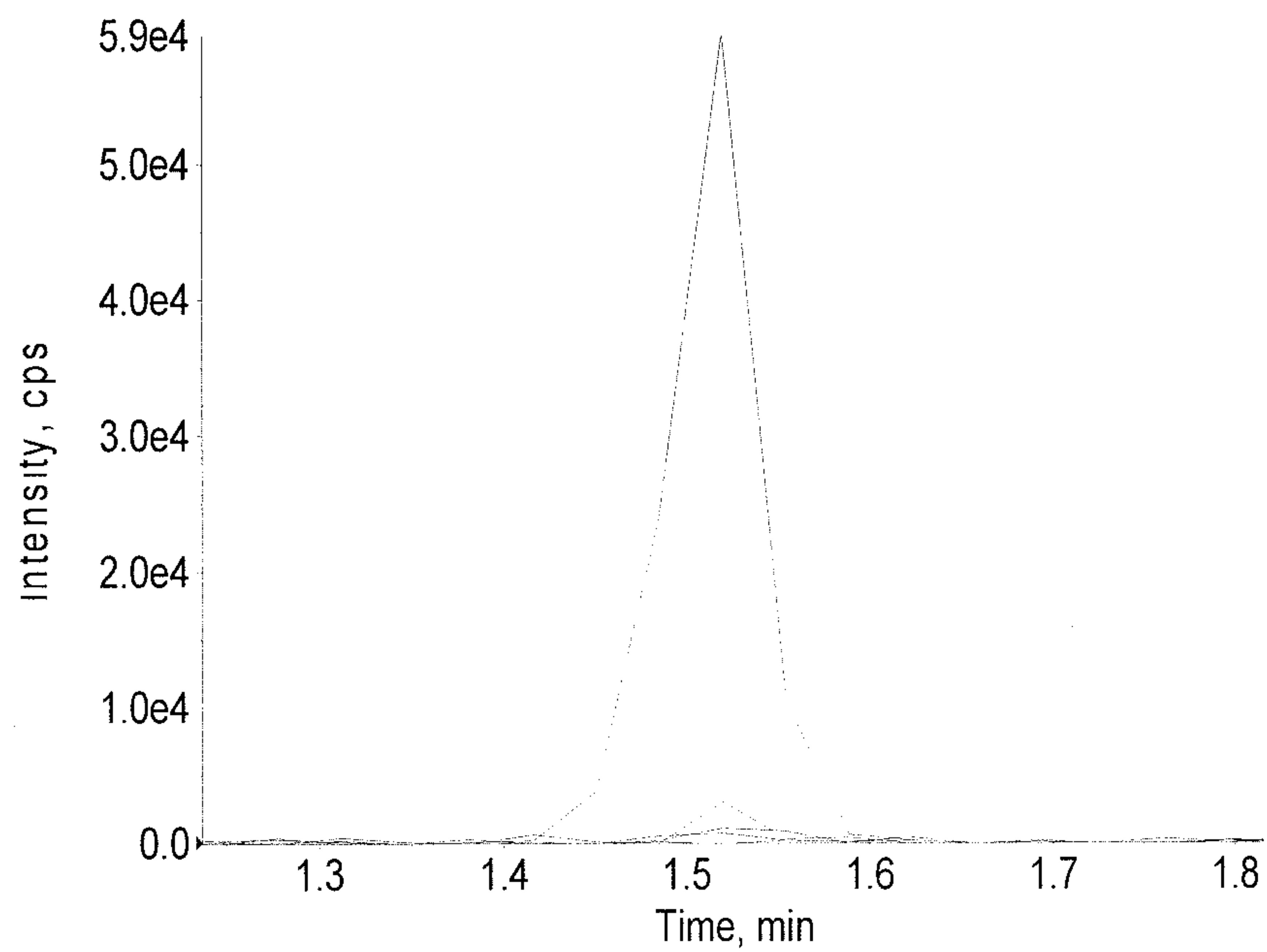
LGFIYDDELAER

FIG. 2B

3/52



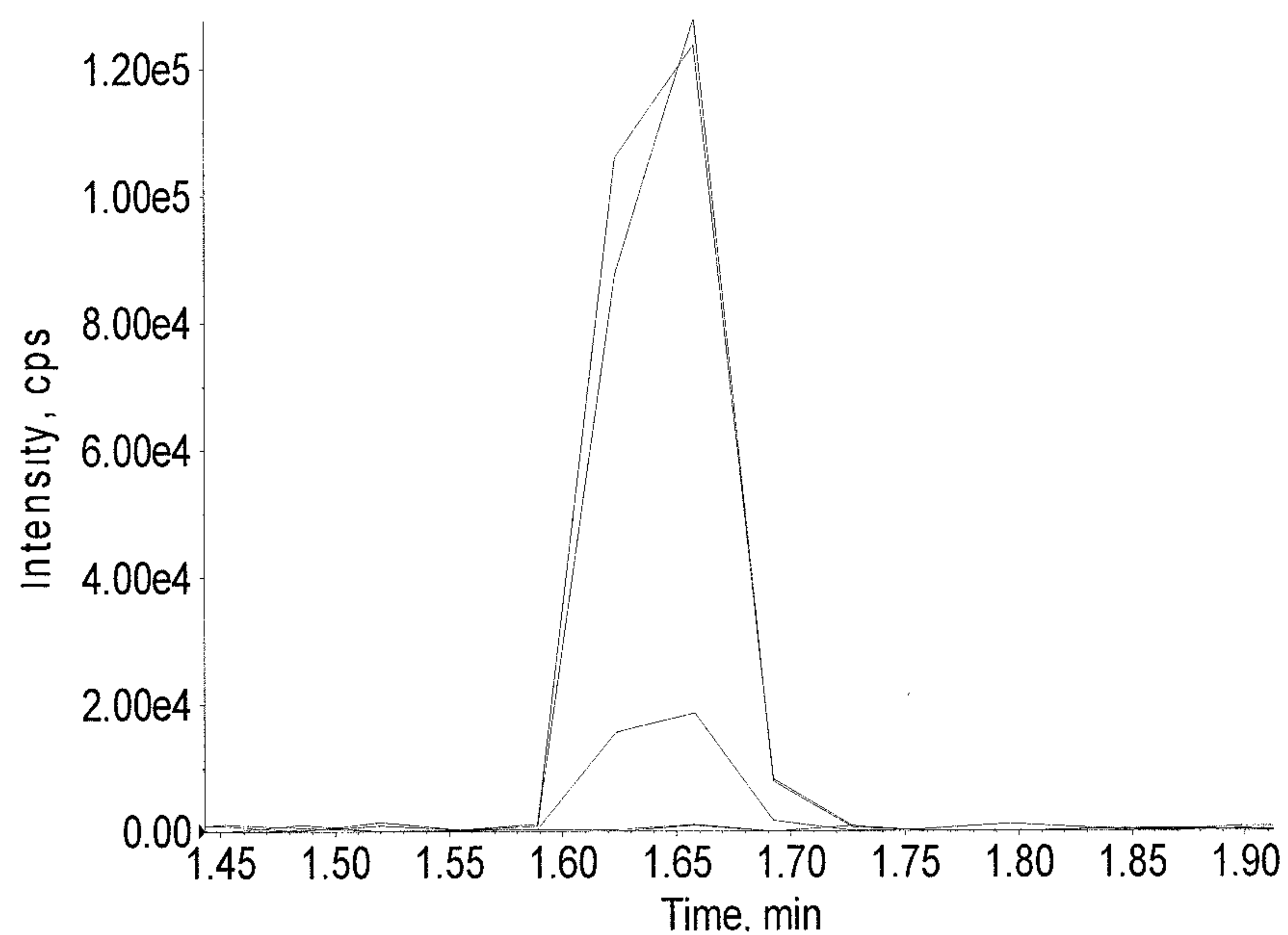
TVVEEIFSK

FIG. 2C

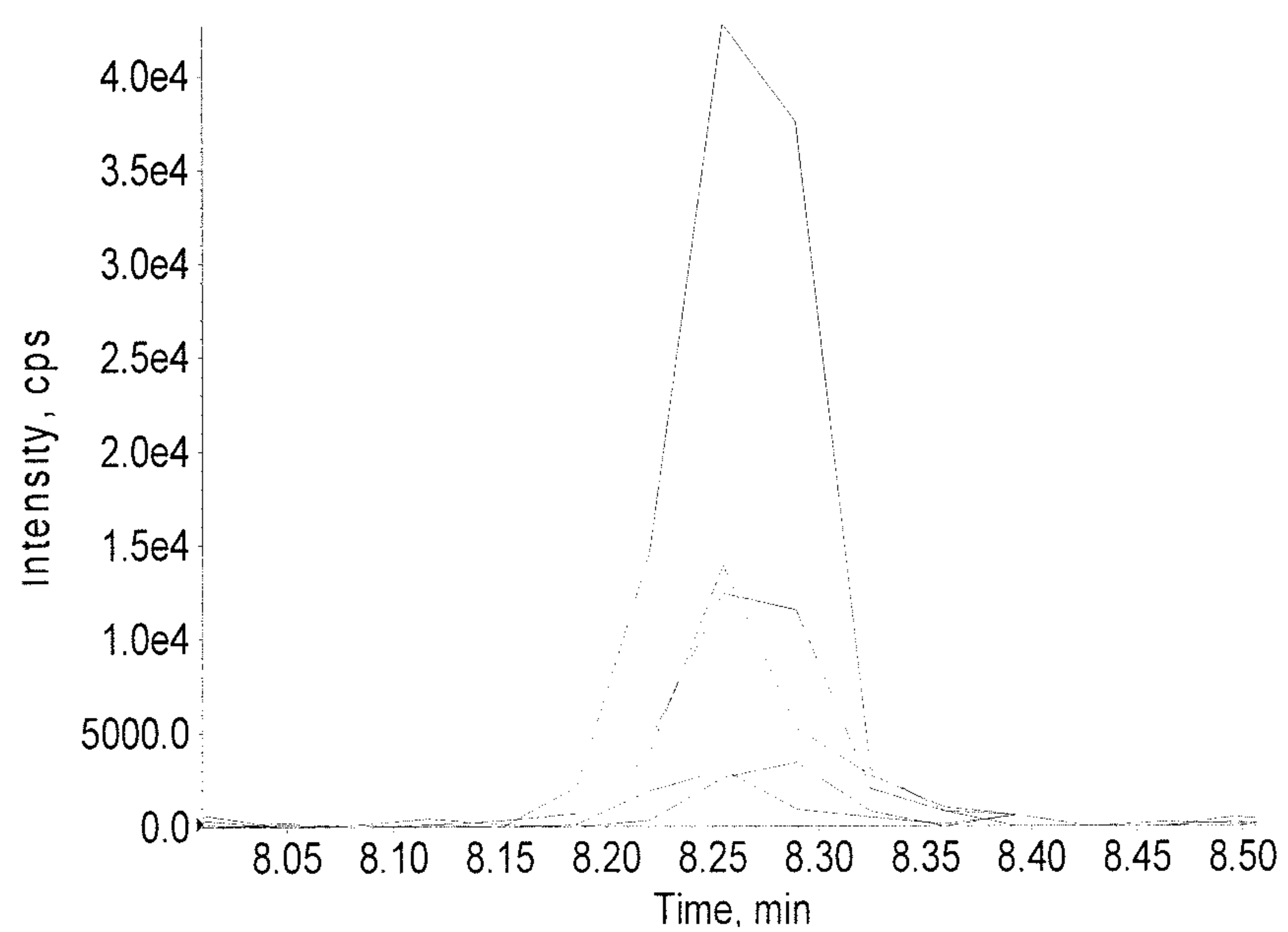
MMQDQEEDEEEK

FIG. 2D

4/52



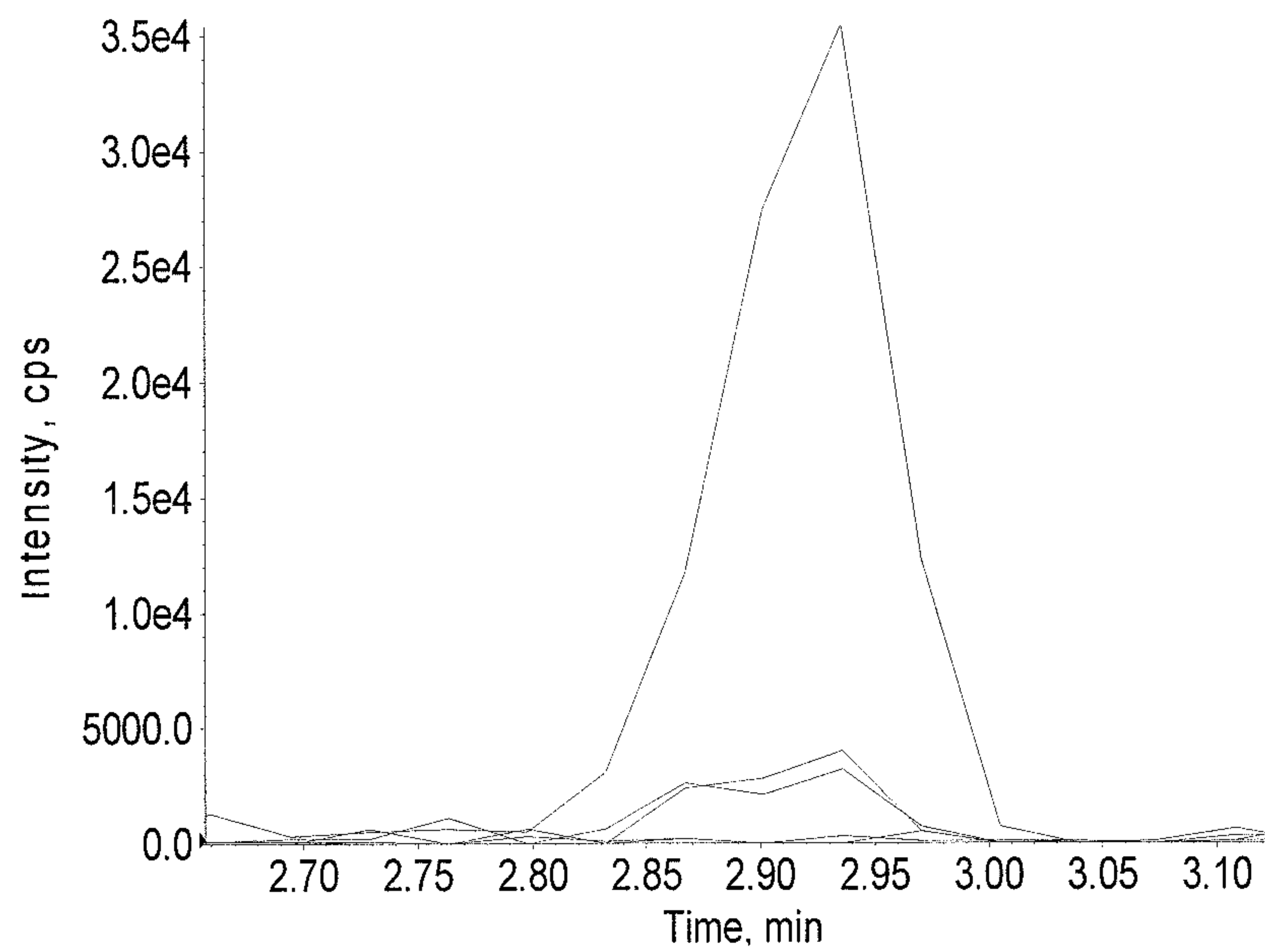
NAYGWSK

FIG. 2E

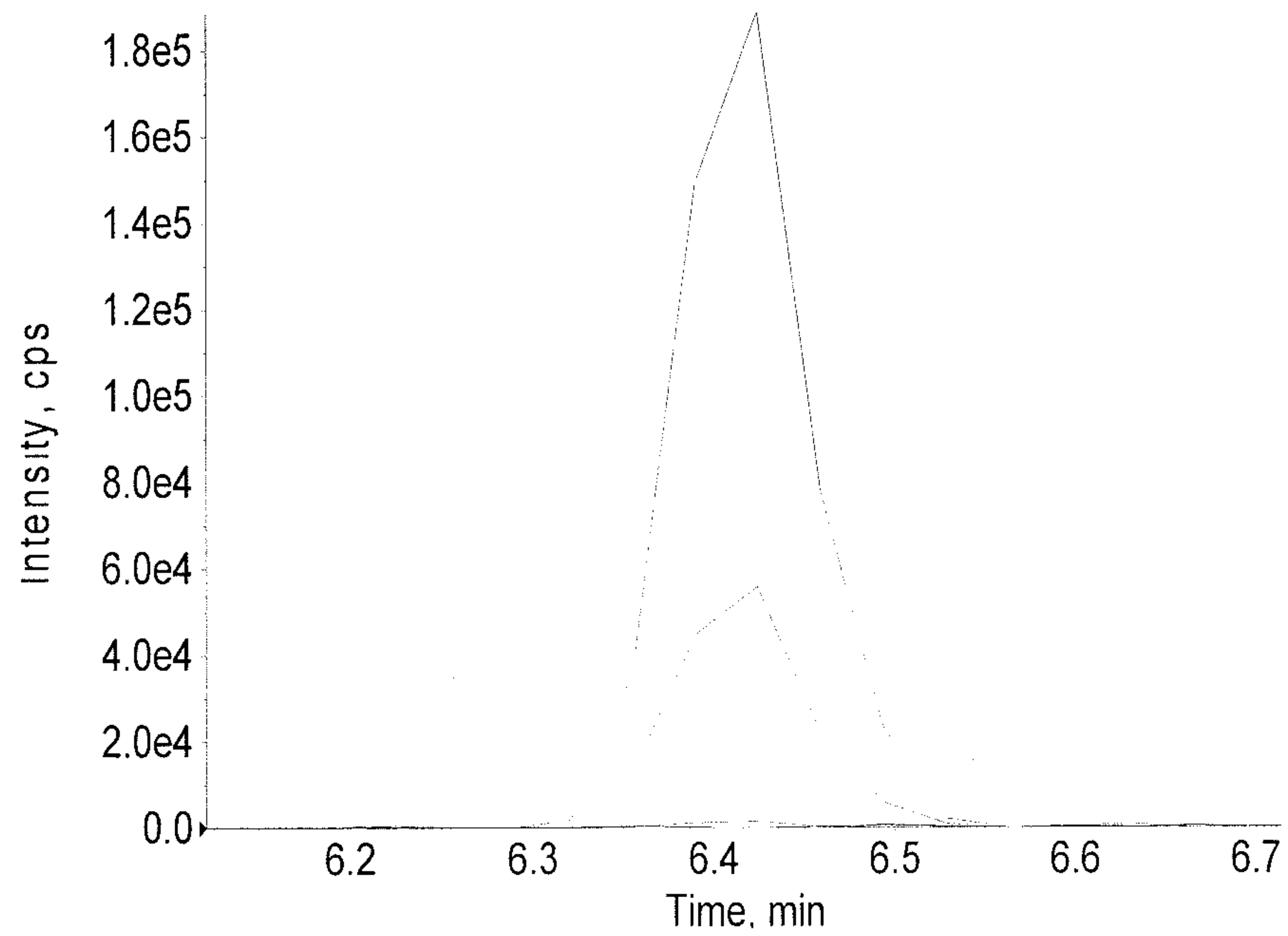
ALHGGEYPPLSEPDIGVLLVK

FIG. 2F

5/52



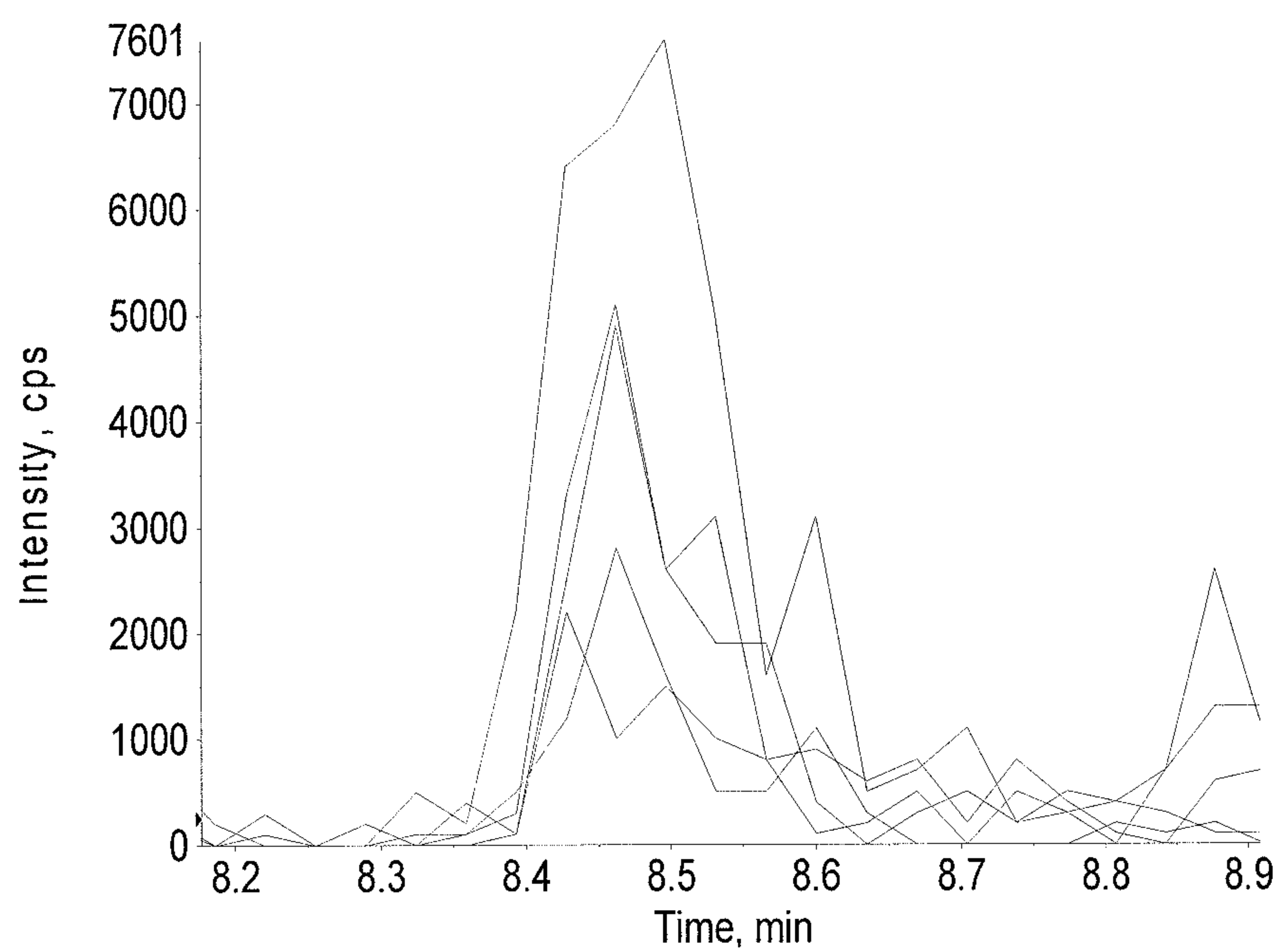
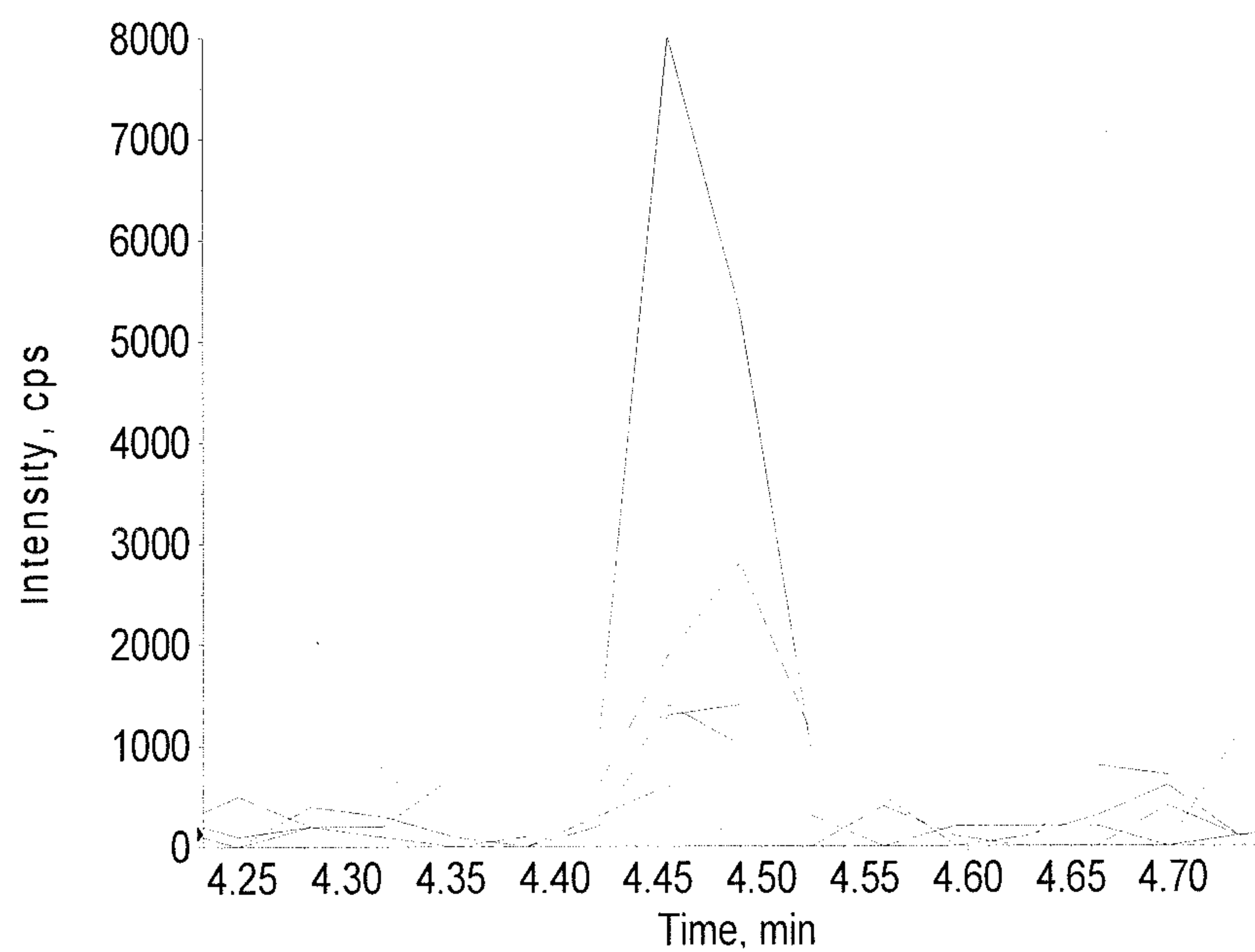
QGDVFVPR

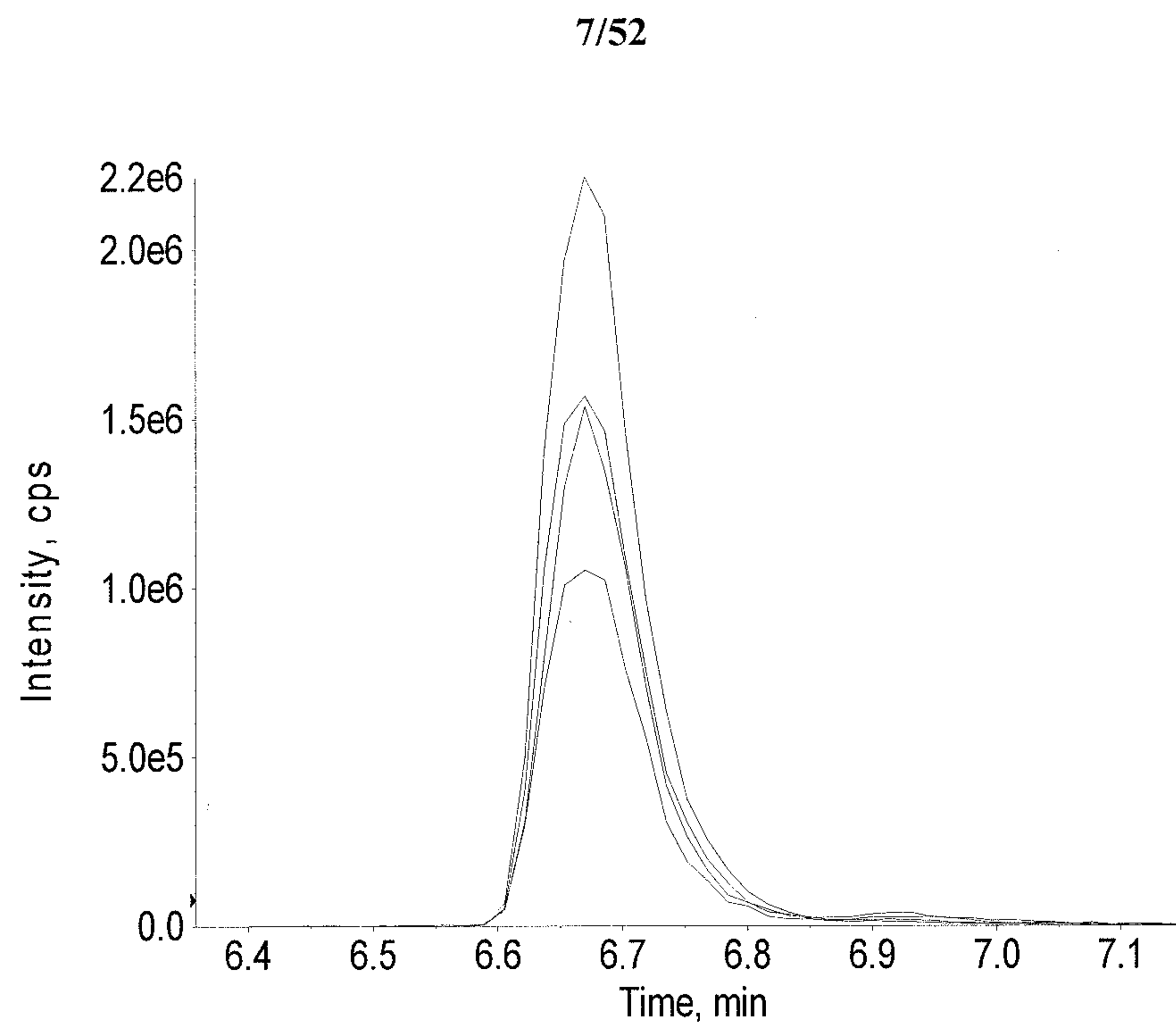
FIG. 2G

YFPFCQVASR

FIG. 2H

6/52

**FIG. 2I****FIG. 2J**



NKPQFLAGAASLLR

FIG. 2K

8/52

	1		50
28K Eric BAB21619	(1)	---KTTLTLLLLFVLCHGVATTTMAFHDDDEGGDKKSPKSLFLMSNSTRVFK	
28K ILSI 12697782	(1)	MGNKTTLTLLLLFVLCHGVATTTMAFHDDDEGGDKKSPKSLFLMSNSTRVFK	
28K Ping ACD36978.1	(1)	-----VLCHGVATTTMAFHDDDEGGDKKSPKSLFLMSNSTRVFK	
28K Ping ACD36975.1	(1)	-----	
28K Ping ACD36976.1	(1)	-----	
28K Ping ACD36974.1	(1)	-----	
Consensus	(1)	VLCHGVATTTMAFHDDDEGGDKKSPKSLFLMSNSTRVFK	
	51		100
28K Eric BAB21619	(48)	TDAGEMRVLKSHGGRIFYRHHIGFISMEPKSLFVPQYLDNLIIIFIRRG	
28K ILSI 12697782	(51)	TDAGEMRVLKSHGGRIFYRHHIGFISMEPKSLFVPQYLDNLIIIFIRRG	
28K Ping ACD36978.1	(39)	TDAGEMRVLKSHGGRIFYRHHIGFISMEPKSLFVPQYLDNLIIIFIRRG	
28K Ping ACD36975.1	(1)	-----LDSNLIIFIRRG	
28K Ping ACD36976.1	(1)	-----LDSNLIIFIRRG	
28K Ping ACD36974.1	(1)	-----LDSNLIIFIRRG	
Consensus	(51)	TDAGEMRVLKSHGGRIFYRHHIGFISMEPKSLFVPQYLDNLIIIFIRRG	
	101		150
28K Eric BAB21619	(98)	EAKLGFIYDDELAERRLKTGDLYMIPSGSAFYLVNIGEGQRLHVICSIDP	
28K ILSI 12697782	(101)	EAKLGFIYDDELAERRLKTGDLYMIPSGSAFYLVNIGEGQRLHVICSIDP	
28K Ping ACD36978.1	(89)	EAKLGFIYDDELAERRLKTGDLYMIPSGSAFYLVNIGEGQRLHVICSIDP	
28K Ping ACD36975.1	(13)	EAKLGFIYDDELAERRLKTGDLYMIPSGSAFYLVNIGEGQRLHVICSIDP	
28K Ping ACD36976.1	(13)	EAKLGFIYDDELAERRLKTGDLYMIPSGSAFYLVNIGEGQRLHVICSIDP	
28K Ping ACD36974.1	(13)	EAKLGFIYDDELAERRLKTGDLYMIPSGSAFYLVNIGEGQRLHVICSIDP	
Consensus	(101)	EAKLGFIYDDELAERRLKTGDLYMIPSGSAFYLVNIGEGQRLHVICSIDP	
	151		200
28K Eric BAB21619	(148)	STSLGLETQSFYIGGGANSHSVLSGFEPAILLETAFNESRTVVVEEIFSKE	
28K ILSI 12697782	(151)	STSLGLETQSFYIGGGANSHSVLSGFEPAILLETAFNESRTVVVEEIFSKE	
28K Ping ACD36978.1	(139)	STSLGLETQSFYIGGGANSHSVLSGFEPAILLETAFNESRTVVVEEIFSKE	
28K Ping ACD36975.1	(63)	STSLGLETQSFYIGGGANSHSVLSGFEPAILLETAFNESRTVVVEEIFSKE	
28K Ping ACD36976.1	(63)	STSLGLETQSFNIGGGANSHSVLSGFEPAILLETAFNESRTVVVEEIFSKE	
28K Ping ACD36974.1	(63)	STSLGLETQSFYIGGGANSHSVLSGFEPAILLETAFNESRTVVVEEIFSKE	
Consensus	(151)	STSLGLETQSFYIGGGANSHSVLSGFEPAILLETAFNESRTVVVEEIFSKE	

FIG. 2L

9/52

	201	250
28K Eric BAB21619	(198)	LDGPIMFVDDSHAPSLWTKFLQLKKDDKEQQLKKMMQDQEEDDEEEKQTSR
28K ILSI 12697782	(201)	LDGPIMFVDDSHAPSLWTKFLQLKKDDKEQQLKKMMQDQEEDDEEEKQTSR
28K Ping ACD36978.1	(189)	LDGPIMFVDDSHAPSLWTKFLQLKKDDKEQQLKKMMQDQEEDDEEEKQTSR
28K Ping ACD36975.1	(113)	LDGPIMFVDDSHAPSLWTKFLQLKKDDKEQQLKKMMQDQEEDDEEEKQTSR
28K Ping ACD36976.1	(113)	LDGPIMFVDDSHAPSLWTKFLQLKKDDKEQQLKKMMQDQEEDDEEEKQTSR
28K Ping ACD36974.1	(113)	LDGPIMFVDDSHVPSLWTKFLQLKKDDKEQQLKKMMQDQEEDDEEEKQTSR
Consensus	(201)	LDGPIMFVDDSHAPSLWTKFLQLKKDDKEQQLKKMMQDQEEDDEEEKQTSR
	251	300
28K Eric BAB21619	(248)	SWRKLETVFGKVNEKIENKDTAGSPASYNLYDDKKADFKNAYGWSKALH
28K ILSI 12697782	(251)	SWRKLETVFGKVNEKIENKDTAGSPASYNLYDDKKADFKNAYGWSKALH
28K Ping ACD36978.1	(239)	SWRKLETVFGKVNEKIENKDTAGSPASYNLYDDKKADFKNAYGWSKALH
28K Ping ACD36975.1	(163)	SWRKLETVFGKVNEKIENKDTAGSPASYNLYDDKKADFKNAYGWSKALH
28K Ping ACD36976.1	(163)	SWRKLETVFGKVNEKIENKDTAGSPASYNLYDDKKADFKNAYGWSKALH
28K Ping ACD36974.1	(163)	SWRKLETVFGKVNEKIENKDTAGSPASYNLYDDKKADFKNAYGWSKALH
Consensus	(251)	SWRKLETVFGKVNEKIENKDTAGSPASYNLYDDKKADFKNAYGWSKALH
	301	350
28K Eric BAB21619	(298)	GGEYPPLSEPDIGVLLVKLSAGSMLAPHVNPISDEYTIVLSGYGELHIGY
28K ILSI 12697782	(301)	GGEYPPLSEPDIGVLLVKLSAGSMLAPHVNPISDEYTIVLSGYGELHIGY
28K Ping ACD36978.1	(289)	GGEYPPLSEPDIGVLLVKLSAGSMLAPHVNPISDEYTIVLSGYGELHIGY
28K Ping ACD36975.1	(213)	GGEYPPLSEPDIGVLLVKLSAGSMLAPHVNPISDEYTIVLSGYGELHIGY
28K Ping ACD36976.1	(213)	GGEYPPLSEPDIGVLLVKLSAGSMLAPHVNPISDEYTIVLSGYGELHIGY
28K Ping ACD36974.1	(213)	GGEYPPLSEPDIGVLLVKLSAGSMLAPHVNPISDEYTIVLSGYGELHIGY
Consensus	(301)	GGEYPPLSEPDIGVLLVKLSAGSMLAPHVNPISDEYTIVLSGYGELHIGY
	351	400
28K Eric BAB21619	(348)	PNGSRAMKTKIKQGDVFVVPRYFPFCQVASRDGPLEFFGFSTSAKKNKPQ
28K ILSI 12697782	(351)	PNGSRAMKTKIKQGDVFVVPRYFPFCQVASRDGPLEFFGFSTSAKKNKPQ
28K Ping ACD36978.1	(339)	PNGSKAMKTKIKQGDVFVVPRYFPFCQVASRDGPLEFFGFSTSAKKNKPQ
28K Ping ACD36975.1	(263)	PNGSKAMKTKIKQGDVFVVPRYFPFCQVASRDGPLEFFGFSTSAKKNKPQ
28K Ping ACD36976.1	(263)	PNGSKAMKTKIKQGDVFVVPRYFPFCQVASRDGPLEFFGFSTSAKKNKPQ
28K Ping ACD36974.1	(263)	PNGSKAMKTKIKQGDVFVVPRYFPFCQVASRDGPLEFFGFSTSAKKNKPQ
Consensus	(351)	PNGSKAMKTKIKQGDVFVVPRYFPFCQVASRDGPLEFFGFSTSAKKNKPQ

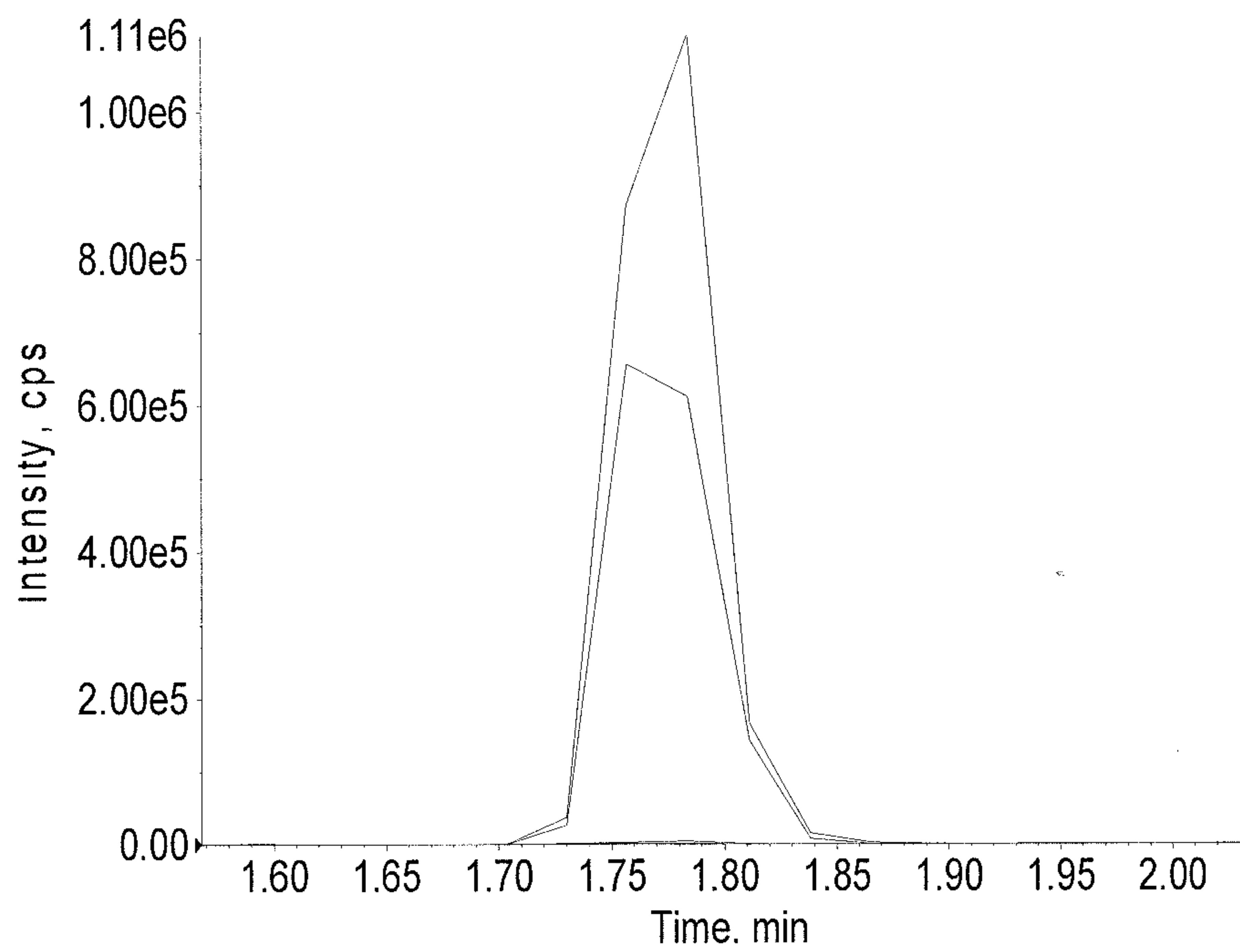
FIG. 2L - continued

10/52

	401	450
28K Eric BAB21619	(398)	FLAGAASLLRTLTMGPELSAAFGVSEDTLRRAVDAQHEAVILPSAWAAPPE
28K ILSI 12697782	(401)	FLAGAASLLRTLTMGPELSAAFGVSEDTLRRAVDAQHEAVILPSAWAAPPE
28K Ping ACD36978.1	(389)	FLAGAASLLRTLTMGPELSAAFGVSEDTLRRAVDAQHEAVILPSAWAAP-R
28K Ping ACD36975.1	(313)	FLAGAASLLRTLTMGPELSAAFGVSEDTLRRAVDAQHEAVILPSAWAAPPE
28K Ping ACD36976.1	(313)	FLAGAASLLRTLTMGPELSAAFGVSEDTLRRAVDAQHAAVILPSAWAAPPE
28K Ping ACD36974.1	(313)	FLAGAASL-----
Consensus	(401)	FLAGAASLLRTLTMGPELSAAFGVSEDTLRRAVDAQHEAVILPSAWAAPPE
	451	476
28K Eric BAB21619	(448)	NAGKLKMEEEPNAIRSFANDVVMDVF
28K ILSI 12697782	(451)	NAGKLKMEEEPNAIRSFANDVVMDVF
28K Ping ACD36978.1	(438)	KMQEAEMEESQMLLKLCQ-----
28K Ping ACD36975.1	(363)	NAGKLKMEEEP-----
28K Ping ACD36976.1	(363)	NAGKLKMEEEP-----
28K Ping ACD36974.1	(321)	-----
Consensus	(451)	NAGKLKMEEEP

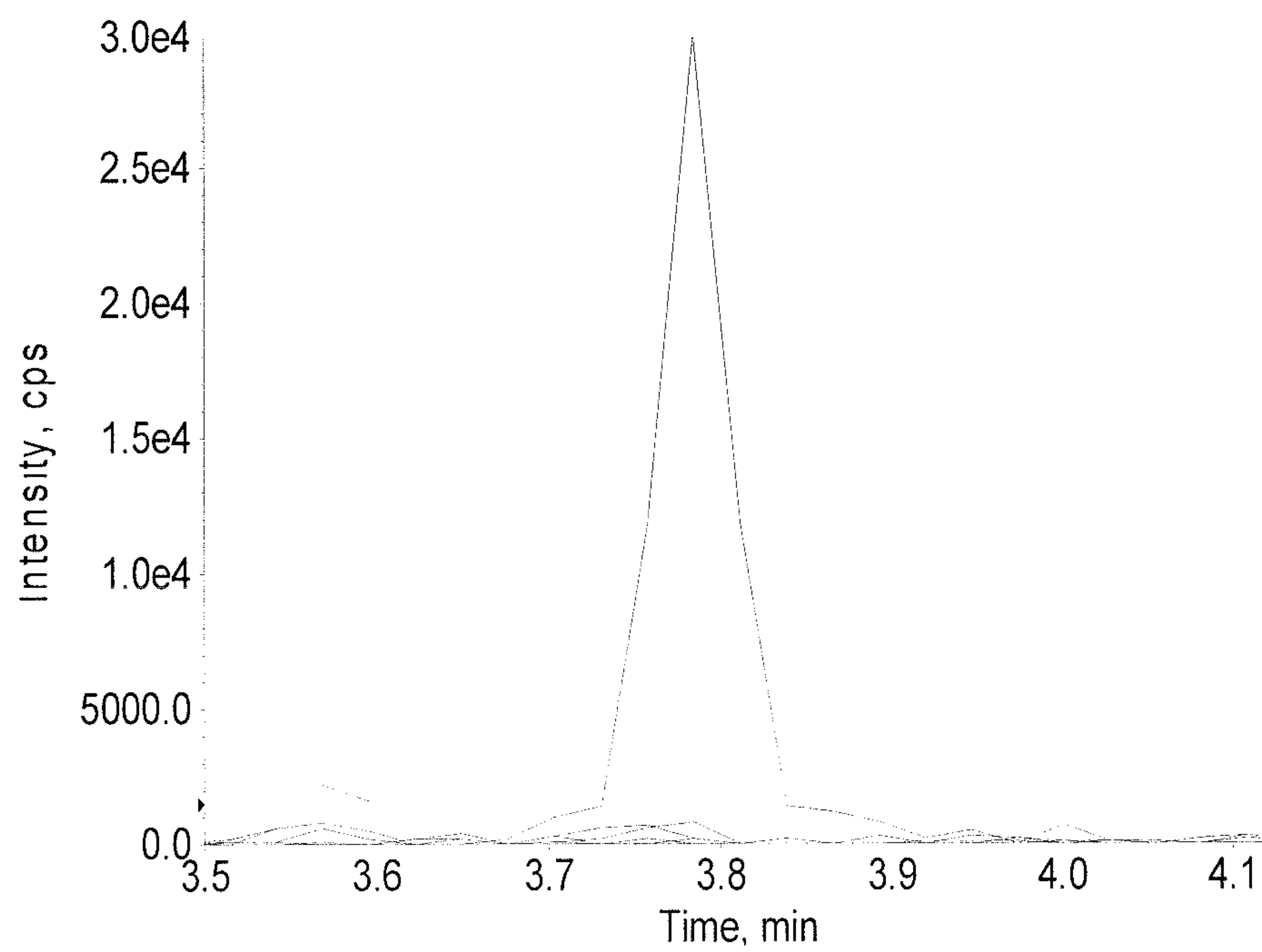
FIG. 2L – continued

11/52



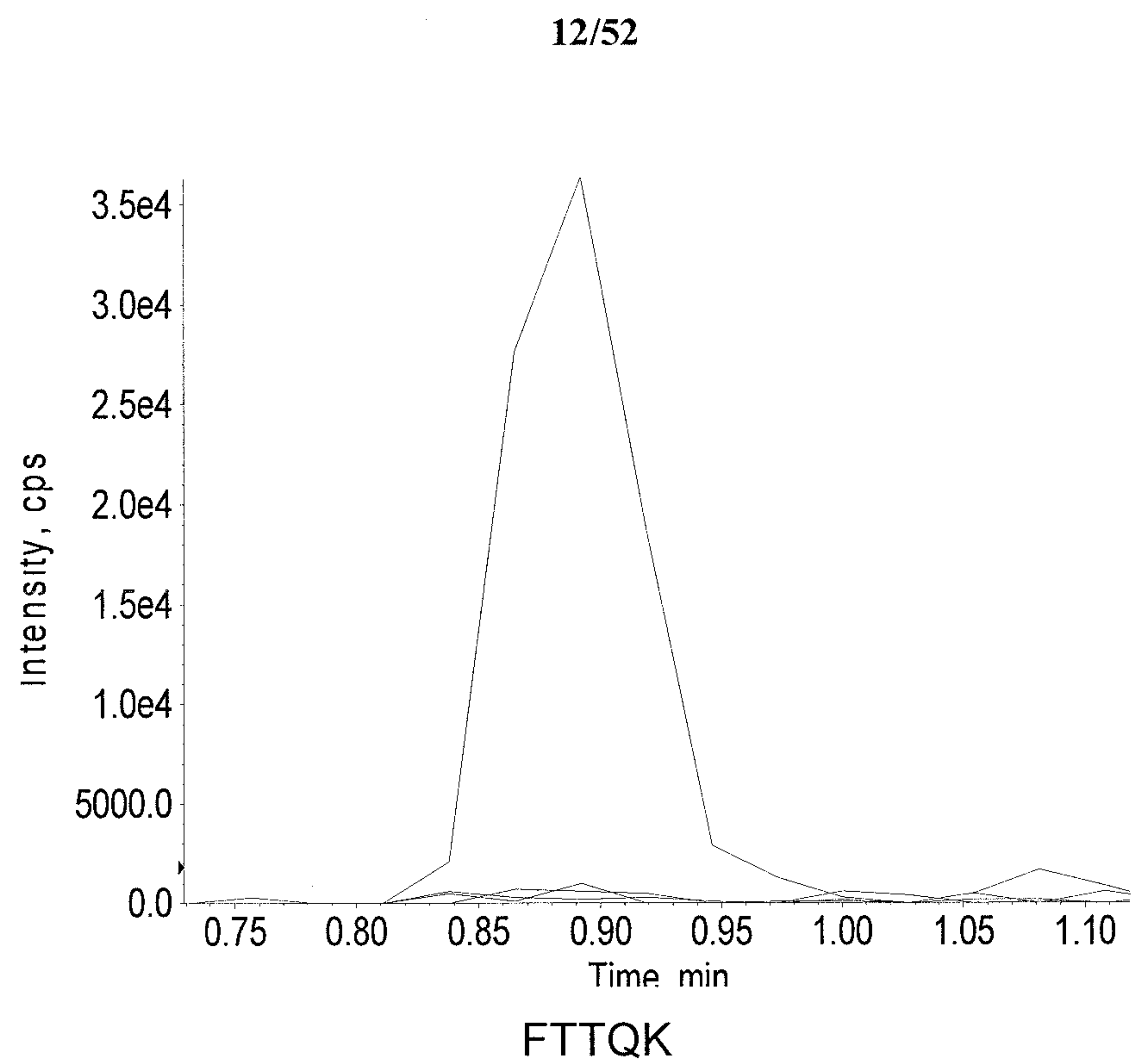
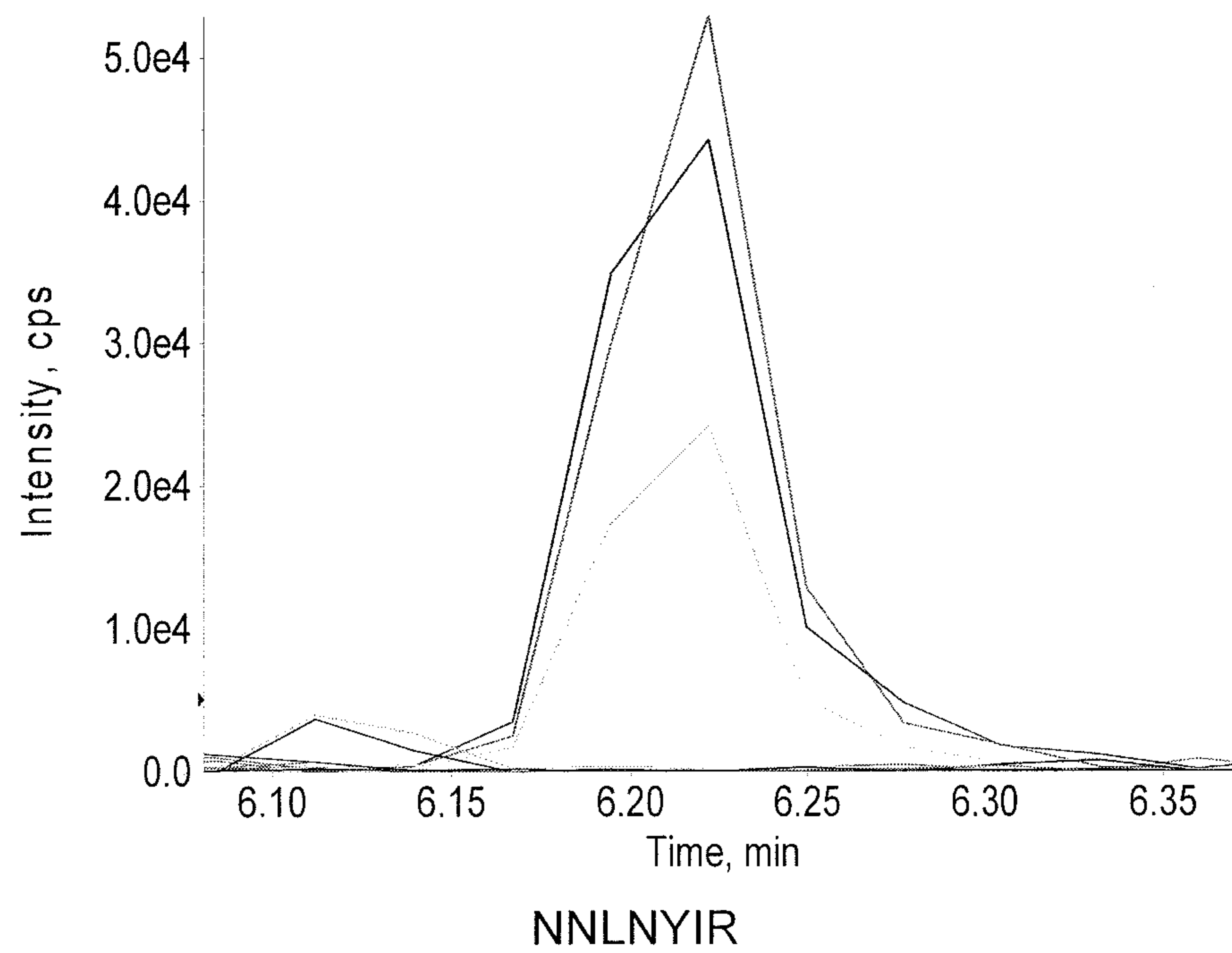
GVITQVK

FIG. 3A

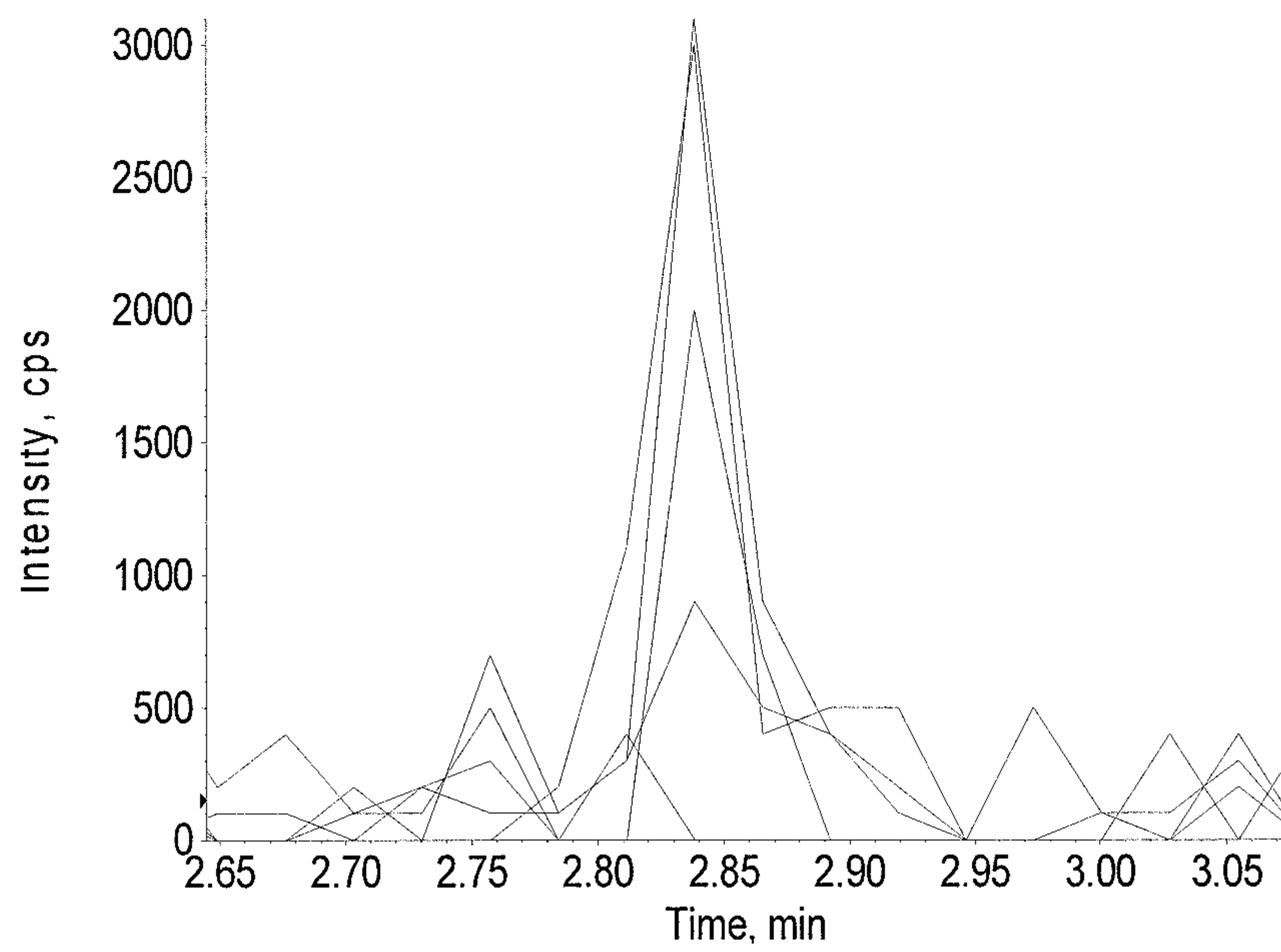


SILDLDLTK

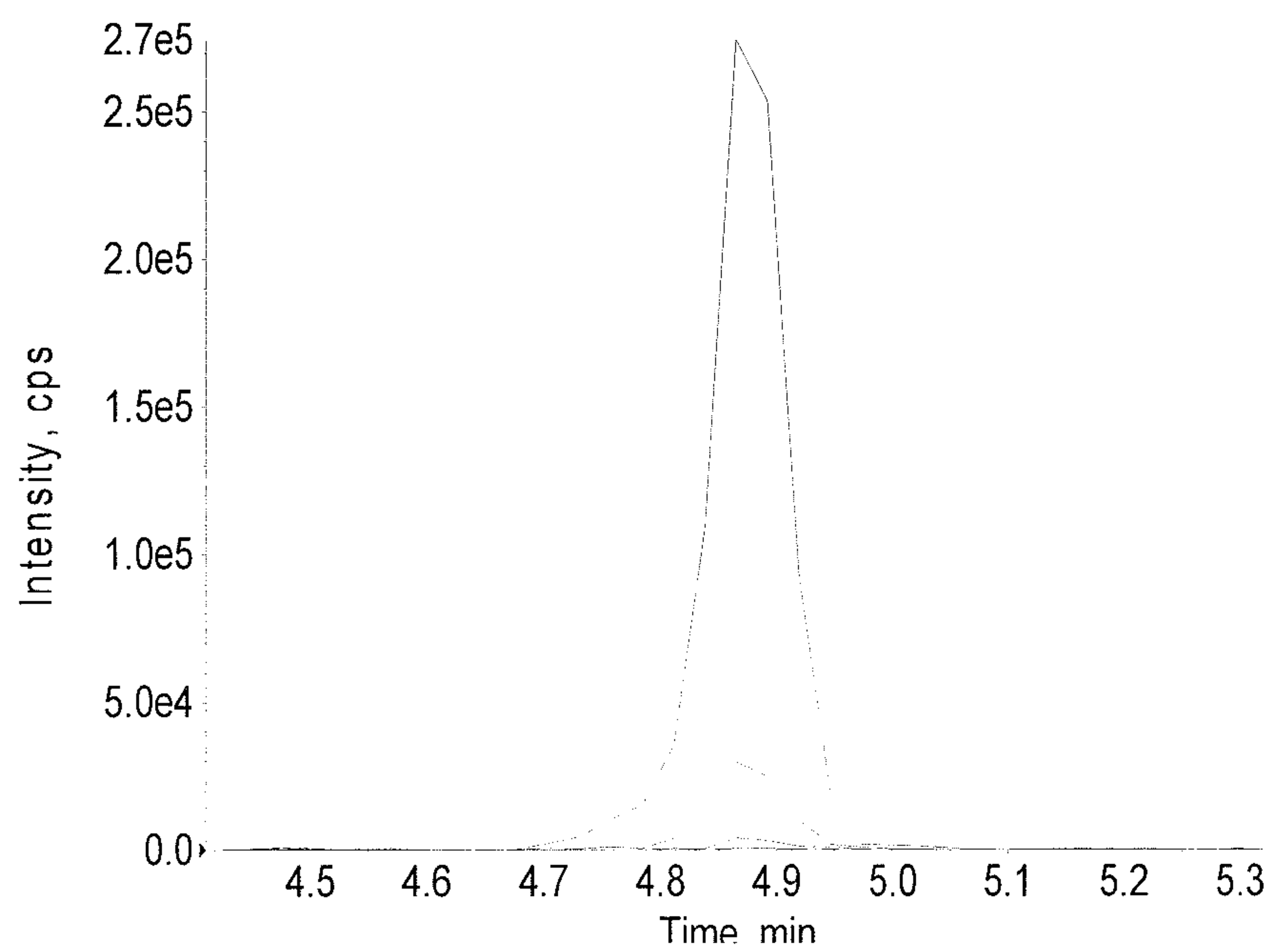
FIG. 3B

**FIG. 3C****FIG. 3D**

13/52



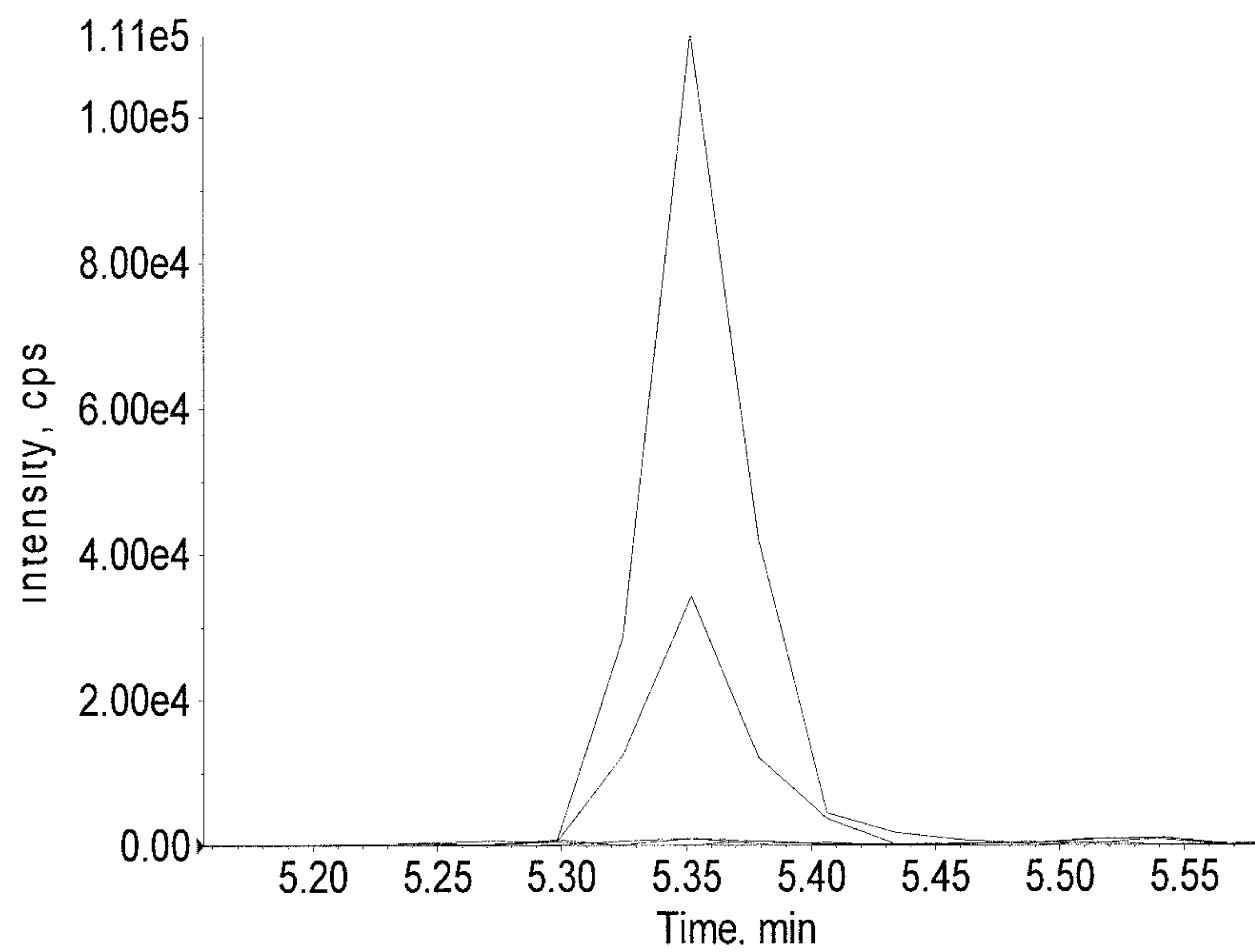
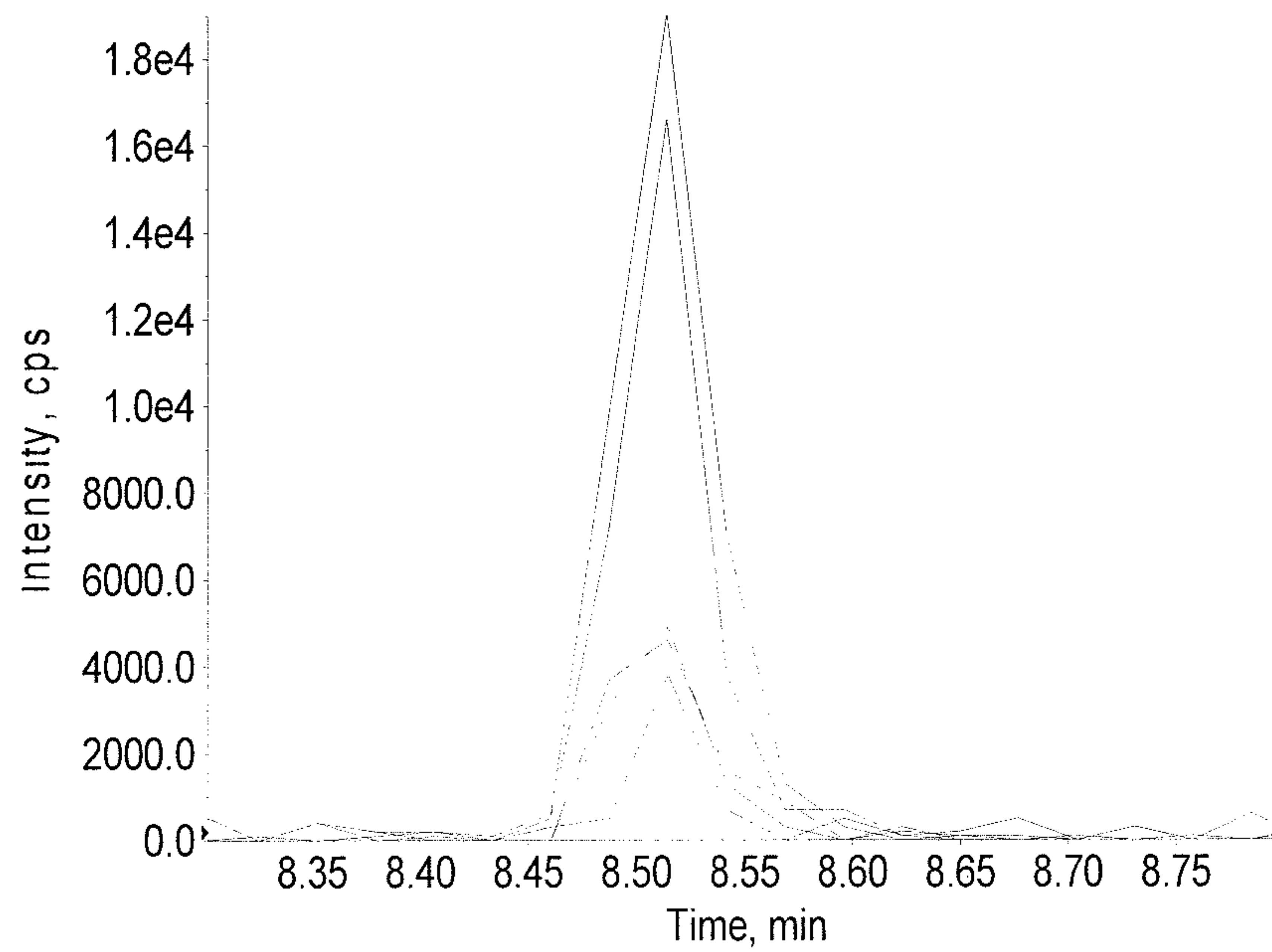
FADITPQEFSK

FIG. 3E

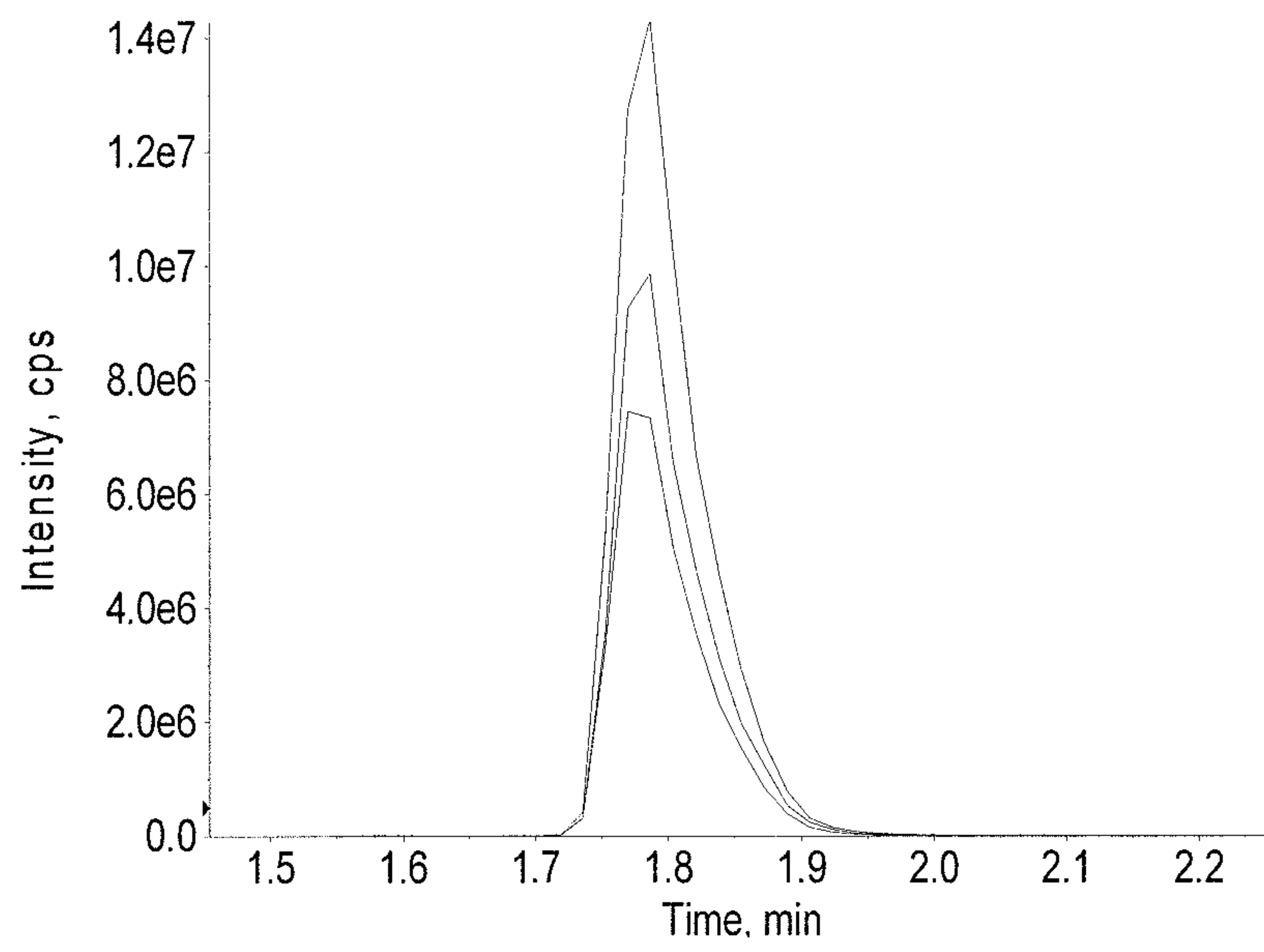
EQYSCDHPPASWDWR

FIG. 3F

14/52

**FIG. 3G****FIG. 3H**

15/52



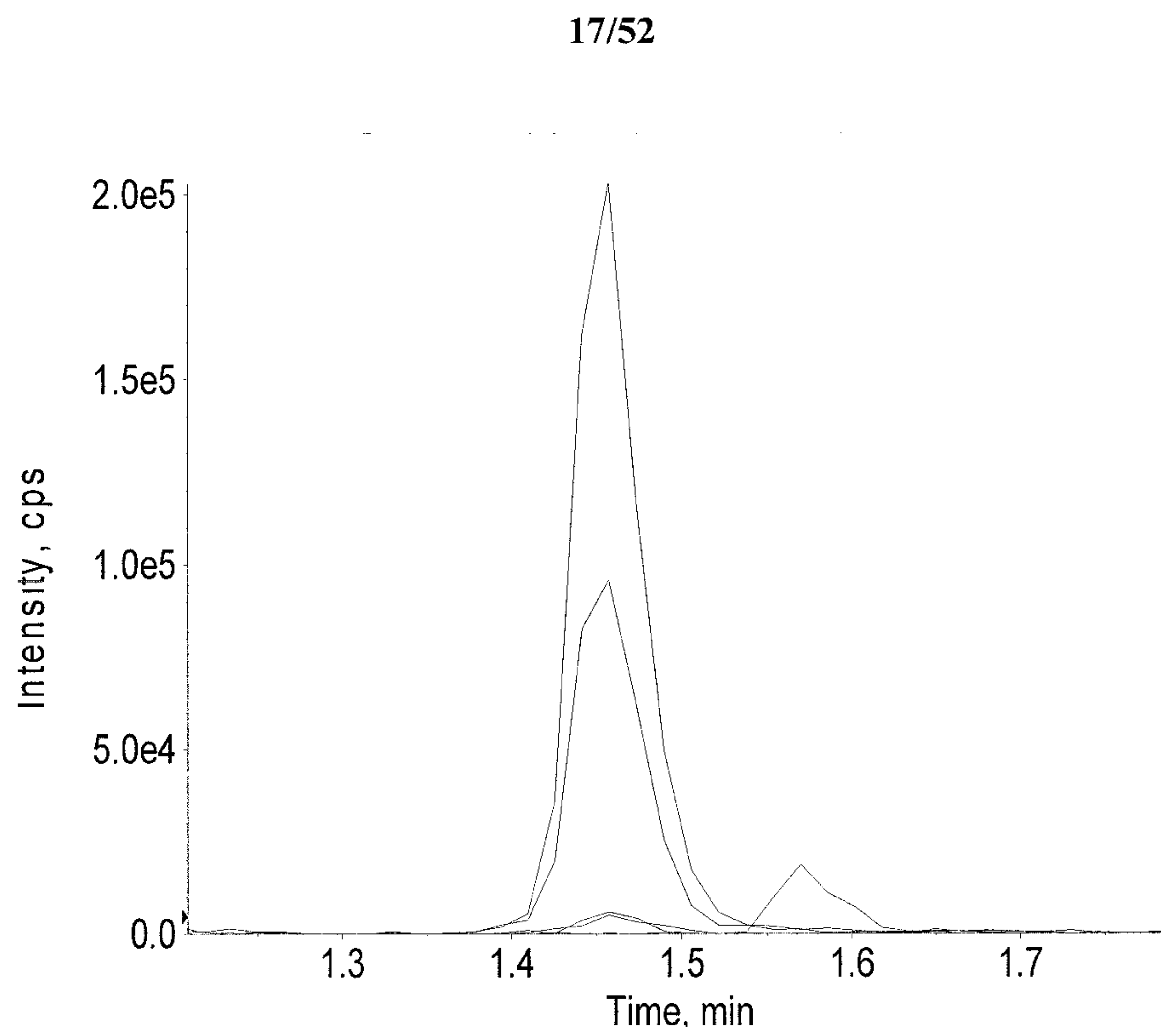
GVITQVK

FIG. 3I

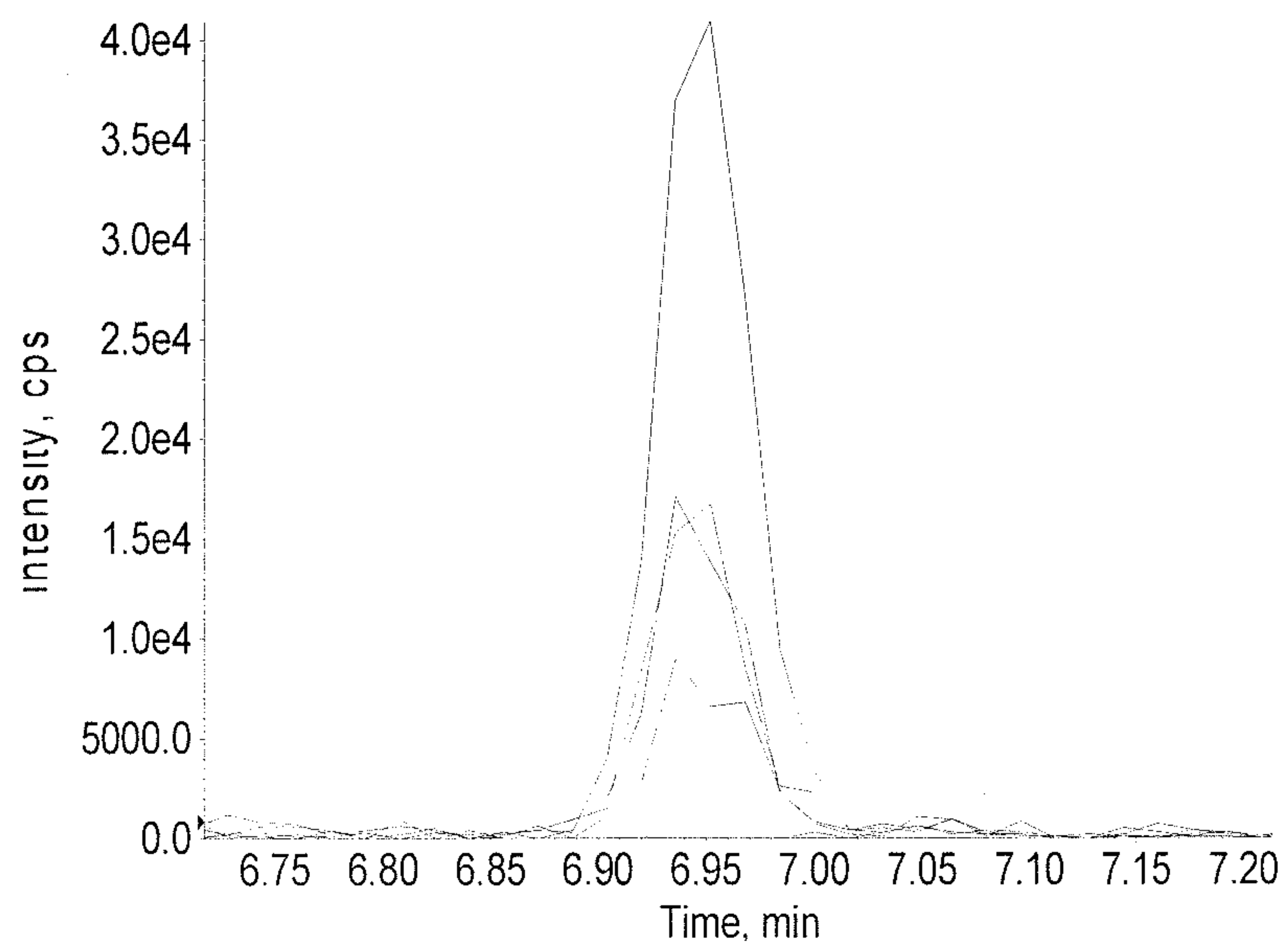
16/52

	1		50
30K ILSI 84371705	(1)	MGFLVLLLFSLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSE	
30K Ping AAB09252.1	(1)	MGFLVLLLFSLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSE	
30K Ping P22895.1	(1)	MGFLVLLLFSLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSE	
Consensus	(1)	MGFLVLLLFSLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSE	
	51		100
30K ILSI 84371705	(51)	HGRVYHNHEEEAKRLEIFKNLNLYIRDMNANRKSPHSHRLGLNKFADITP	
30K Ping AAB09252.1	(51)	HGRVYHNHEEEAKRLEIFKNNSNYIRDMNANRKSPHSHRLGLNKFADITP	
30K Ping P22895.1	(51)	HGRVYHNHEEEAKRLEIFKNNSNYIRDMNANRKSPHSHRLGLNKFADITP	
Consensus	(51)	HGRVYHNHEEEAKRLEIFKNNSNYIRDMNANRKSPHSHRLGLNKFADITP	
	101		150
30K ILSI 84371705	(101)	QEFSSKKYLQAPKDVSSQIQMANKKMKKEQYSCDHPPASWDWRKKGVITQV	
30K Ping AAB09252.1	(101)	QEFSSKKYLQAPKDVSSQIQMANKKMKKEQYSCDHPPASWDWRKKGVITQV	
30K Ping P22895.1	(101)	QEFSSKKYLQAPKDVSSQIQMANKKMKKEQYSCDHPPASWDWRKKGVITQV	
Consensus	(101)	QEFSSKKYLQAPKDVSSQIQMANKKMKKEQYSCDHPPASWDWRKKGVITQV	
	151		200
30K ILSI 84371705	(151)	KYQGGCGSGWAFSATGAIEAAHAIAITGDLVSLSEQELVDCVEESEGCSYNG	
30K Ping AAB09252.1	(151)	KYQGGCGSGWAFSATGAIEAAHAIAITGDLVSLSEQELVDCVEESEGCSYNG	
30K Ping P22895.1	(151)	KYQGGCGSGWAFSATGAIEAAHAIAITGDLVSLSEQELVDCVEESEGCSYNG	
Consensus	(151)	KYQGGCGSGWAFSATGAIEAAHAIAITGDLVSLSEQELVDCVEESEGCSYNG	
	201		250
30K ILSI 84371705	(201)	WHYQSFQWVLEHGGIATDDDDYPYRAKEGRCKANKIQDKVTIDGYETLIMS	
30K Ping AAB09252.1	(201)	WQYQSFQWVLEHGGIATDDDDYPYRAKEGRCKANKIQDKVTIDGYETLIMS	
30K Ping P22895.1	(201)	WQYQSFQWVLEHGGIATDDDDYPYRAKEGRCKANKIQDKVTIDGYETLIMS	
Consensus	(201)	WQYQSFQWVLEHGGIATDDDDYPYRAKEGRCKANKIQDKVTIDGYETLIMS	
	251		300
30K ILSI 84371705	(251)	DESTESQEQAFLSAILEQPIVSIDAKDFHLYTGGIYDGENCTSPYGIN	
30K Ping AAB09252.1	(251)	DESTESQEQAFLSAILEQPIVSIDAKDFHLYTGGIYDGENCTSPYGIN	
30K Ping P22895.1	(251)	DESTESQEQAFLSAILEQPIVSIDAKDFHLYTGGIYDGENCTSPYGIN	
Consensus	(251)	DESTESQEQAFLSAILEQPIVSIDAKDFHLYTGGIYDGENCTSPYGIN	
	301		350
30K ILSI 84371705	(301)	HFVLLVGYGSADGVYWIAKNSWGEDWGEDGYIWIQRNTGNLLGVCGMNY	
30K Ping AAB09252.1	(301)	HFVLLVGYGSADGVYWIAKNSWGEDWGEDGYIWIQRNTGNLLGVCGMNY	
30K Ping P22895.1	(301)	HFVLLVGYGSADGVYWIAKNSWGEDWGEDGYIWIQRNTGNLLGVCGMNY	
Consensus	(301)	HFVLLVGYGSADGVYWIAKNSWGEDWGEDGYIWIQRNTGNLLGVCGMNY	
	351		379
30K ILSI 84371705	(351)	FASYPTKEESETLV SARVKGHRRVDHSPL	
30K Ping AAB09252.1	(351)	FASYPTKEESETLV SARVKGHRRVDHSPL	
30K Ping P22895.1	(351)	FASYPTKEESETLV SARVKGHRRVDHSPL	
Consensus	(351)	FASYPTKEESETLV SARVKGHRRVDHSPL	

FIG. 3J



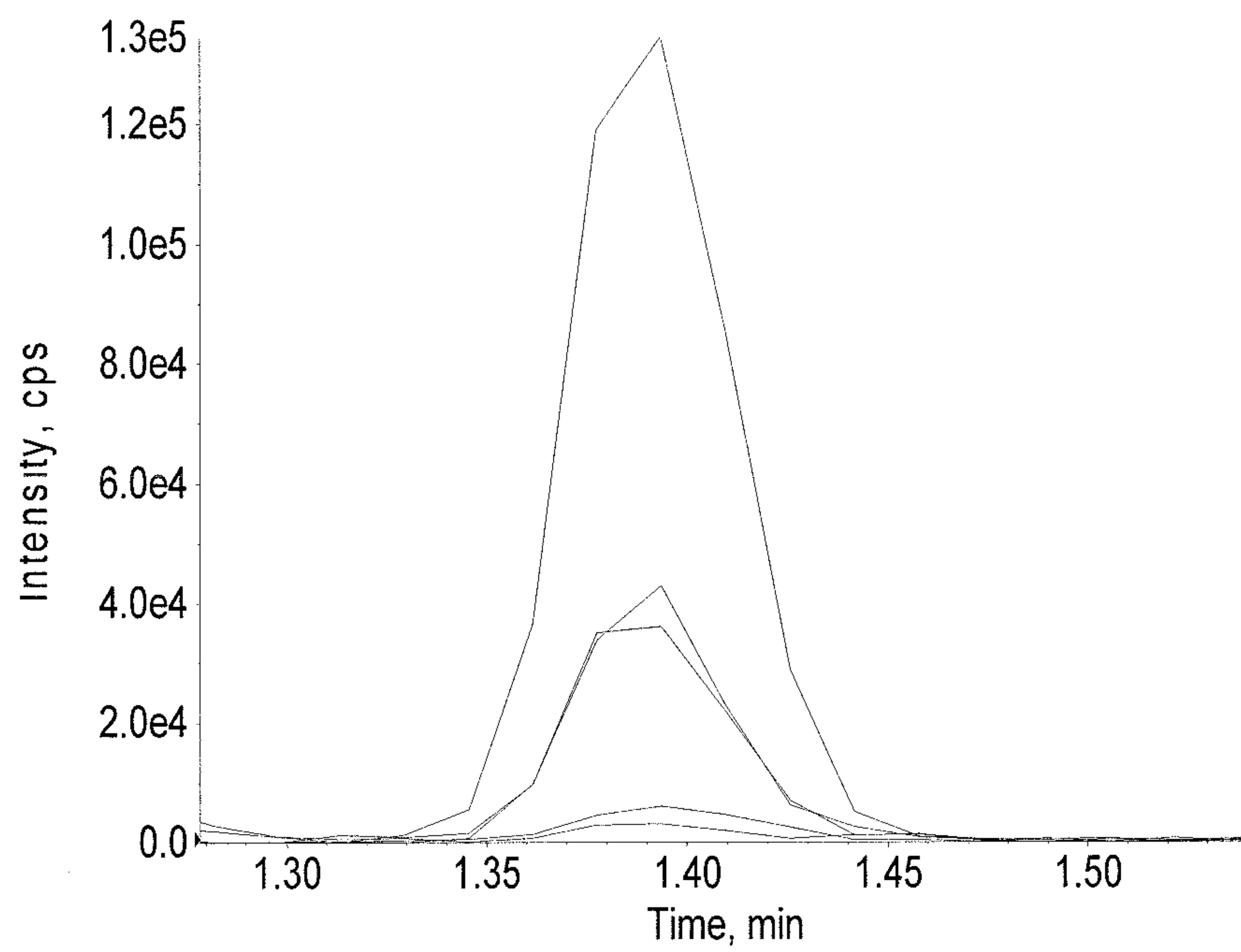
GGGIEVDSTGK

FIG. 4A

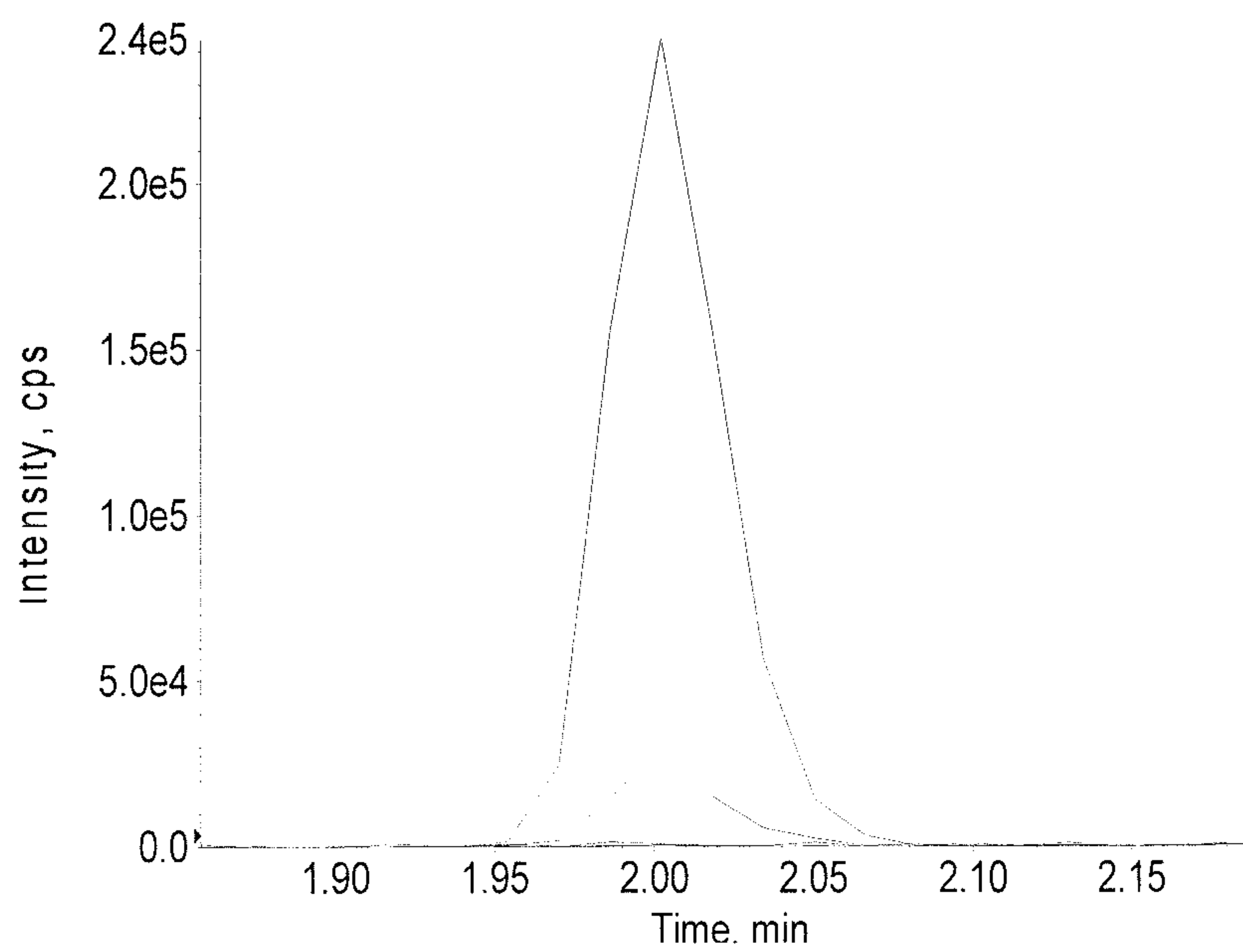
EICPLTVVQSPNELDK

FIG. 4B

18/52



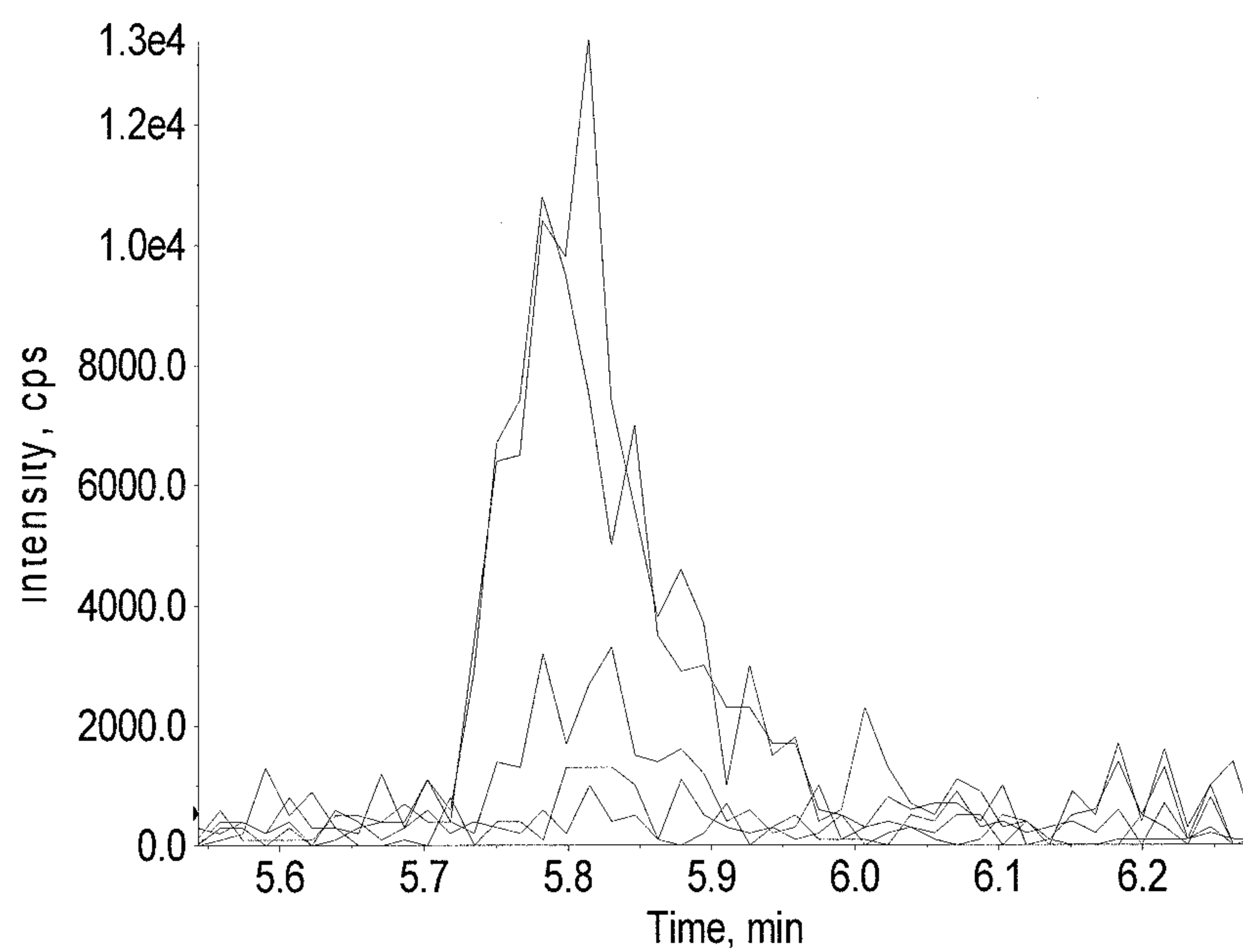
EGLQAVK

FIG. 4C

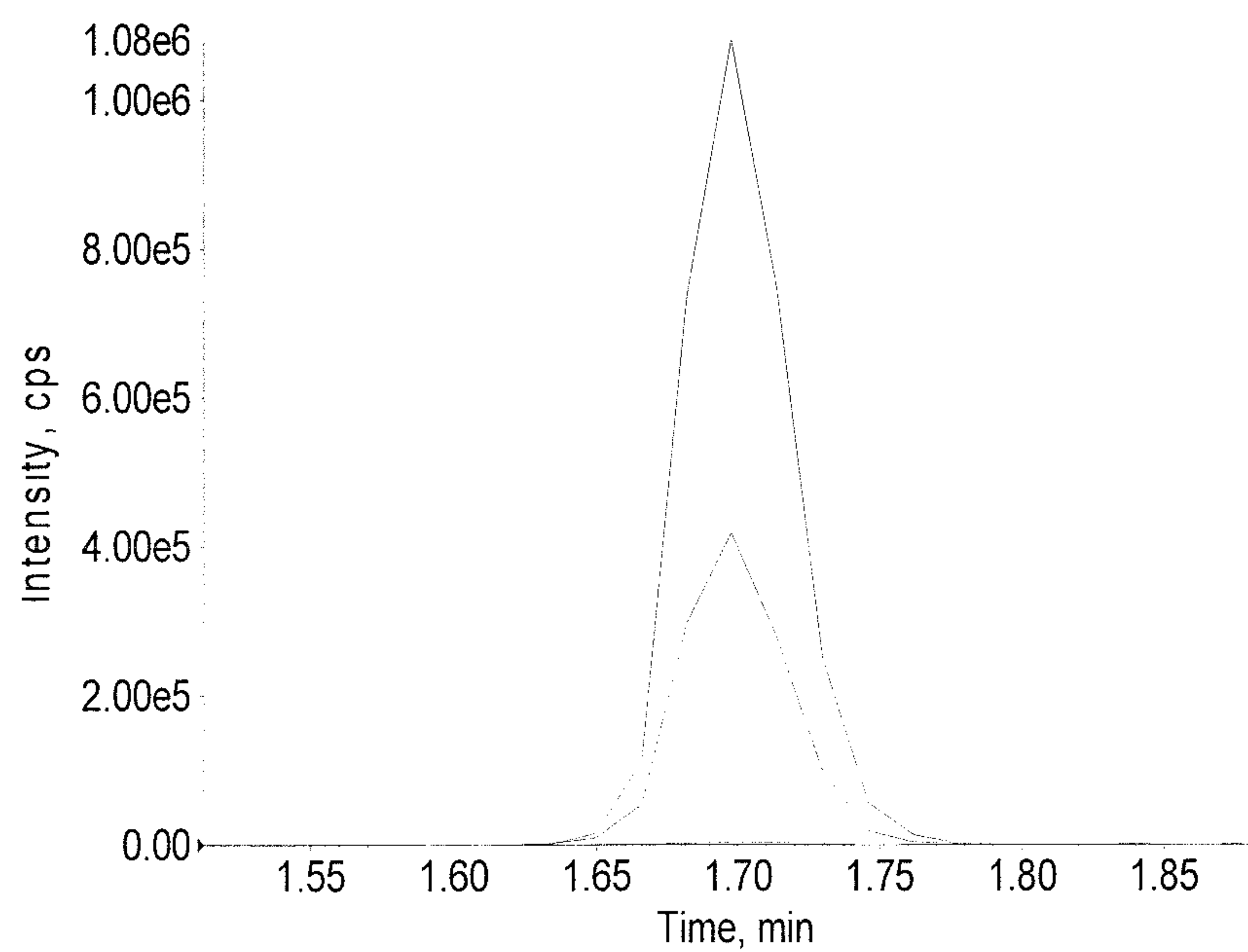
LVFCPQQAEDNK

FIG. 4D

19/52

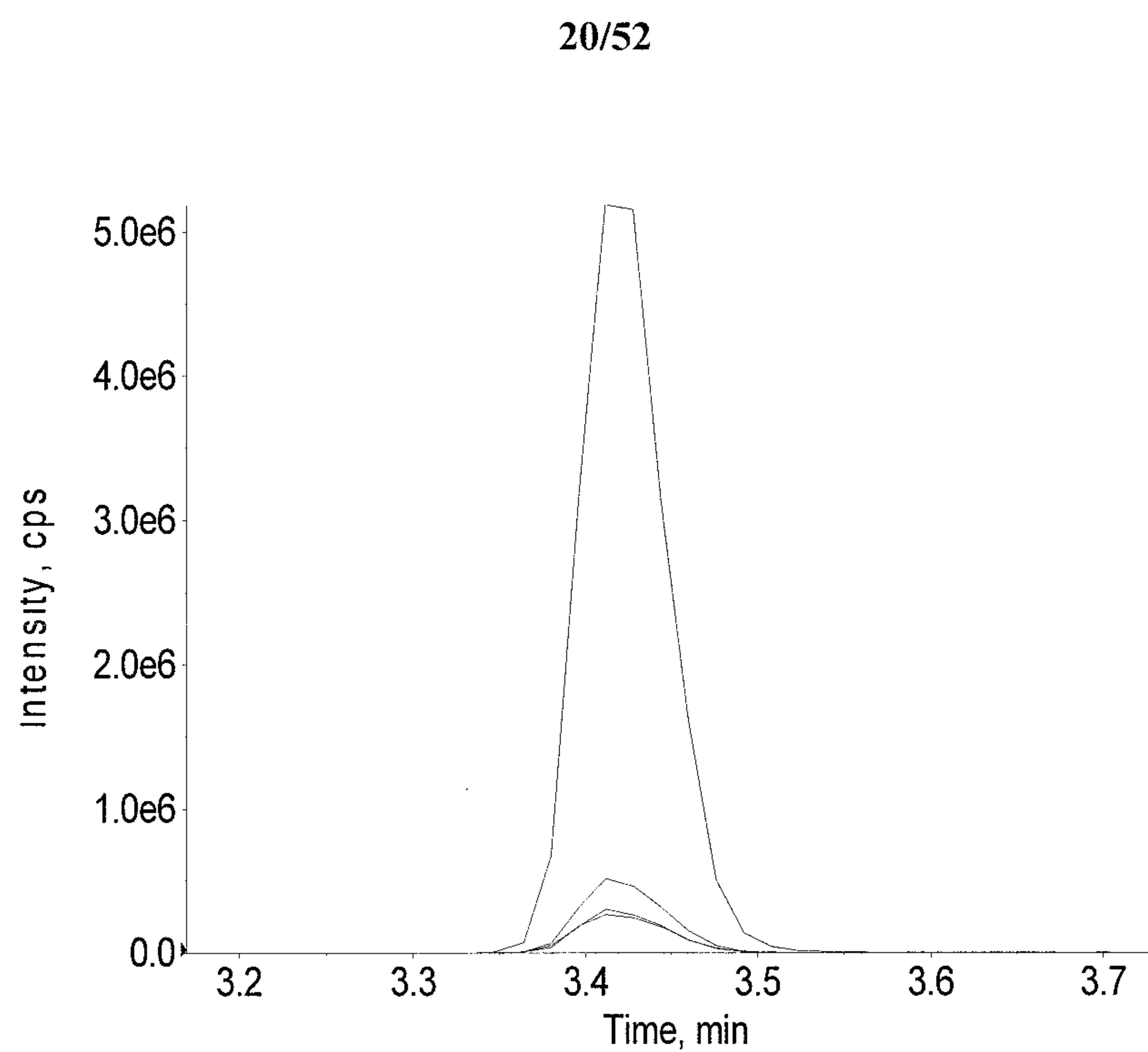


CEDIGIQIDDDGIR

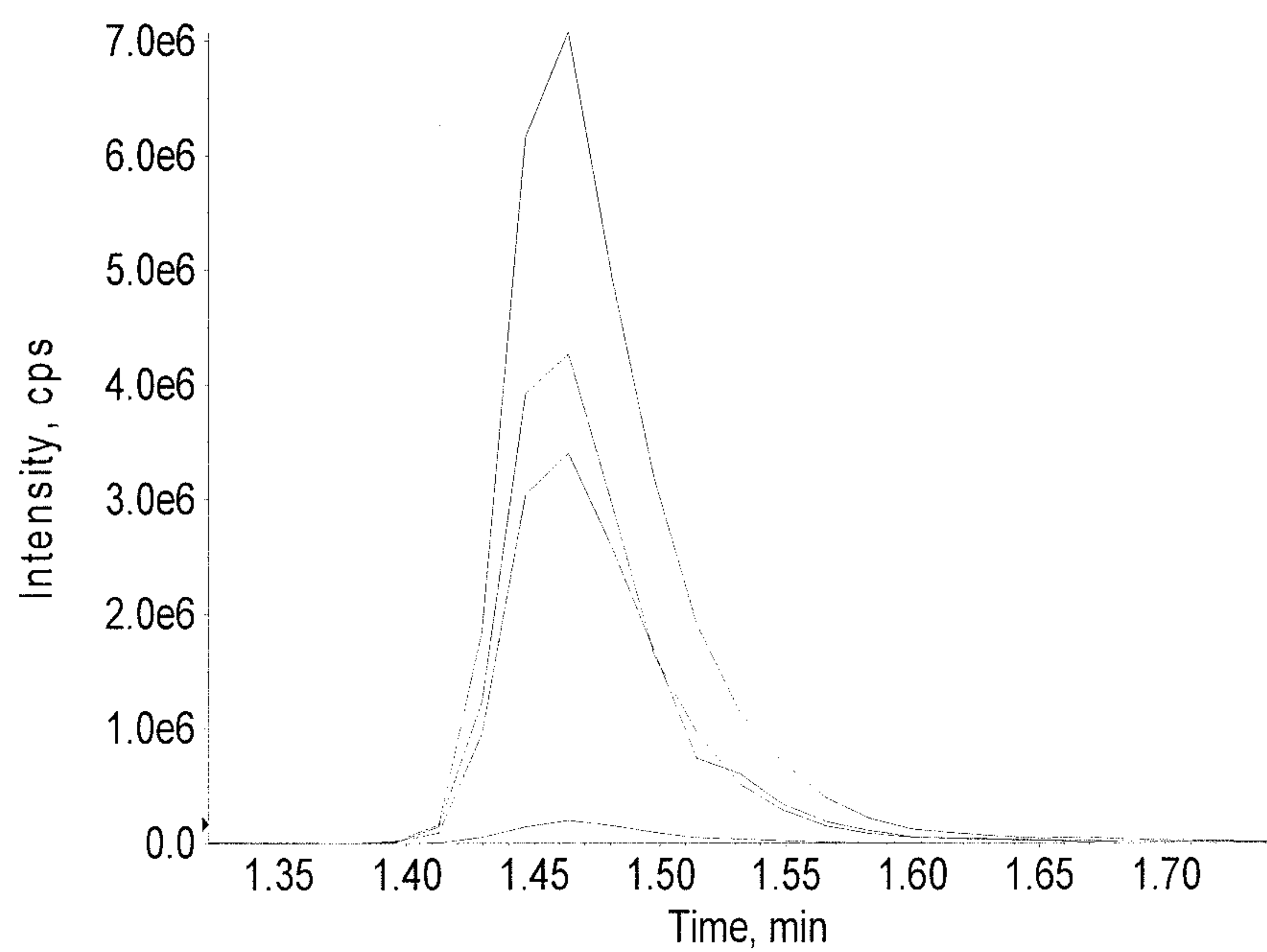
FIG. 4E

LVLSK

FIG. 4F



NKPLVVQFQK

FIG. 4G

GGGIEVDSTGK

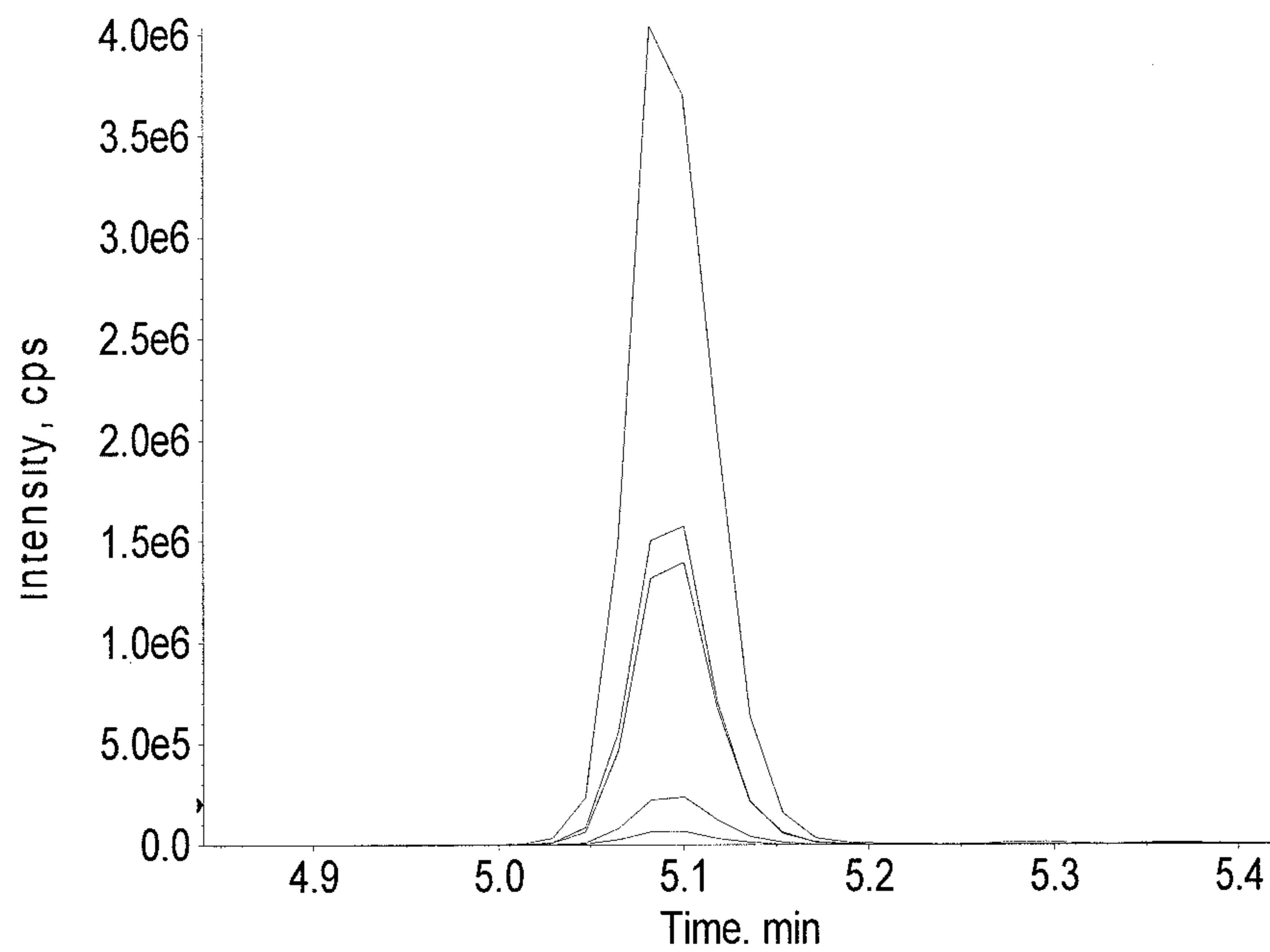
FIG. 4H

21/52

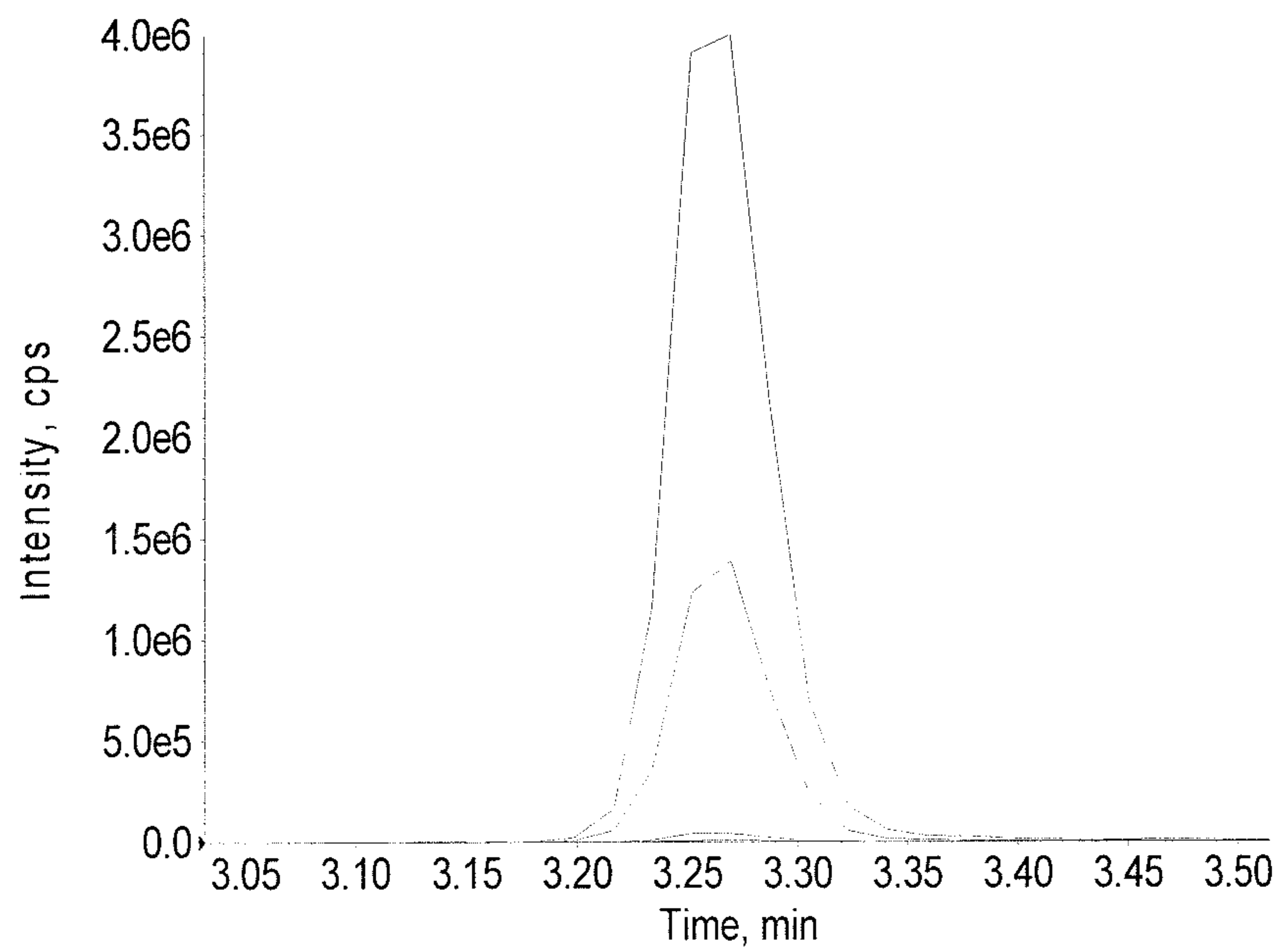
	1	50
KTI 1 AAB23482.1	(1)	MKSTIFFALFLVCAFTISYLP SATAQFVLDTDDDPLQNGGTY YMLPV MRG
KTI 1 AAB23483.1	(1)	MKSTIFFALFLVCAFTISYLP SATAQFVLDTDDDPLQNGGTY YMLPV MRG
KTI 1 CAA56343.1	(1)	MKSTTSLALFLLCALTSSYQPSATADIVFDTEGNPIRNGGTY YVLPV IRG
Consensus	(1)	MKSTIFFALFLVCAFTISYLP SATAQFVLDTDDDPLQNGGTY YMLPV MRG
	51	100
KTI 1 AAB23482.1	(51)	KGGGIEVDSTGKEICPLTVVQSPNE-LDKGIGLVFTSPLHALFIAERYPL
KTI 1 AAB23483.1	(51)	KSGGIEGNSTGKEICPLTVVQSPNK-HNKGIGLVFKSPLHALFIAERYPL
KTI 1 CAA56343.1	(51)	KGGGIEFAKTETETCPLTVVQSPFEGQLQ RGLPLIIS SPFKILDITEGLIL
Consensus	(51)	KGGGIEGASTGKEICPLTVVQSPNE LDKGIGLVFSSPLHALFIAERYPL
	101	150
KTI 1 AAB23482.1	(100)	SIKFGSFAVITLCAGMPTEWAIVEREGLOAVKLAARDTVDGWFNIERVSR
KTI 1 AAB23483.1	(100)	SIKFDSFAVIPLCGVMPTKWAIVEREGLOAVTLAARDTVDGWFNIERVSR
KTI 1 CAA56343.1	(101)	SLKFHLCTPLSLNSFSVD RYSQGSARRTPCQTHWLQKHNR CWFRIQRASS
Consensus	(101)	SIKFGSFAVISLCGGMPTKWAIVEREGLOAVTLAARDTVDGWFNIERVSR
	151	200
KTI 1 AAB23482.1	(150)	EYND-YKLVFCPQQAEDNKCEDIGIQIDDDGIRRLVLS--KNKPLVVQFQ
KTI 1 AAB23483.1	(150)	EYNDYYKLVFCPQQAEDNKCEDIGIQIDNDGIRRLVLS--KNKPLVVEFQ
KTI 1 CAA56343.1	(151)	ESNY-YKLVFCTSN-DDSSCGDIVAPIDREGNRPLIVTHDQNHPLLVQFQ
Consensus	(151)	EYND YKLVFCPQNAEDNKCEDIGIQIDNDGIRRLVLS KNKPLVVQFQ
	201	
KTI 1 AAB23482.1	(197)	KFRSSTA---
KTI 1 AAB23483.1	(198)	KFRSSTA---
KTI 1 CAA56343.1	(199)	KVEAYESSTA
Consensus	(201)	KFRSSTA

FIG. 4I

22/52

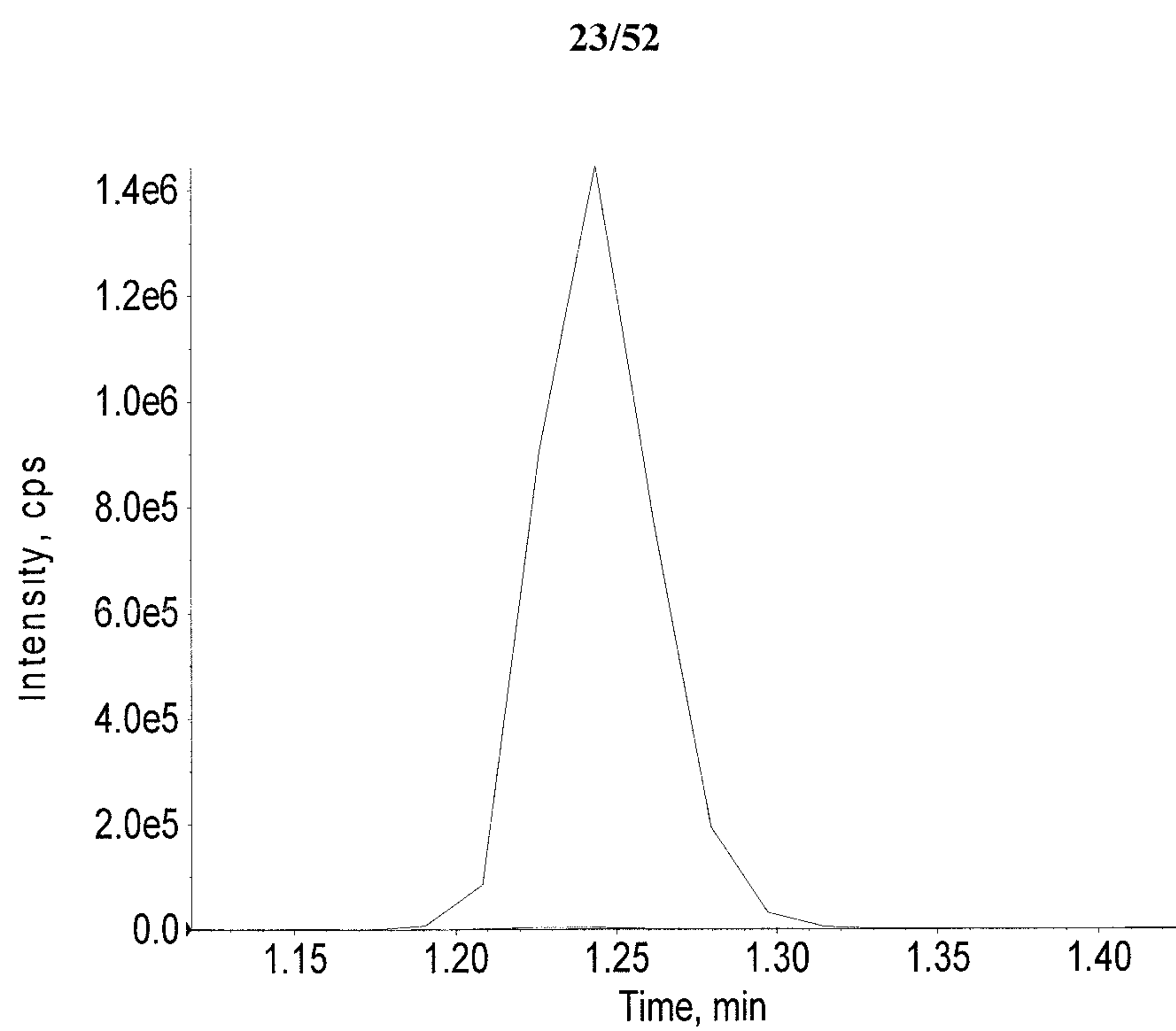


GIGTISSPYR

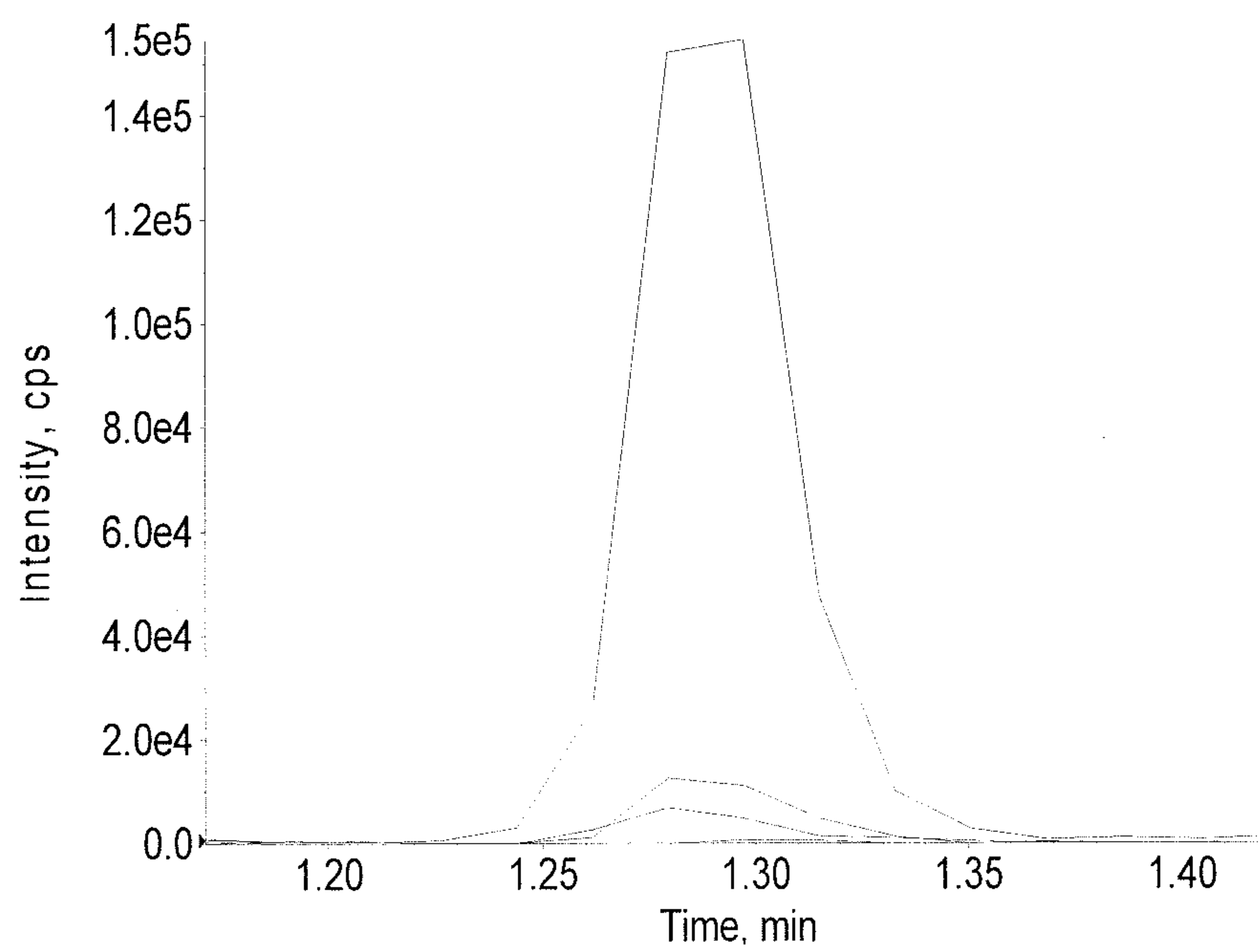
FIG. 5A

CPLTVVQSR

FIG. 5B



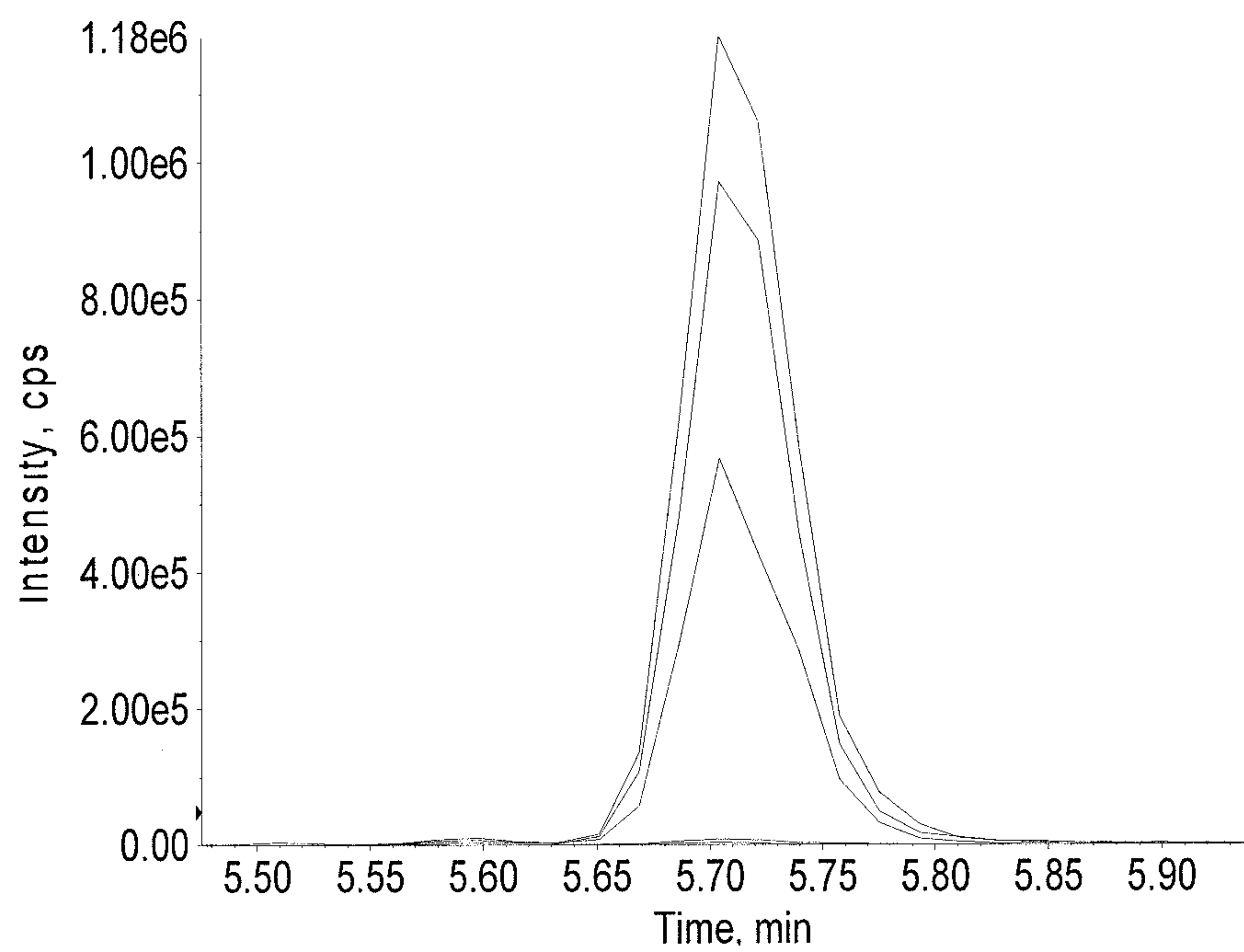
NELDK

FIG. 5C

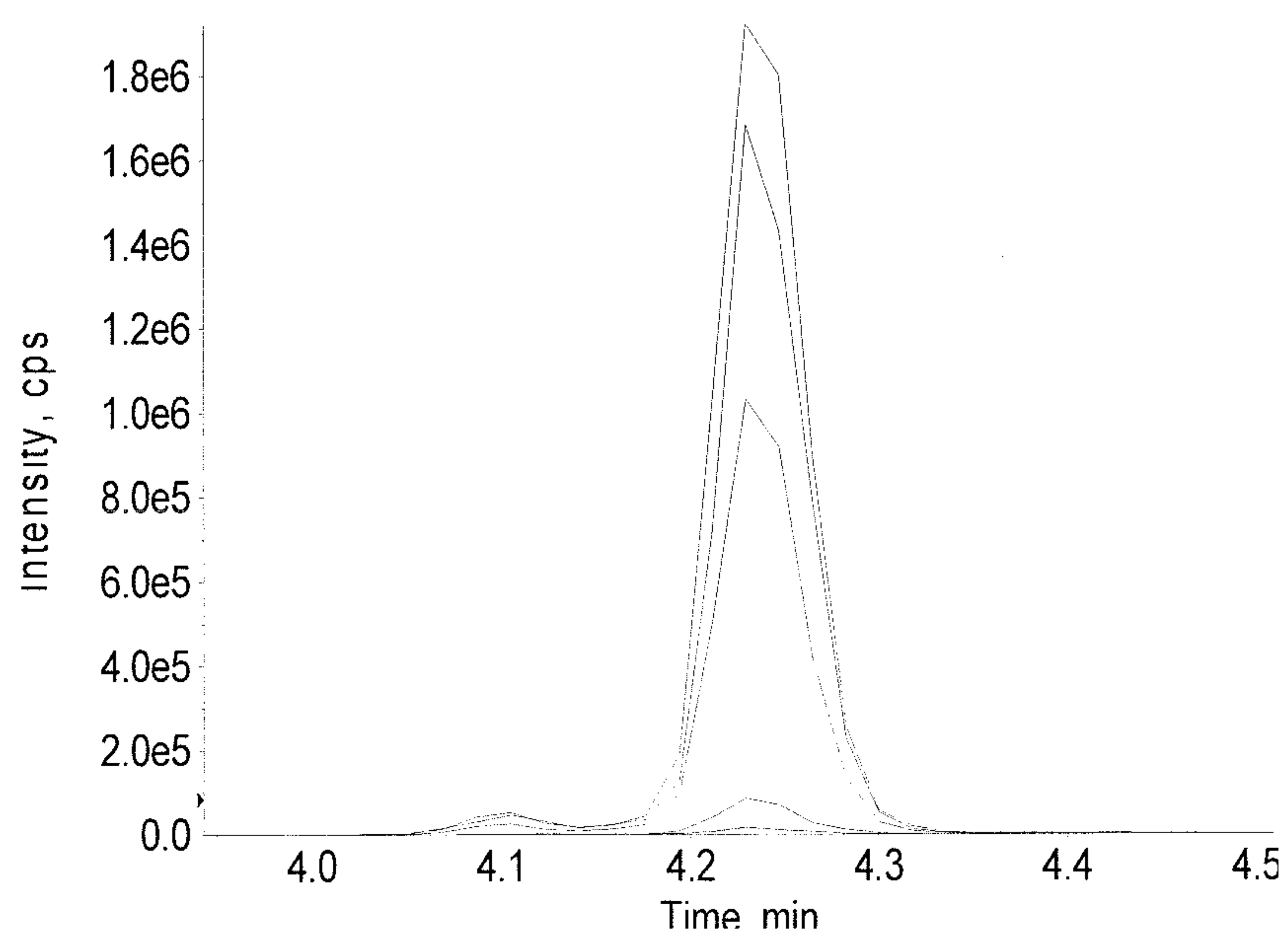
IGENK

FIG. 5D

24/52

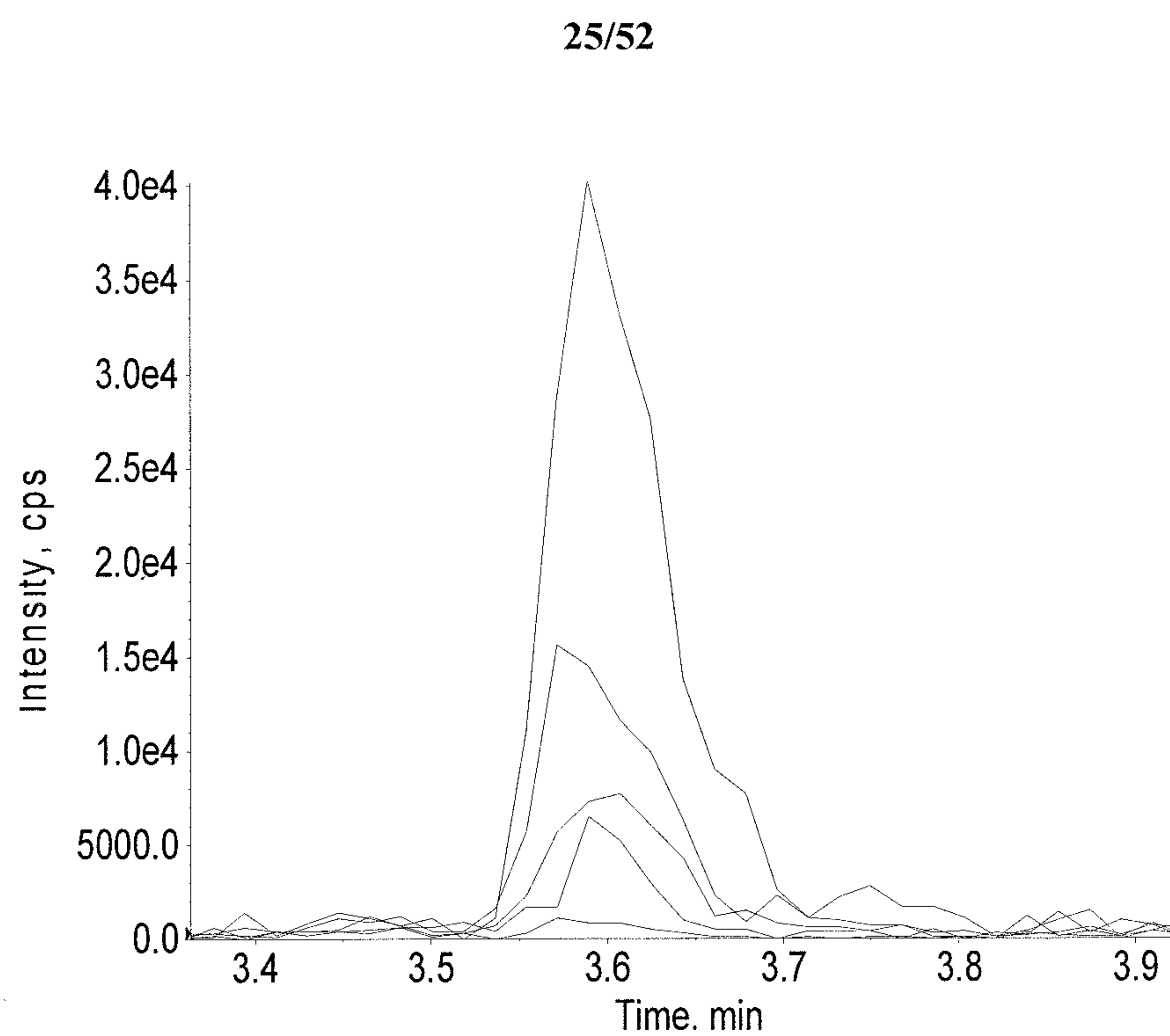
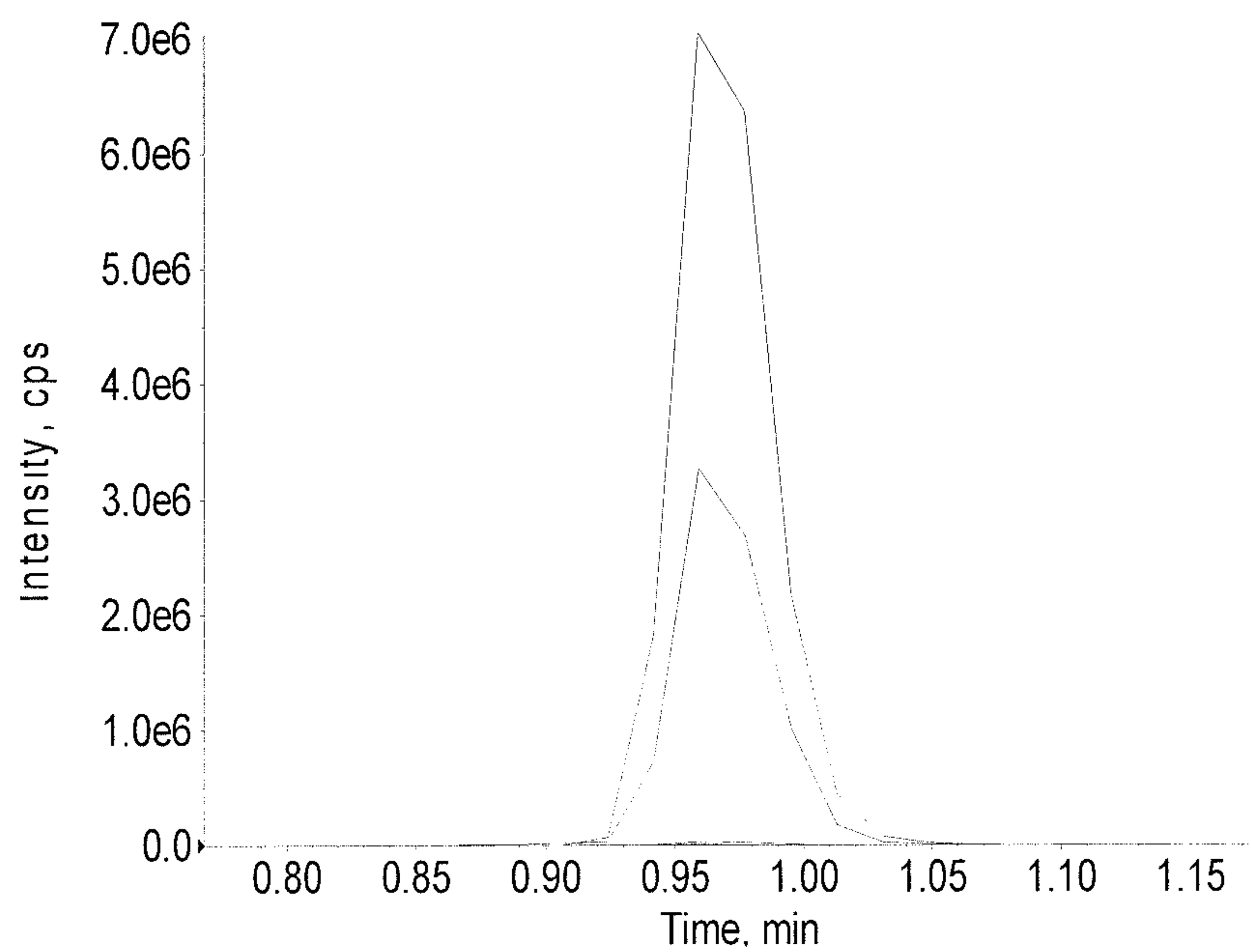


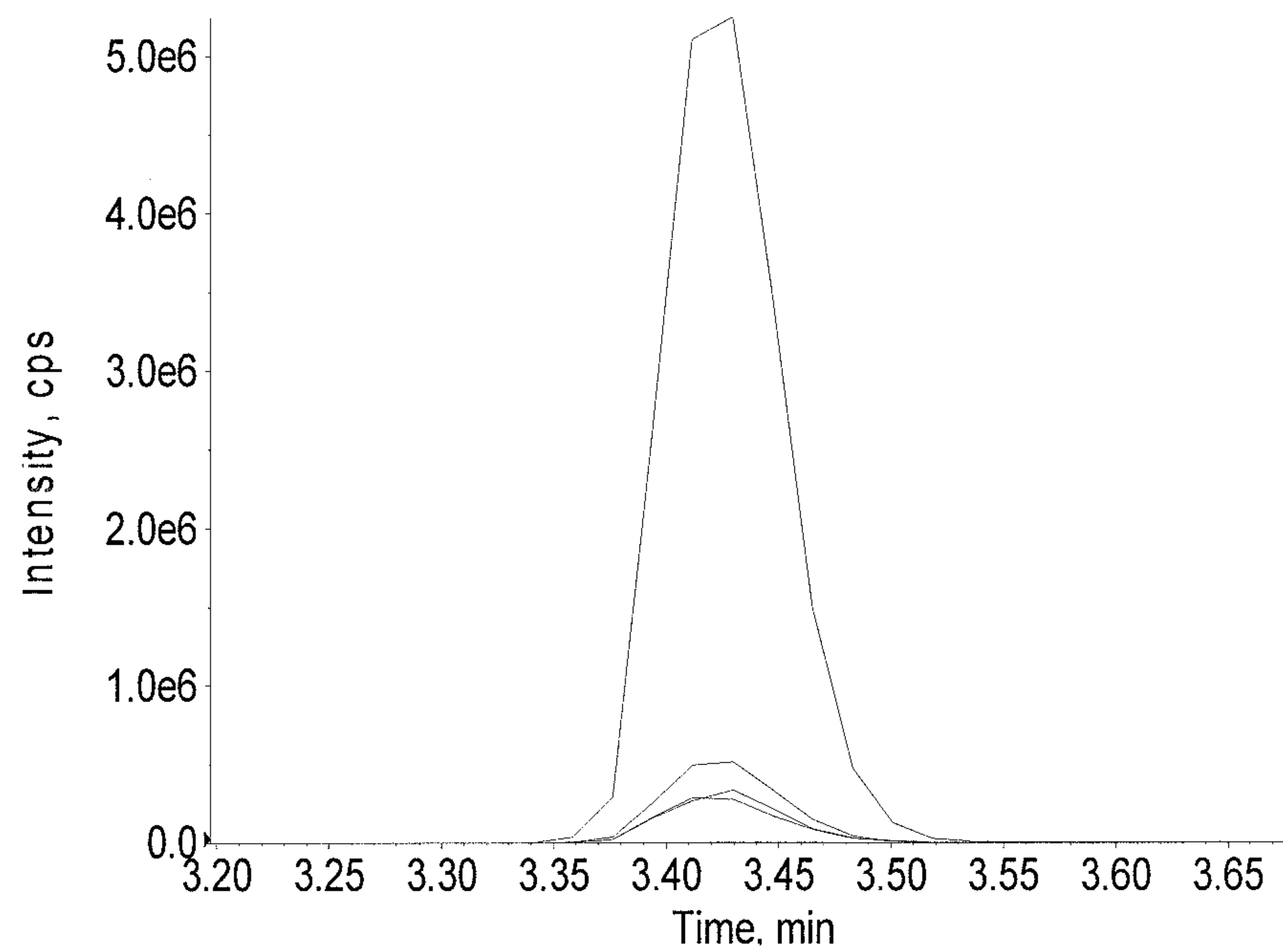
DAMDGWFR

FIG. 5E

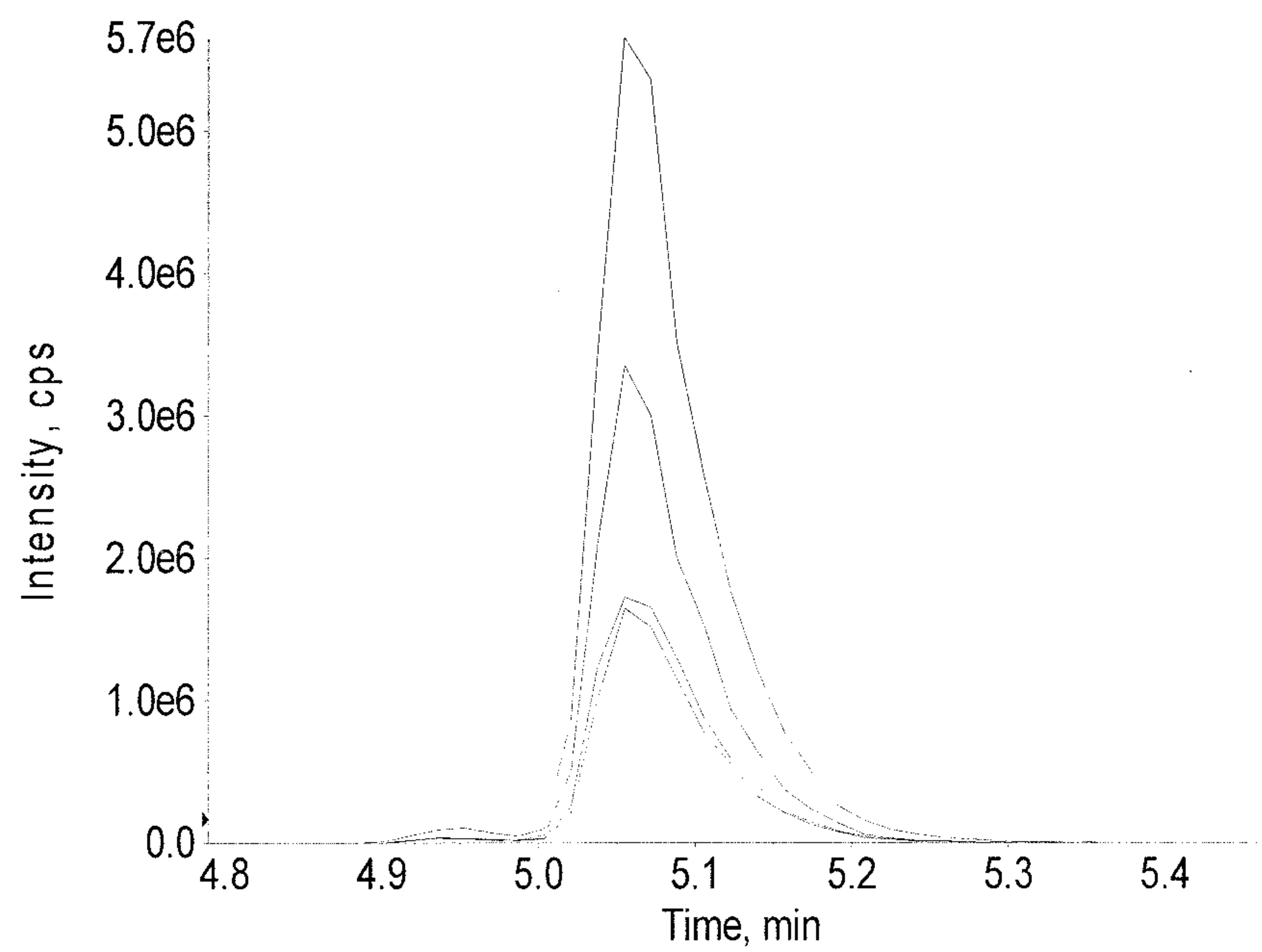
LVFCPQQAEDDK

FIG. 5F

**FIG. 5G****FIG. 5H**

26/52

NKPLVQFQK

FIG. 5I

GIGTISSPYR

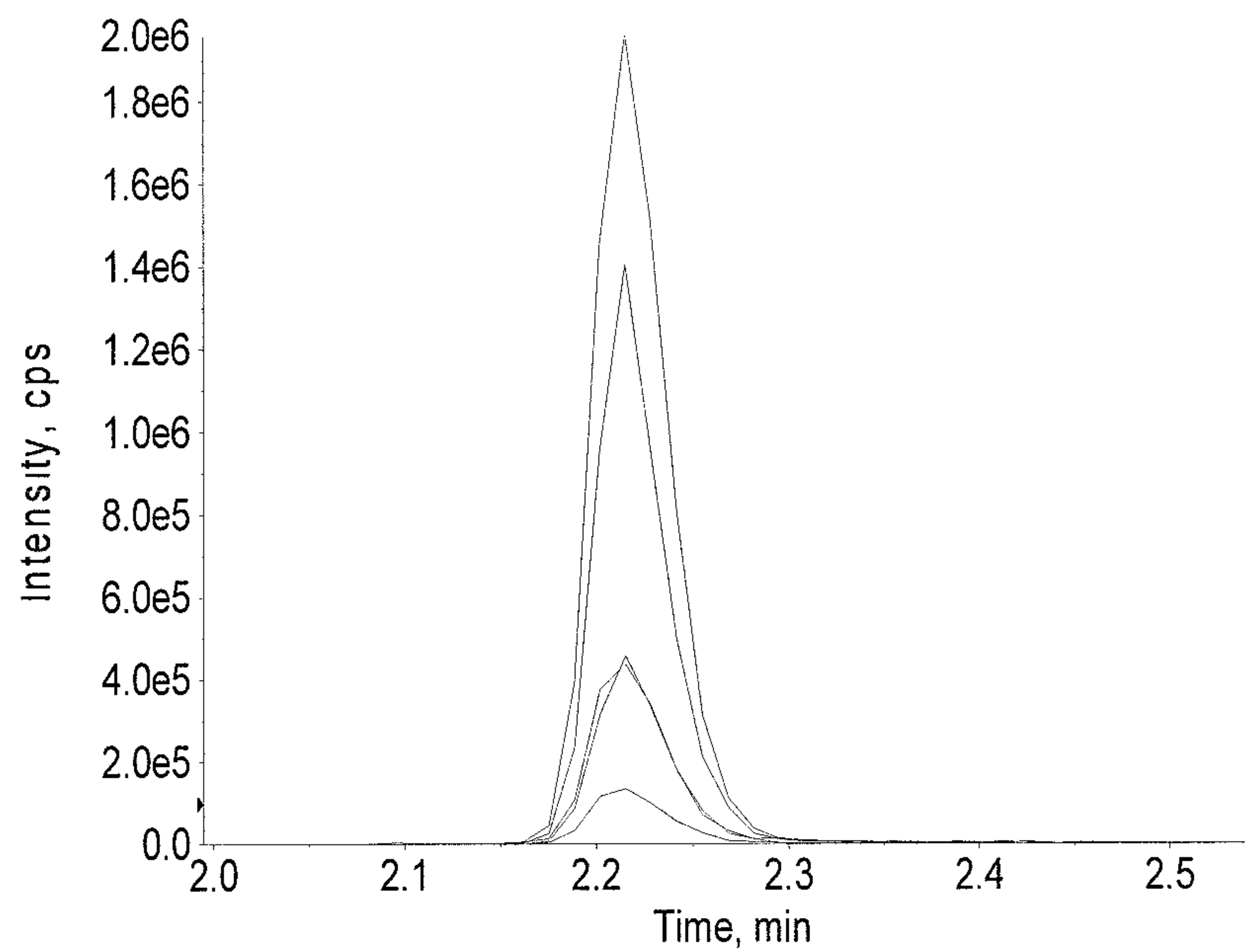
FIG. 5J

27/52

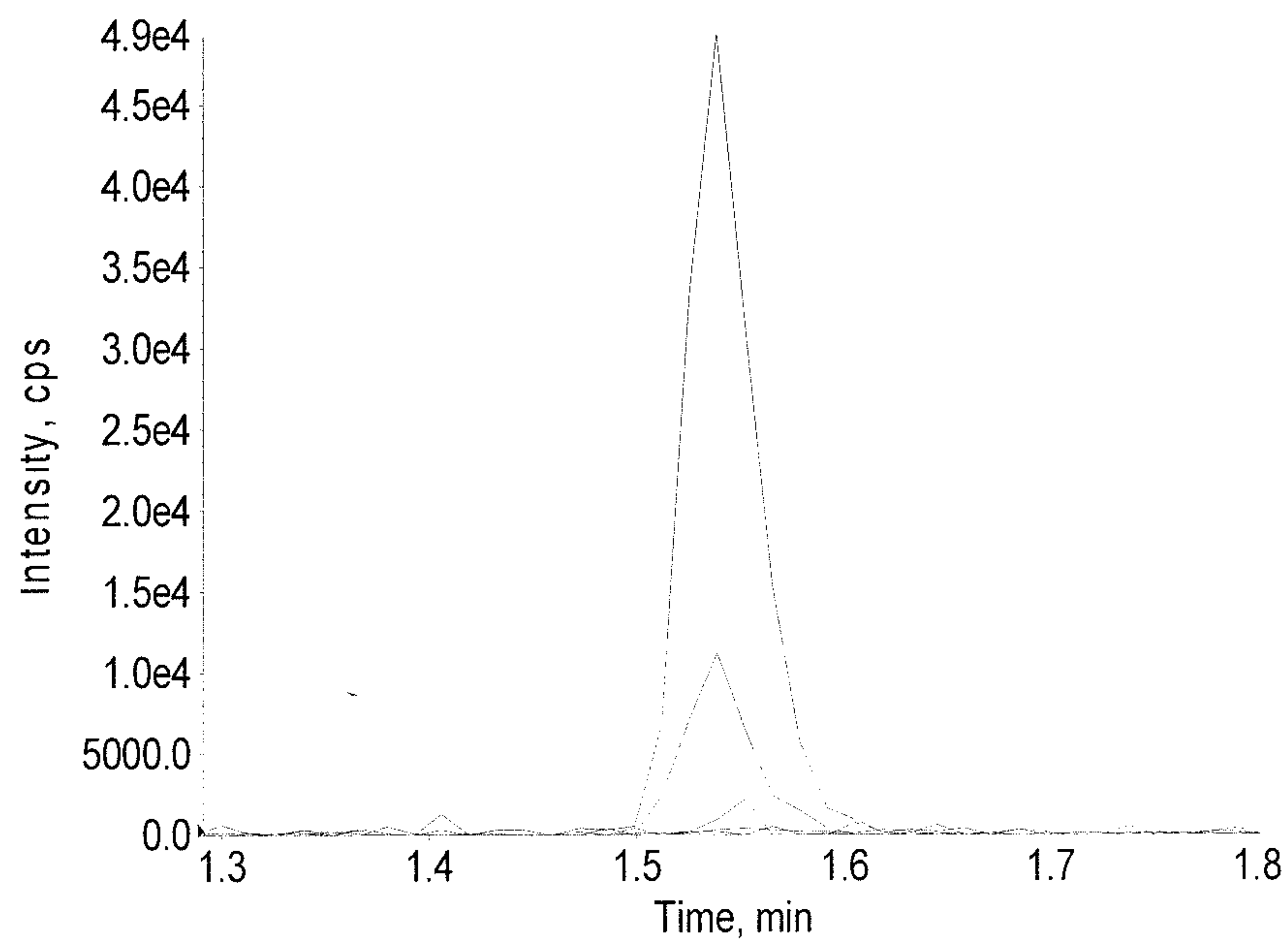
		1		50
KTI 3 AAB23464.1	(1)	MKSTIFF-LFLFCAFTTSYLPSAIADFVLDNEGNPLENGGTTYIILSDITA		
KTI 3 CAA45777.1	(1)	MKSTIFFALFLFCAFTTSYLPSAIADFVLDNEGNPLENGGTTYIILSDITA		
KTI 3 CAA45778.1	(1)	MKSTIFFALFLFCAFTTSYLPSAIADFVLDNEGNPLDSGGTTYIILSDITA		
Consensus	(1)	MKSTIFFALFLFCAFTTSYLPSAIADFVLDNEGNPLENGGTTYIILSDITA		
		51		100
KTI 3 AAB23464.1	(50)	FGGIRAAPTGNERCPLTVVQSRNELDKGIGTIISSPYRIRFIAEGHPLSL		
KTI 3 CAA45777.1	(51)	FGGIRAAPTGNERCPLTVVQSRNELDKGIGTIISSPYRIRFIAEGHPLSL		
KTI 3 CAA45778.1	(51)	FGGIRAAPTGNERCPLTVVQSRNELDKGIGTIISSPFIRIRFIAEGNPLRL		
Consensus	(51)	FGGIRAAPTGNERCPLTVVQSRNELDKGIGTIISSPYRIRFIAEGHPLSL		
		101		150
KTI 3 AAB23464.1	(100)	KFDSFAVIMLCVGIPTEWSVVEDLPEGPAVKIGENKDAMDGWFRLERVSD		
KTI 3 CAA45777.1	(101)	KFDSFAVIMLCVGIPTEWSVVEDLPEGPAVKIGENKDAMDGWFRLERVSD		
KTI 3 CAA45778.1	(101)	KFDSFAVIMLCVGIPTEWSVVEDLPEGPAVKIGENKDAVDGWFRIERVSD		
Consensus	(101)	KFDSFAVIMLCVGIPTEWSVVEDLPEGPAVKIGENKDAMDGWFRLERVSD		
		151		200
KTI 3 AAB23464.1	(150)	DEFNNYKLVFCPQQAEDDKCGDIGISIDHDDGTRRLVVSKNKPLVVQFQK		
KTI 3 CAA45777.1	(151)	DEFNNYKLVFCPQQAEDDKCGDIGISIDHDDGTRRLVVSKNKPLVVQFQK		
KTI 3 CAA45778.1	(151)	DEFNNYKLVFCTQQAEDDKCGDIGISIDHDDGTRRLVVSKNKPLVVQFQK		
Consensus	(151)	DEFNNYKLVFCPQQAEDDKCGDIGISIDHDDGTRRLVVSKNKPLVVQFQK		
		201	217	
KTI 3 AAB23464.1	(200)	LDKESLAKKNHGLSRSE		
KTI 3 CAA45777.1	(201)	LDKESLAKKNHGLSRSE		
KTI 3 CAA45778.1	(201)	VDKESLAKKNHGLSRSE		
Consensus	(201)	LDKESLAKKNHGLSRSE		

FIG. 5K

28/52



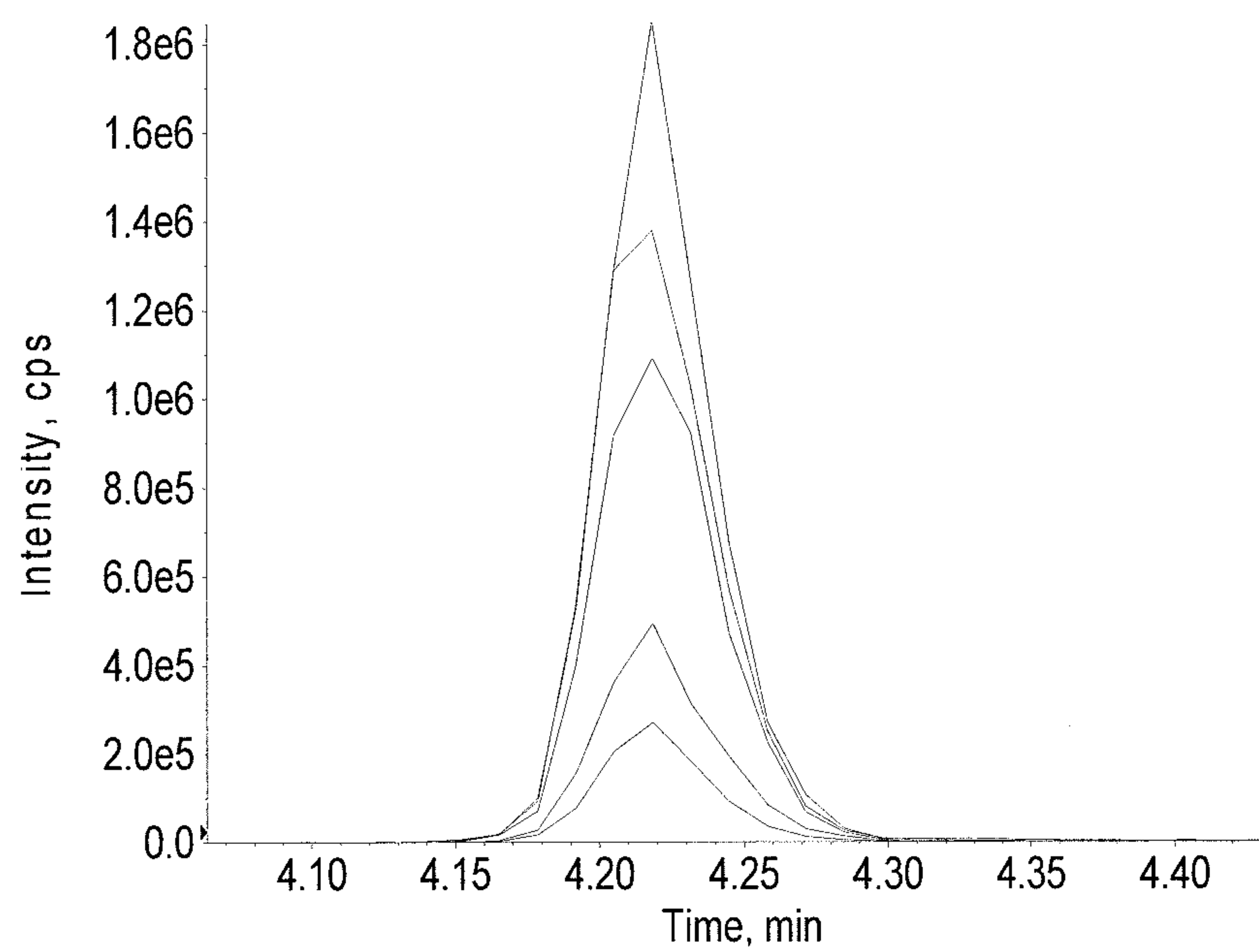
IMENQSEELEEK

FIG. 6A

WQHQQDSCR

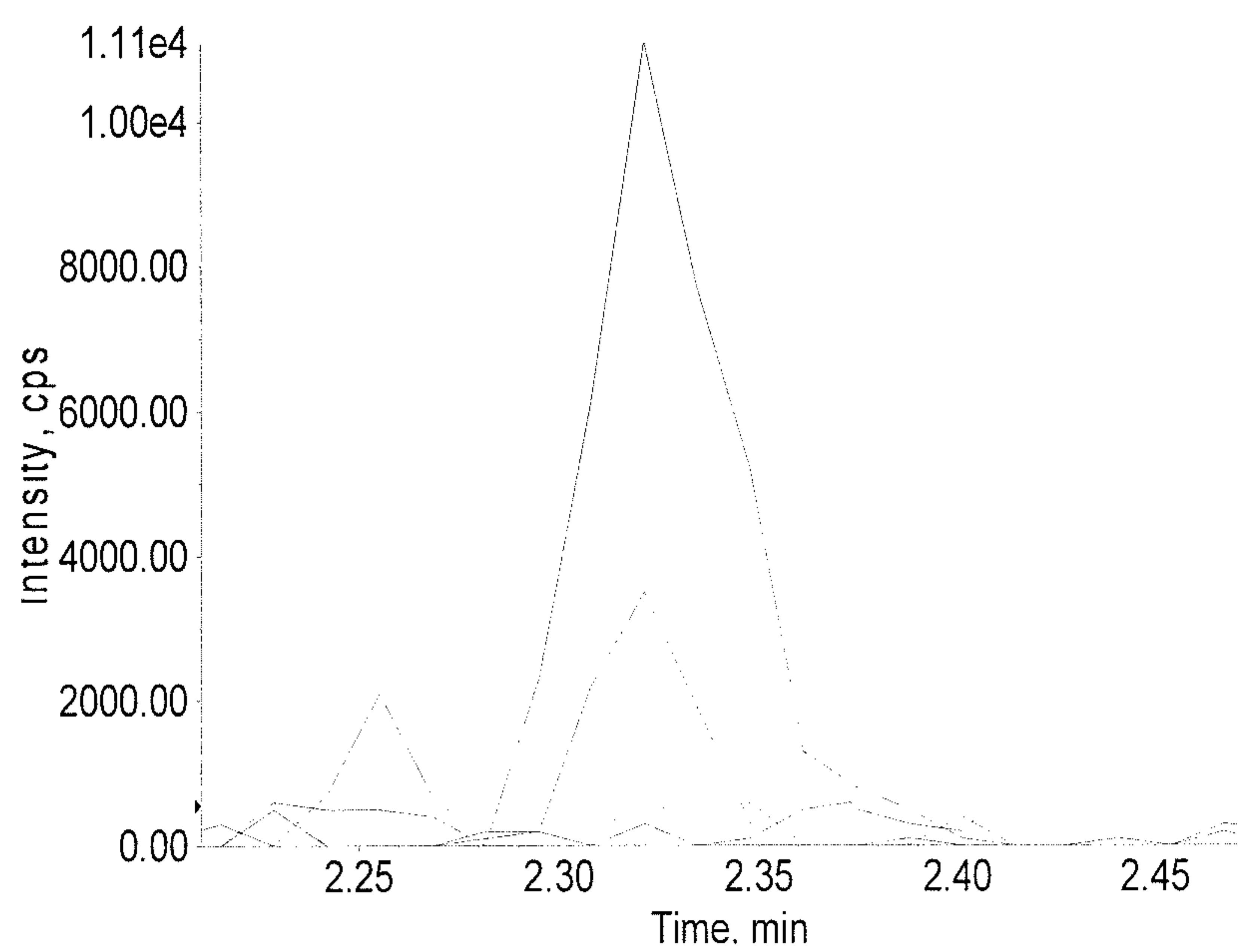
FIG. 6B

29/52



QLQGVNLTPEK

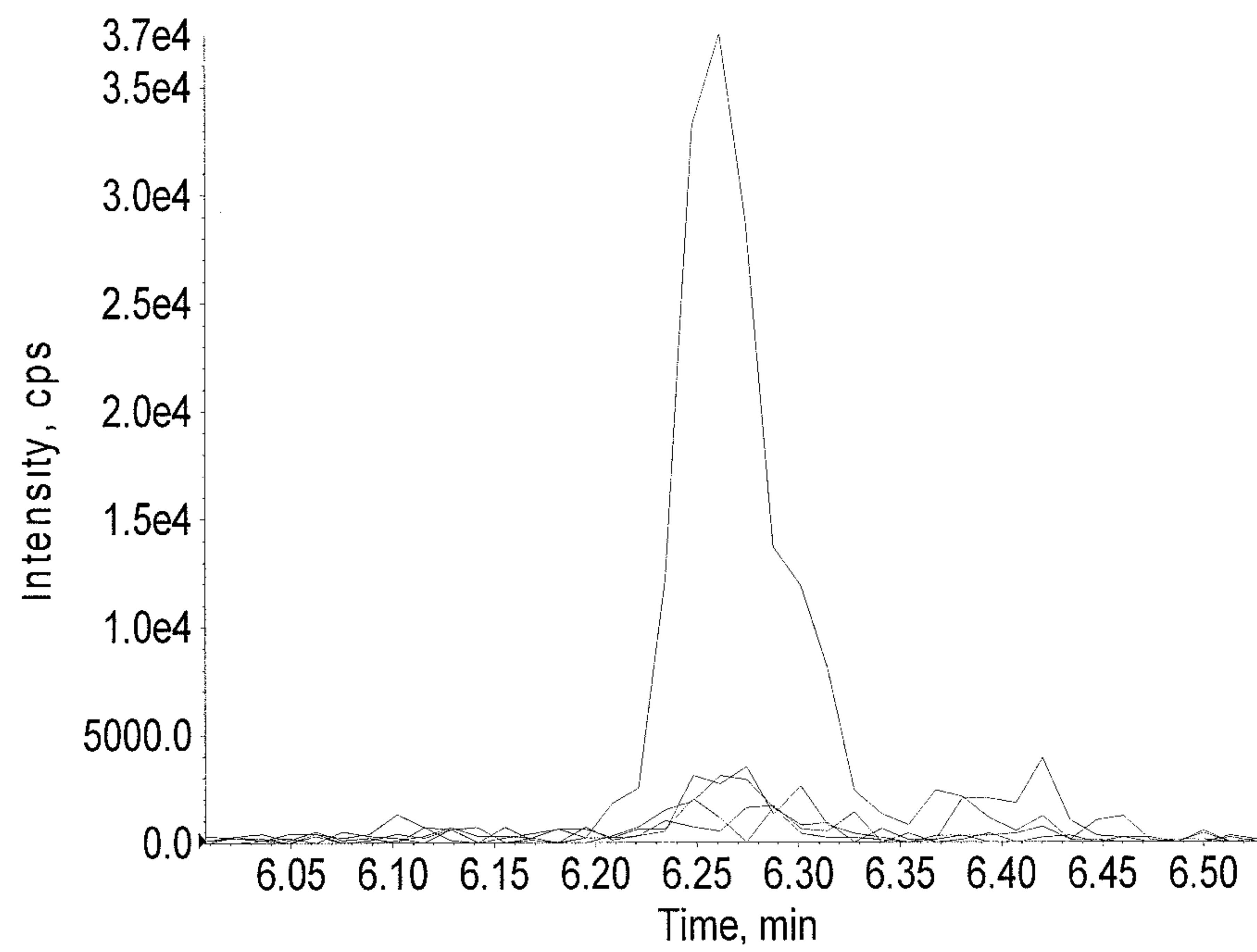
FIG. 6C



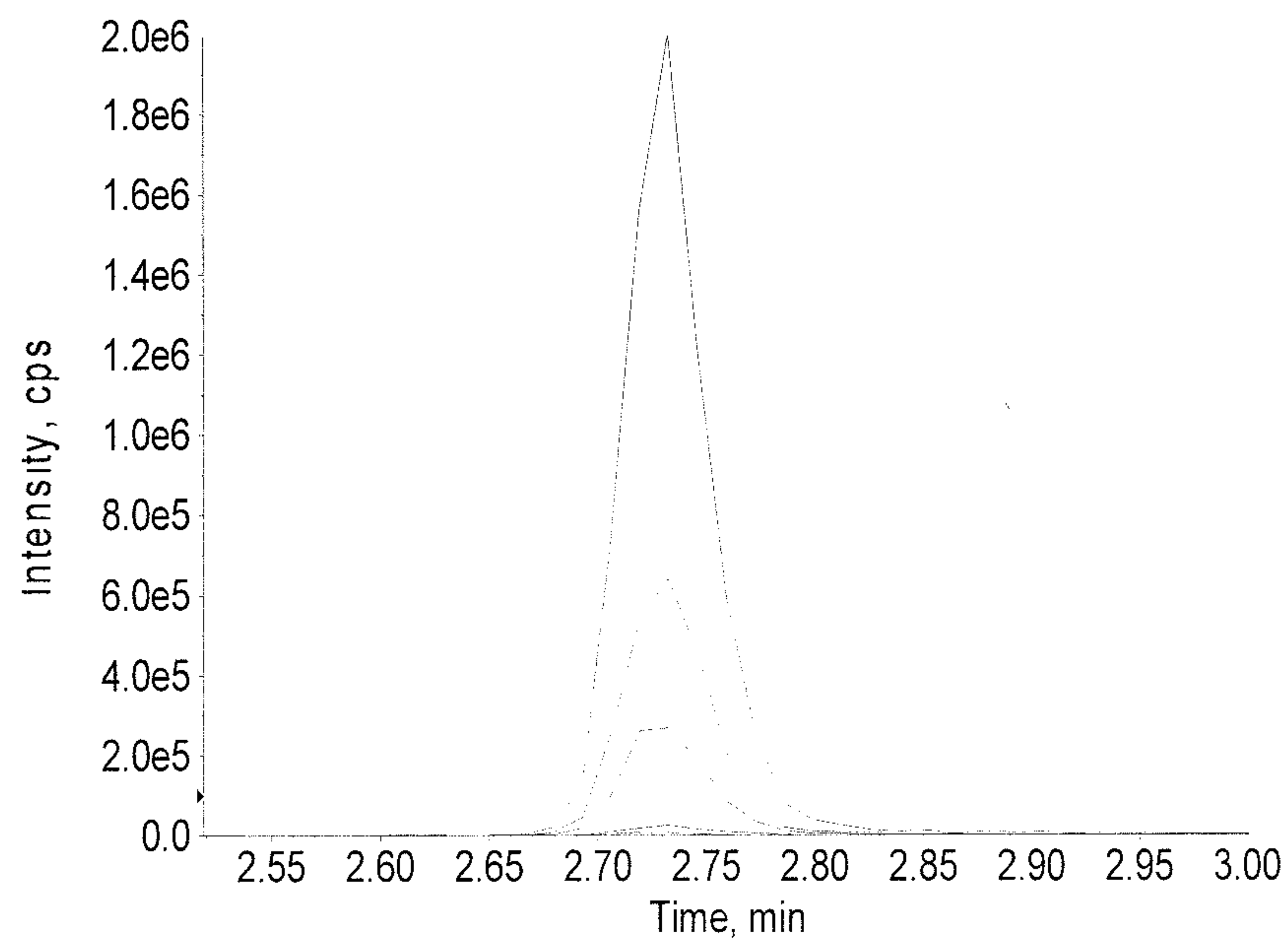
HIMEK

FIG. 6D

30/52



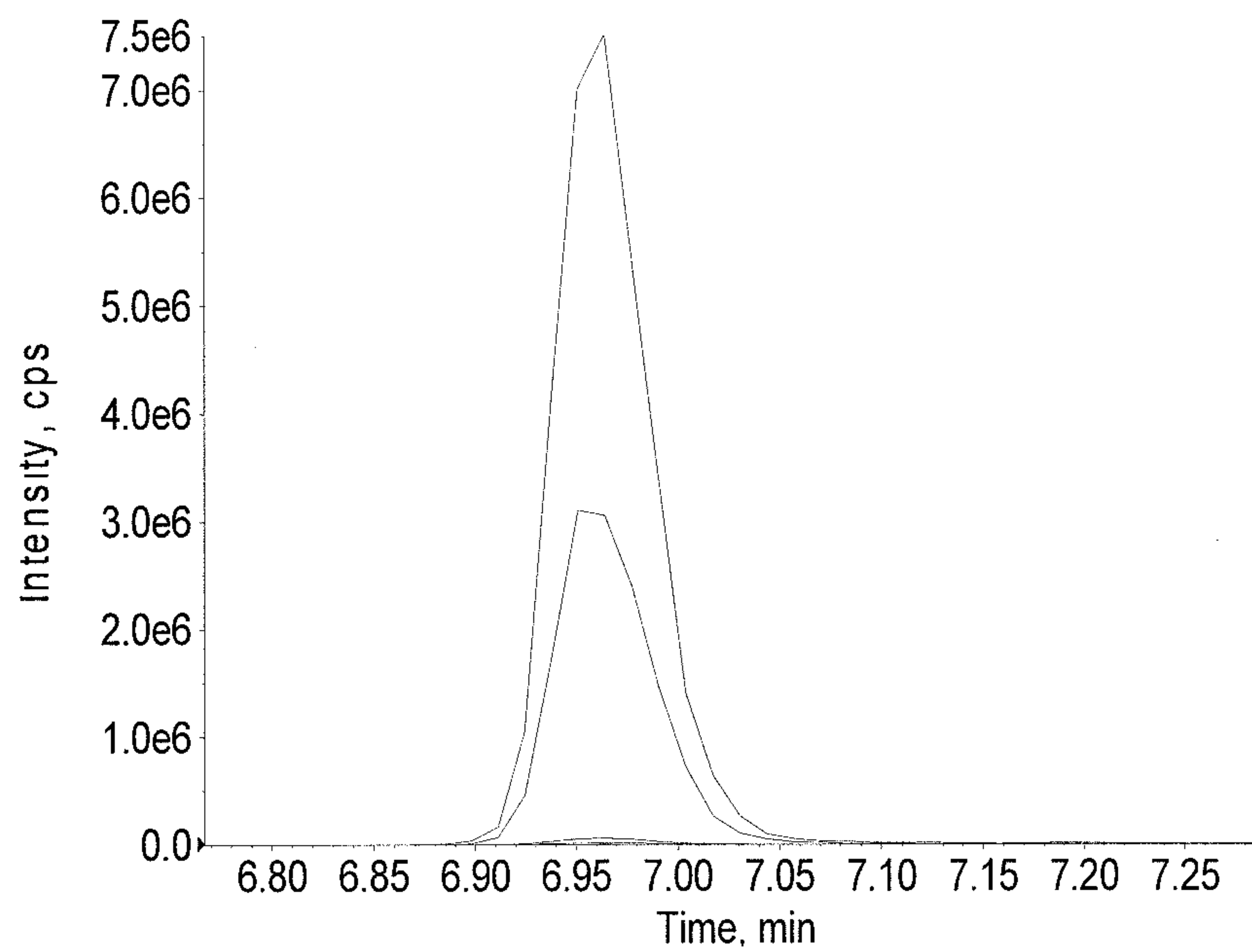
DEDEEEEGHMQK

FIG. 6E

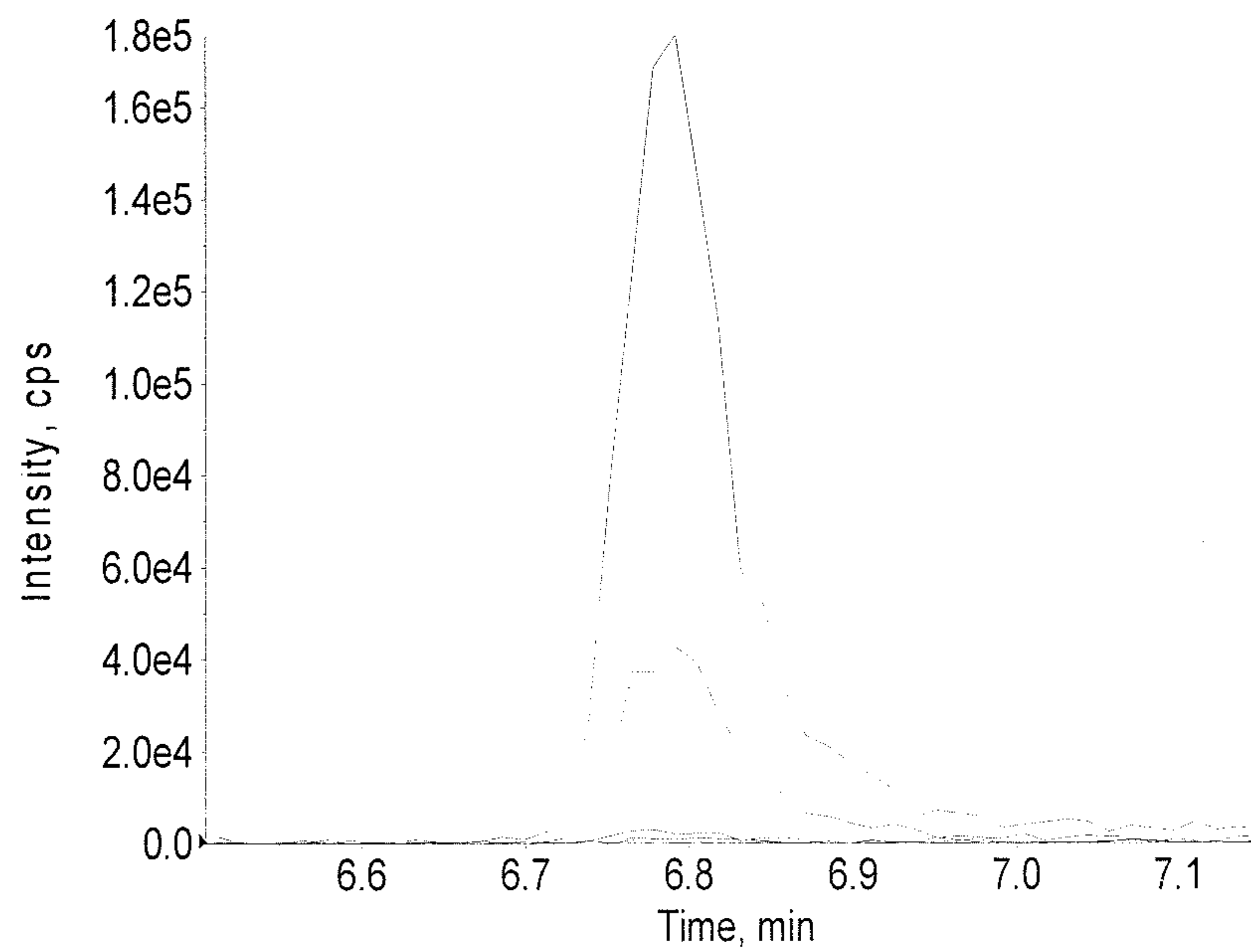
CCTEMSELR

FIG. 6F

31/52

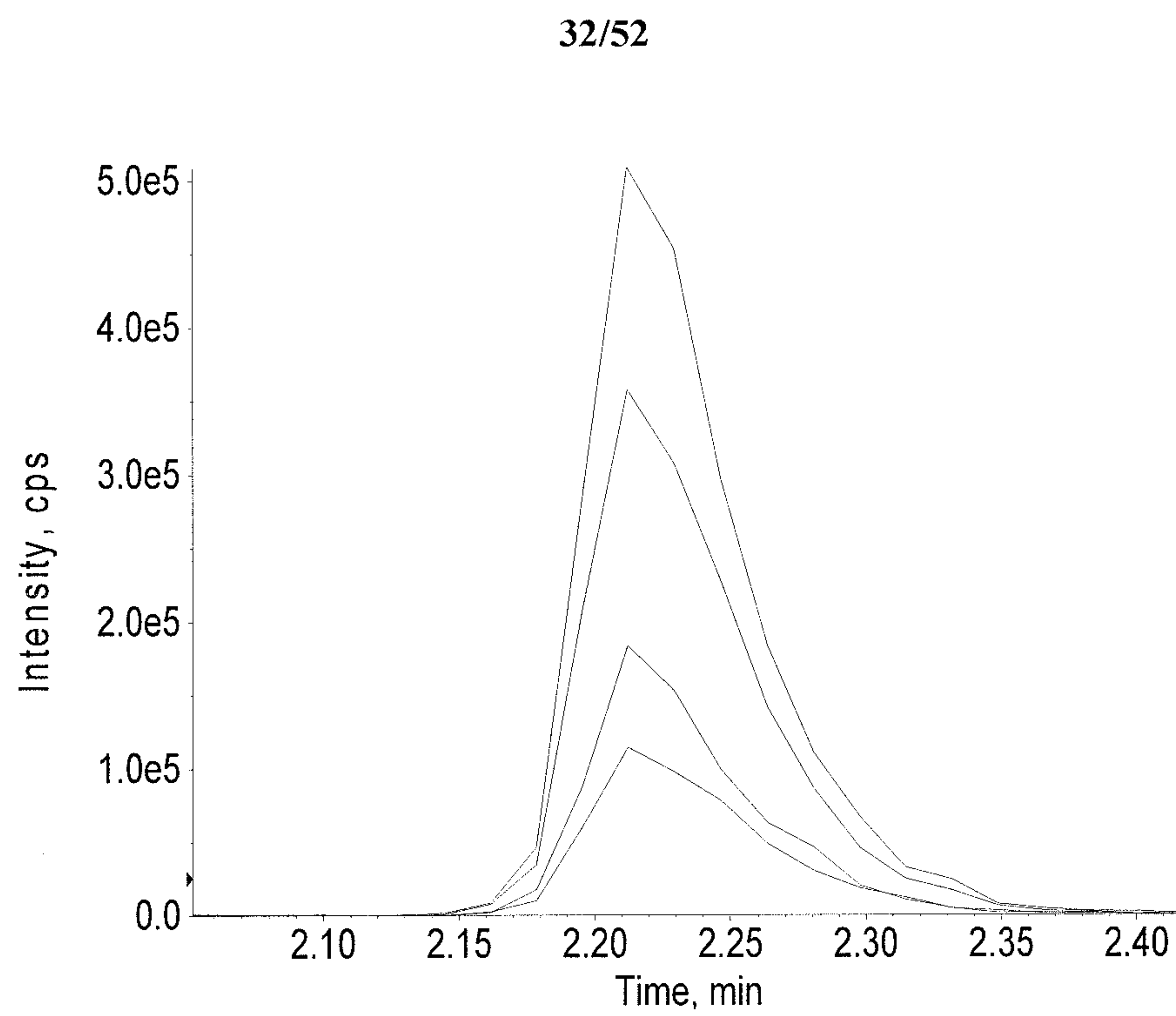


ELINLATMCR

FIG. 6G

FGPMIQCDLSSDD

FIG. 6H



IMENQSEELEEK

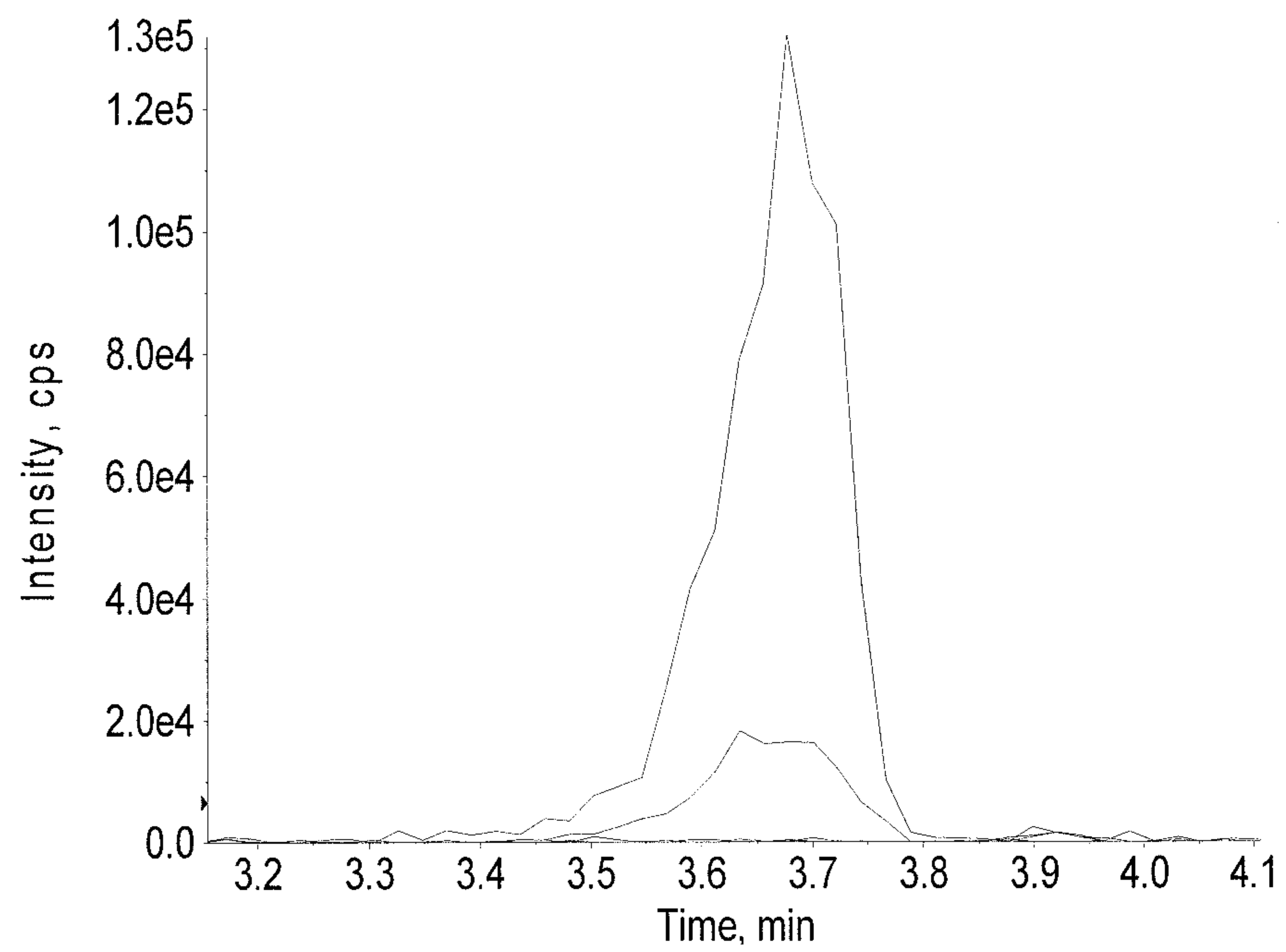
FIG. 6I

33/52

	1	50
2S albumin Eric NP_001238443	(1) -----MTKFTILLISLLFCIAHTCSASKWQHQQDSCRKQL	
2S albumin Glyma13g36400.1 BLAST	(1) MPPPSLHFTSPINSKMTKFTILLISLLFCIAHTCSASKWQHQQDSCRKQL	
2S albumin NP_001234950 BLAST	(1) -----MTKFTILLISLLFCIAHTCSASEWQHQQDSCRKQL	
2S albumin Glyma12g34160.1 BLAST	(1) -----MTKLTILLIALLFIAHTCCASKWQQHQQESCREQL	
Consensus	(1) MTKFTILLISLLFCIAHTCSASKWQHQQDSCRKQL	
	51	100
2S albumin Eric NP_001238443	(36) QGVNLTPECHEHIMEKIQGRGDDDDDDDDNHILRTMRGRINYIRRNEGKD	
2S albumin Glyma13g36400.1 BLAST	(51) QGVNLTPECHEHIMEKIQGRGDDDDDDDDNHILRTMRGRINYIRRNEGKD	
2S albumin NP_001234950 BLAST	(36) QGVNLTPECHEHIMEKIQGRGDDDDDDDDNHILRTMRGRINYIRRNEGKD	
2S albumin Glyma12g34160.1 BLAST	(36) KGINLNPCE-HIMEKIQAGRRGEDGSDEDHILRTMPGRINYIRKKEGKE	
Consensus	(51) QGVNLTPECHEHIMEKIQGRGDDDDDDDDNHILRTMRGRINYIRRNEGKD	
	101	150
2S albumin Eric NP_001238443	(86) EDEEEEGHMQKCCTEMSELRSKPCQCKALQKIMENQSEELEEKQKKKMEK	
2S albumin Glyma13g36400.1 BLAST	(101) EDEEEEGHMQKCCTEMSELRSKPCQCKALQKIMENQSEELEEKQKKKMEK	
2S albumin NP_001234950 BLAST	(86) EDEEEEGHMQKCRTEMSELRSKPCQCKALQKIMENQSEELEEKQKKKMEK	
2S albumin Glyma12g34160.1 BLAST	(85) EEE--EGHMQKCCSEMSELKSPICQCKALQKIMDNQSEQLEGKEKKQMER	
Consensus	(101) EDEEEEGHMQKCCTEMSELRSKPCQCKALQKIMENQSEELEEKQKKKMEK	
	151	174
2S albumin Eric NP_001238443	(136) ELINLATMCRFGPMIQCDLSSDD-	
2S albumin Glyma13g36400.1 BLAST	(151) ELINLATMCRFGPMIQCDLSSDD-	
2S albumin NP_001234950 BLAST	(136) ELINLATMCRFGPMIQCDLSSDD-	
2S albumin Glyma12g34160.1 BLAST	(133) ELMNLAIRCLGPMIGCDLSSDD-	
Consensus	(151) ELINLATMCRFGPMIQCDLSSDD	

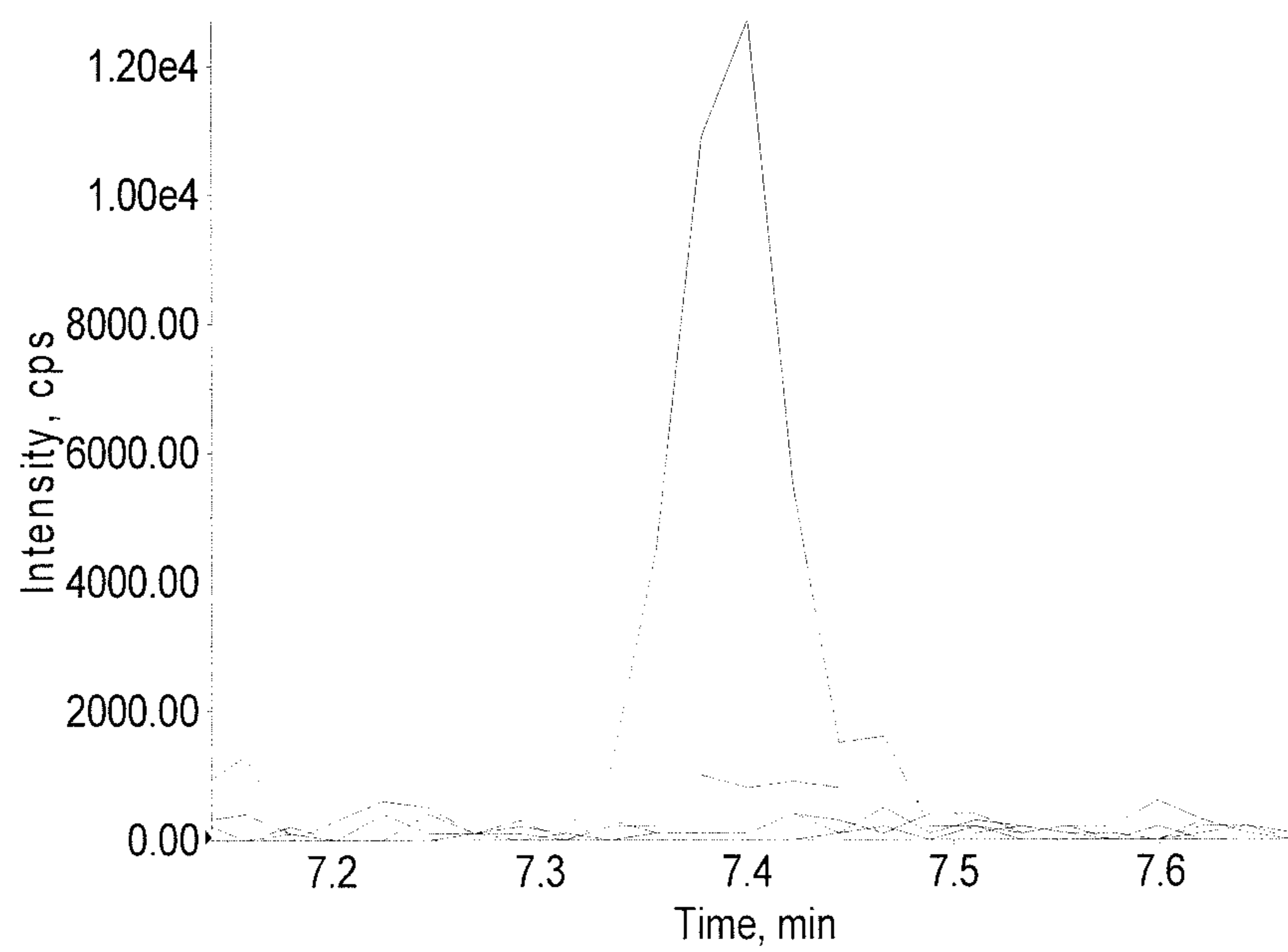
FIG. 6J

34/52



VFSPNK

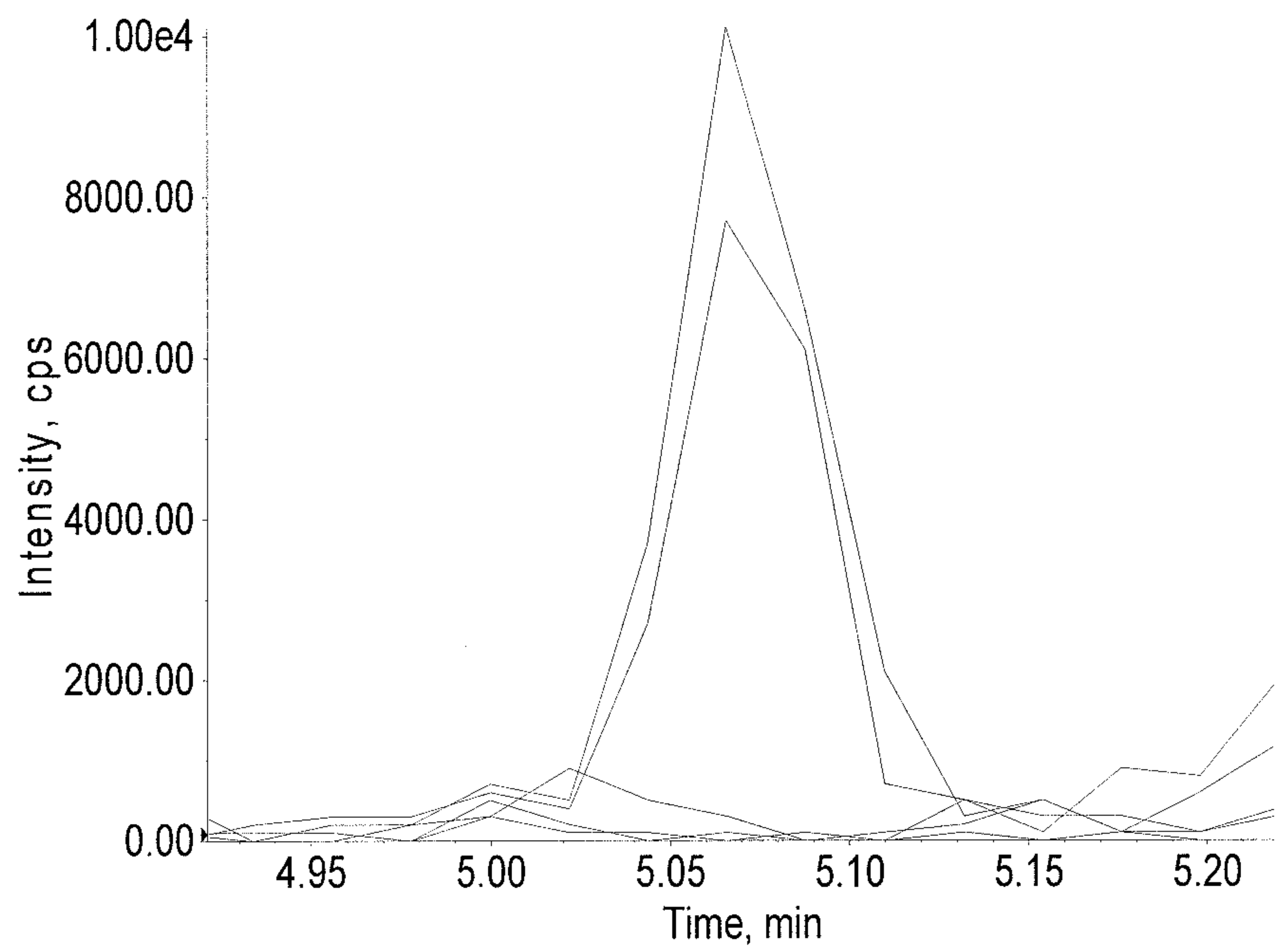
FIG. 7A



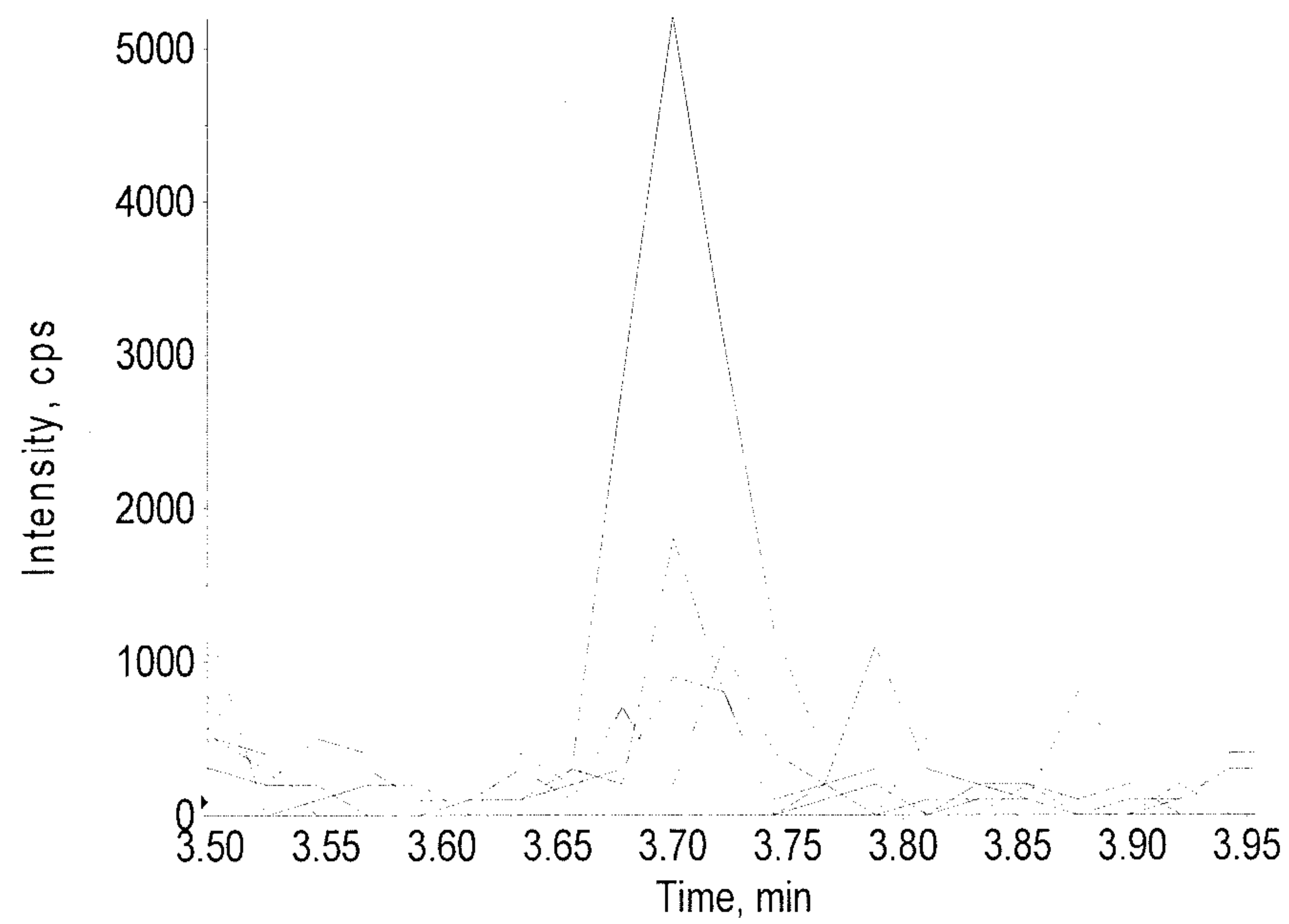
ANSTNTVSFTVSK

FIG. 7B

35/52

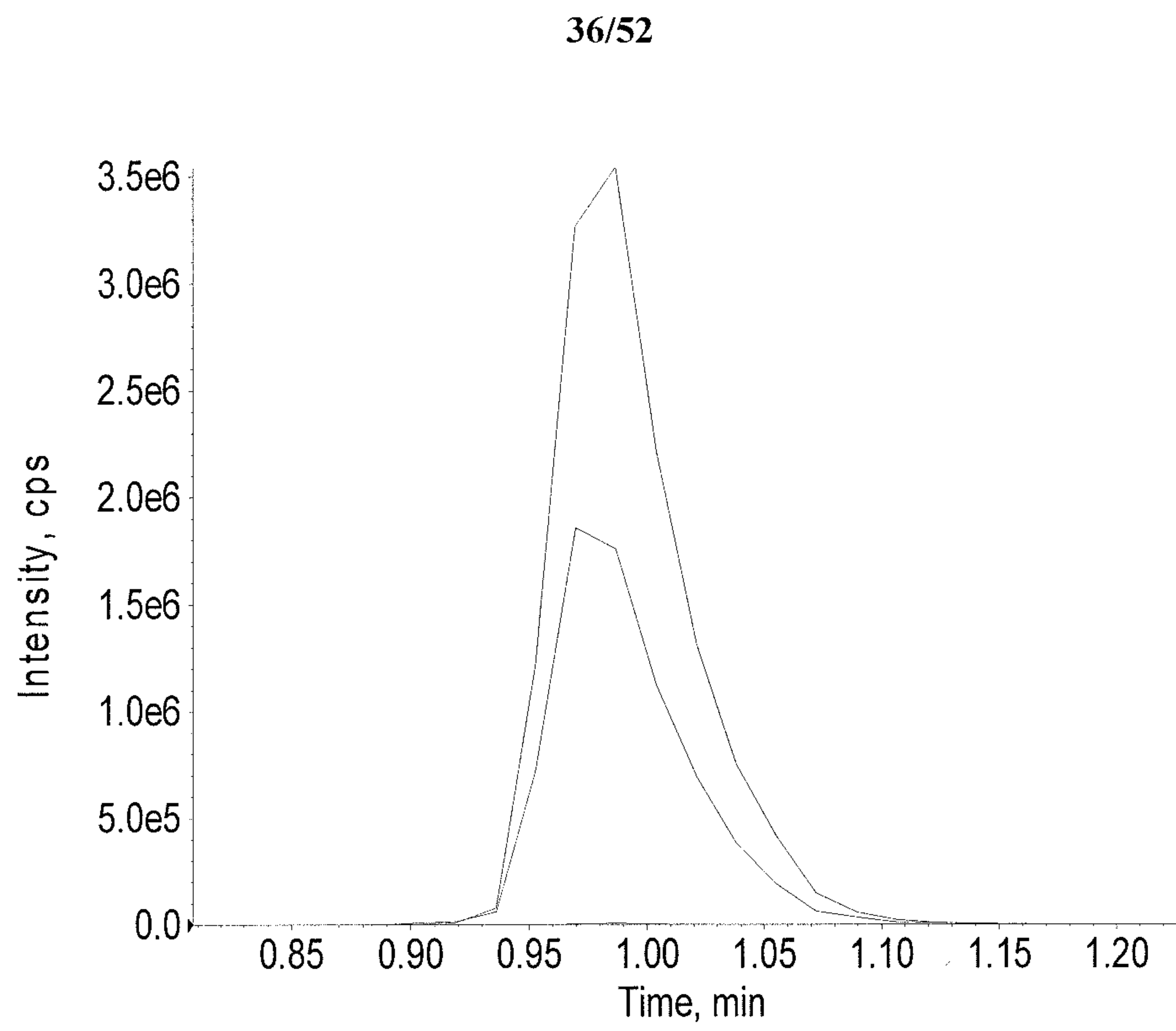


QQNLIFQGDAISP SGVLR

FIG. 7C

TADGLAFFLAPVGSKPQSK

FIG. 7D



VFSPNK

FIG. 7E

37/52

	1	50
Lectin Eric ADC94422	(1) -----MATS NFS-----IVLSLSLAFFLVLLTKANSTNT	
Lectin XP_003535884 BLAST	(1) -----MATS NFS-----IVLSLSLALFLMLLTKANSTNT	
Lectin Glyma10g15480.1 BLAST	(1) -----MATS NFS-----IVLSLSLALFLMLLTKANSTNT	
Lectin NP_001237210 BLAST	(1) -----MATS NFS-----IVLSVSLAFFLVLLTKAHSTDT	
Lectin Glyma02g18090.1 BLAST	(1) MKVLCIIFEFKQIKAMATS NFS-----IVLSVSLAFFLVLLTKAHSTDT	
Lectin ACU23599 BLAST	(1) -----MATS NFS-----IVLSVSLAFFLVLLTKAHPTDT	
Lectin CAH60173 BLAST	(1) -----MASSKFS-----TVISFSLALFLVLLTQANSTNI	
Lectin Glyma10g01620.2 BLAST	(1) -----MATSKFHTQKPLFVVLS--VVVVLLTMTKVNSTKP	
Lectin Glyma02g01590.1 BLAST	(1) -----MATSKLKTONVVVSLSLTLTLVLVLLTSKANS AET	
Consensus	(1) MATS NFS IVLSLSLAFFLVLLTKANSTNT	
	51	100
Lectin Eric ADC94422	(30) VSFTVSKFSPRQONLIFQGDAAISPSGVLRLTKVDSIDVPTTGS LGRALY	
Lectin XP_003535884 BLAST	(30) VSFTTSKFSPRQONLILQGDAAISPSGVLRLTKVDSYGVPTSRSLGRALY	
Lectin Glyma10g15480.1 BLAST	(30) VSFTTSKFSPRQONLILQGDAAISPSGVLRLTKVDSYGVPTSRSLGRALY	
Lectin NP_001237210 BLAST	(30) VSFTFNKFNPVQPNIMLQKDASISSSGVLQLTKVGSNGVPTSGSLGRALY	
Lectin Glyma02g18090.1 BLAST	(45) VSFTFNKFNPVQPNIMLQKDASISSSGVLQLTKVGSNGVPTSGSLGRALY	
Lectin ACU23599 BLAST	(30) VSFTFNKFNPVQPNIMLQKDASISSSGVLQLTKVGSNGVPTSGSLGRALY	
Lectin CAH60173 BLAST	(30) FSENFQTFDS--PNLIFQGDASVSSSGQLRLTKVKGNKGKPTAASLGRAFY	
Lectin Glyma10g01620.2 BLAST	(34) FLSPGTSSCRTNRTLILQGDALVTSSR-----KSLGRALY	
Lectin Glyma02g01590.1 BLAST	(36) VSFSWNKFVPKQPNMILQGD AIVTSSGKLQLNKVDENGTPKPSLGRALY	
Consensus	(51) VSFTFSKFSPRQPNLILQGDAAISSSGVLRLTKVDSNGVPTSGSLGRALY	
	101	150
Lectin Eric ADC94422	(80) ATPIQIWDSETGKVASWATSEKFKVFSPNK--TADGLAFFLAPVGSKPQS	
Lectin XP_003535884 BLAST	(80) AAPIQIWDSETGKVASWATSEKFNVFSPDK--TADGLAFFLAPVGSKPQY	
Lectin Glyma10g15480.1 BLAST	(80) AAPIQIWDSETGKVASWATSEKFNVFSPDK--TADGLAFFLAPVGSKPQY	
Lectin NP_001237210 BLAST	(80) AAPIQIWDSETGKVASWATSEKFNIFAPNKSNSADGLAFFLAPVGSQPQS	
Lectin Glyma02g18090.1 BLAST	(95) AAPIQIWDSETGKVASWATSEKFNIFAPNKSNSADGLAFFLAPVGSQPQS	
Lectin ACU23599 BLAST	(80) AAPIQIWDSETGKVASWATSEKFNIFAPNKSNSADGLAFFLAPVGSQPQS	
Lectin CAH60173 BLAST	(78) SAPIQIWDSTTGNVASFATSFTFNILAPNKSNSADGLAFALVPVGSQPKS	
Lectin Glyma10g01620.2 BLAST	(69) STPIHIWDSEIGSVASFAASENFVYASDIANLADGLAFFLAPIDTQPQT	
Lectin Glyma02g01590.1 BLAST	(86) STPIHIWDKETGSVASFAASENFVYAPDTRKRLADGLAFFLAPIDTKPQT	
Consensus	(101) AAPIQIWDSETGKVASWATSEKFNVFAPNK TADGLAFFLAPVGSKPQS	
	151	200
Lectin Eric ADC94422	(128) KGGFLGLFNSDSKNKSVQTVAVEFDITYYNAKWDPANRHIGIDVNSIKSVK	
Lectin XP_003535884 BLAST	(128) KAGFLGLFNSDSKNMSLQTVAVEFDITYYNQKWDPASRHIGIDVNSIKSVK	
Lectin Glyma10g15480.1 BLAST	(128) KAGFLGLFNSDSKNMSLQTV A-----WDPASRHIGIDVNSIKSVK	
Lectin NP_001237210 BLAST	(130) DDGFLGLFNSPLKDKSLQTV AIEFDTF SNKKWDPANRHIGIDVNSIKSVK	
Lectin Glyma02g18090.1 BLAST	(145) DDGFLGLFNSPLKDKSLQTV AIEFDTF SNKKWDPANRHIGIDVNSIKSVK	
Lectin ACU23599 BLAST	(130) DDGFLGLFNSPLKDKSLQTV AIEFDTF SNKKWDPANRHIGIDVDSIKSIK	
Lectin CAH60173 BLAST	(128) NGGFLGLFDNATYDSSAQTVAVEFDITYSNPKWDPENRHIGIDVNSIESIR	
Lectin Glyma10g01620.2 BLAST	(119) RGGYLGLYN-----N-----PSN-----	
Lectin Glyma02g01590.1 BLAST	(136) HAGYLGLFENEN--ESGDQVVAVEFDTF-RNSWDFPNPHIGINVNSIRSIK	
Consensus	(151) KGGFLGLFNSDSKDKSLQTVAVEFDITYSNKKWDPANRHIGIDVNSIKSVK	

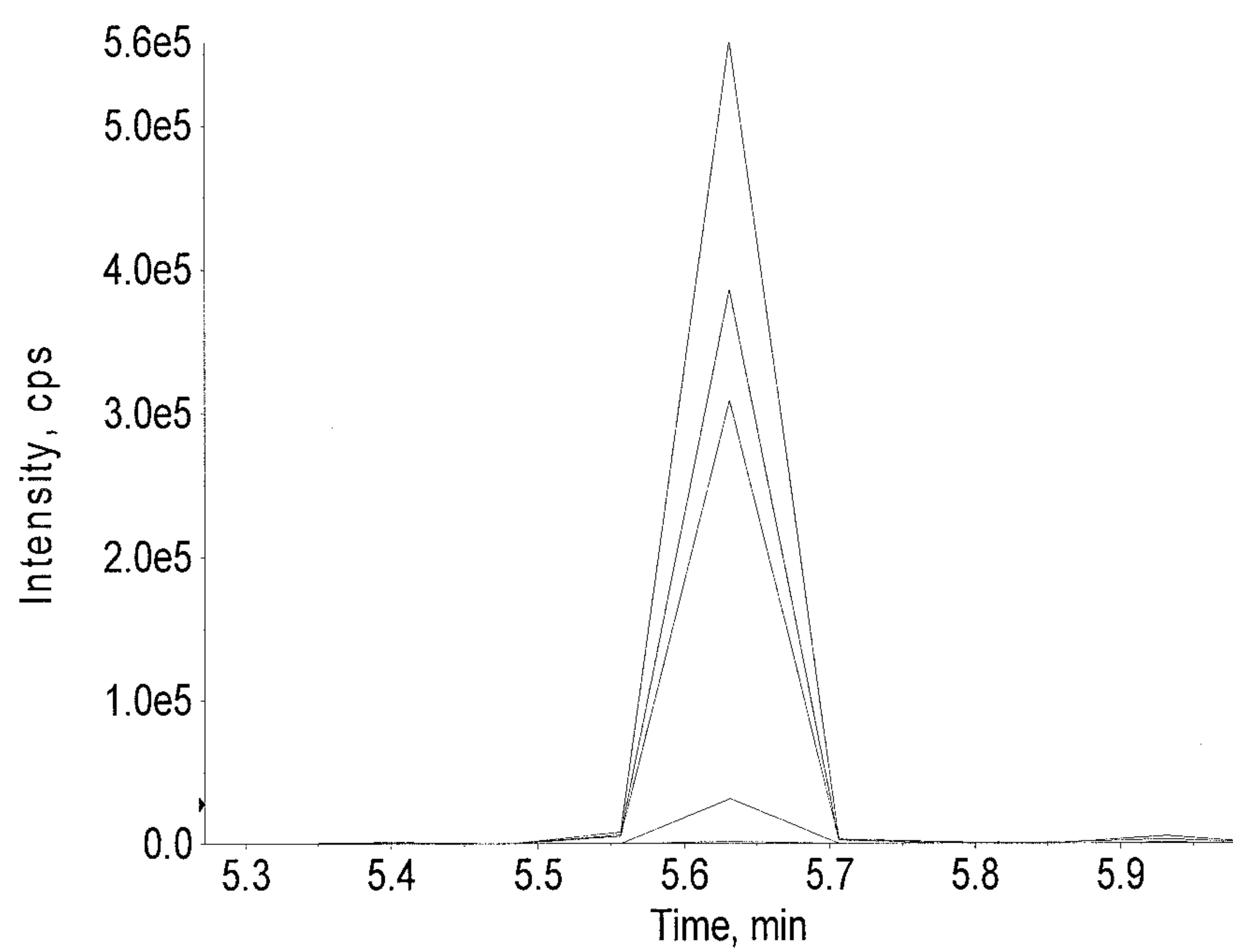
FIG. 7F

38/52

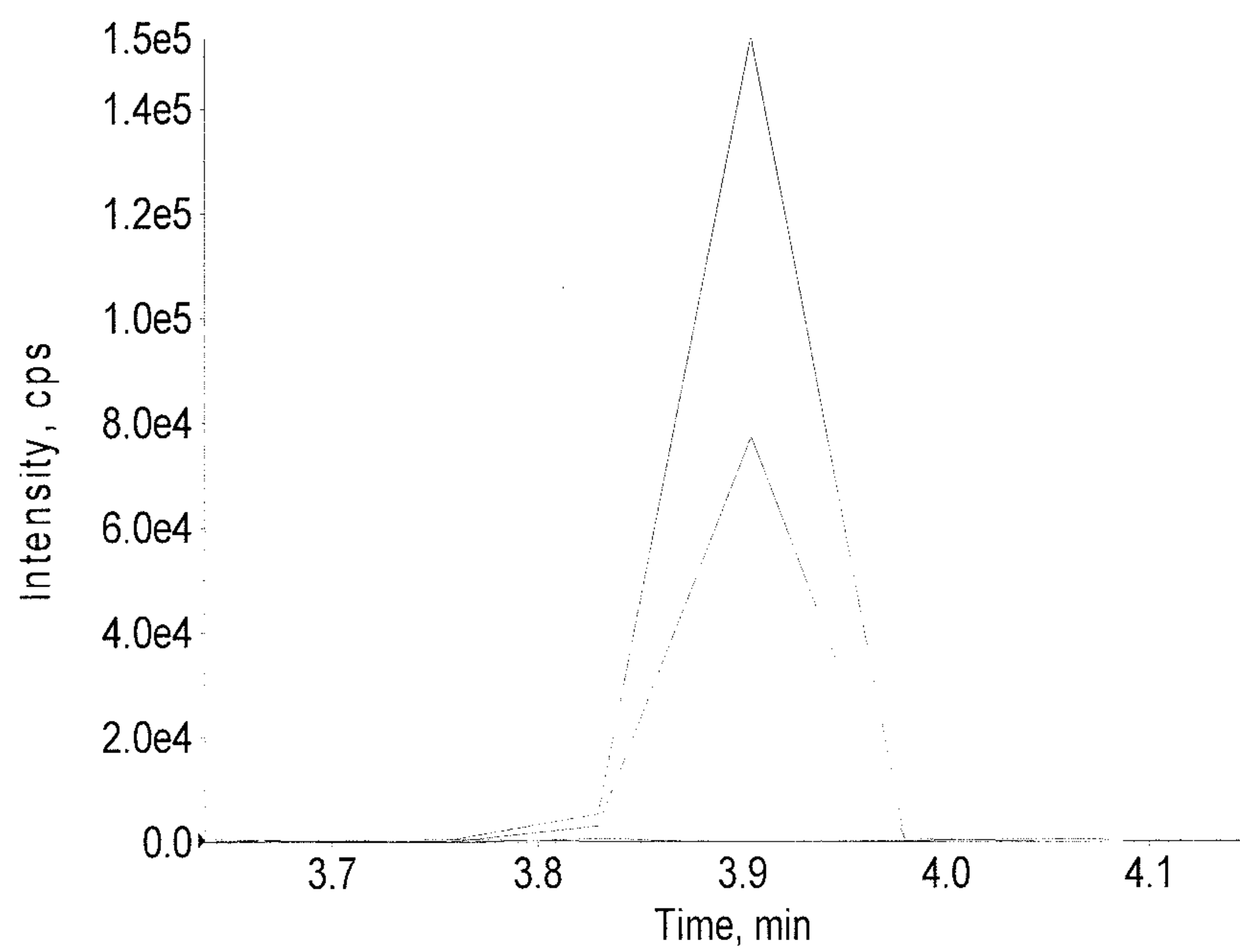
	201	250
Lectin Eric ADC94422	(178) TASWGLANGQIAQILITYDADTSLLVASLIHPSRKTSYILSETVSLKSNL	
Lectin XP_003535884 BLAST	(178) TAPWGFANGQVAQILITYNADTSLLVASLVHPSRKTSYILSETVSLKSNL	
Lectin Glyma10g15480.1 BLAST	(168) TAPWGFANGQVAQILITYNADTSLLVASLVHPSRKTSYILSETVSLKSNL	
Lectin NP_001237210 BLAST	(180) TASWGLSNGQVAEILVTYNAATSLLVASLIHPSKKT SYILSDTVNLKSNL	
Lectin Glyma02g18090.1 BLAST	(195) TASWGLSNGQVAEILVTYNAATSLLVASLIHPSKKT SYILSDTVNLKSNL	
Lectin ACU23599 BLAST	(180) TASWGLSNGQVAEILVTYNAATSLLVASLIHPSKKT SYILSDTVNLKSNL	
Lectin CAH60173 BLAST	(178) TASWGLANGQNAEILITYDSSTKLLVASLVHPSRRTSYIVSERVDLKSVL	
Lectin Glyma10g01620.2 BLAST	(132) -SSWGLANDQVTNVLITYDASTNLLVASLVHPSQRSSYILSDVLDLKVAL	
Lectin Glyma02g01590.1 BLAST	(183) TTSWDLANNKVAKVLITYDASTSLLVASLVYPSQRTSNILSDVVDLKTSL	
Consensus	(201) TASWGLANGQVAQILITYDAATSLLVASLIHPSRKTSYILSETVSLKSNL	
	251	300
Lectin Eric ADC94422	(228) PEWVNIGFSATTGLNKGFEVETHDVFSWSFASKLSDGST-SDTLDLPSFLL	
Lectin XP_003535884 BLAST	(228) PEWVNIGFSATTGANKGFAETHDVFSWSFASKLSDGST-SDTLDLASFLL	
Lectin Glyma10g15480.1 BLAST	(218) PEWVNIGFSATTGANKGFAETHDVFSWSFASKLSDGST-SDTLDLASFLL	
Lectin NP_001237210 BLAST	(230) PEWVSVGFSAATTGLHEGSEVETHDVISWSFASKLSDGSS-NDALDLPSFVL	
Lectin Glyma02g18090.1 BLAST	(245) PEWVSVGFSAATTGLHEGSEVETHDVISWSFASKLSDGSS-NDALDLPSFVL	
Lectin ACU23599 BLAST	(230) PEWVSVGFSAATTGLHEGSEVETHDVISWSFASKLSDGSS-NDALDLPSFVL	
Lectin CAH60173 BLAST	(228) PEWVSIGFSATTGLLEGSIETHDVLWSFASKLSDDTT-SEGLNLANFVL	
Lectin Glyma10g01620.2 BLAST	(181) PEWVRIGFSATTGLNVA-SETHDVHWSFSSNLPFGSSNTNPSDFAI-FI-	
Lectin Glyma02g01590.1 BLAST	(233) PEWVRIGFSAATGLDIP-GESHDVLSWSFASNLPHASSNIDPLDLTSEVL	
Consensus	(251) PEWVSIGFSATTGLNEGSEVETHDVISWSFASKLSDGST SDALDLPSFLL	
	301	
Lectin Eric ADC94422	(277) NEAI-	
Lectin XP_003535884 BLAST	(277) NEAI-	
Lectin Glyma10g15480.1 BLAST	(267) NEAI-	
Lectin NP_001237210 BLAST	(279) NEAI-	
Lectin Glyma02g18090.1 BLAST	(294) NEAI-	
Lectin ACU23599 BLAST	(279) NEAI-	
Lectin CAH60173 BLAST	(277) NKIL-	
Lectin Glyma10g01620.2 BLAST	(229) -----	
Lectin Glyma02g01590.1 BLAST	(282) HEAI-	
Consensus	(301) NEAI	

FIG. 7F - continued

39/52



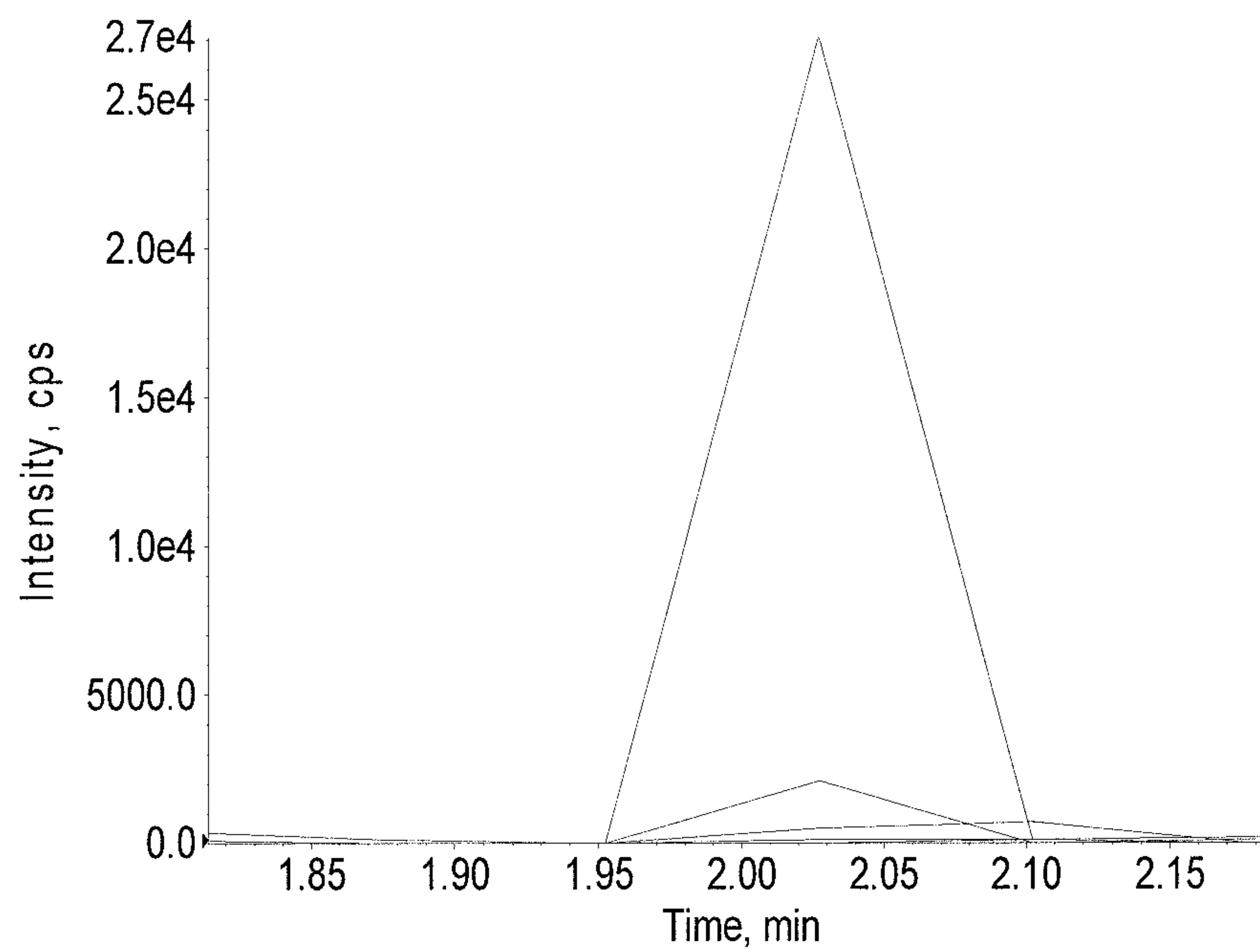
SSDFLTYGIK

FIG. 8A

GTVVLMPK

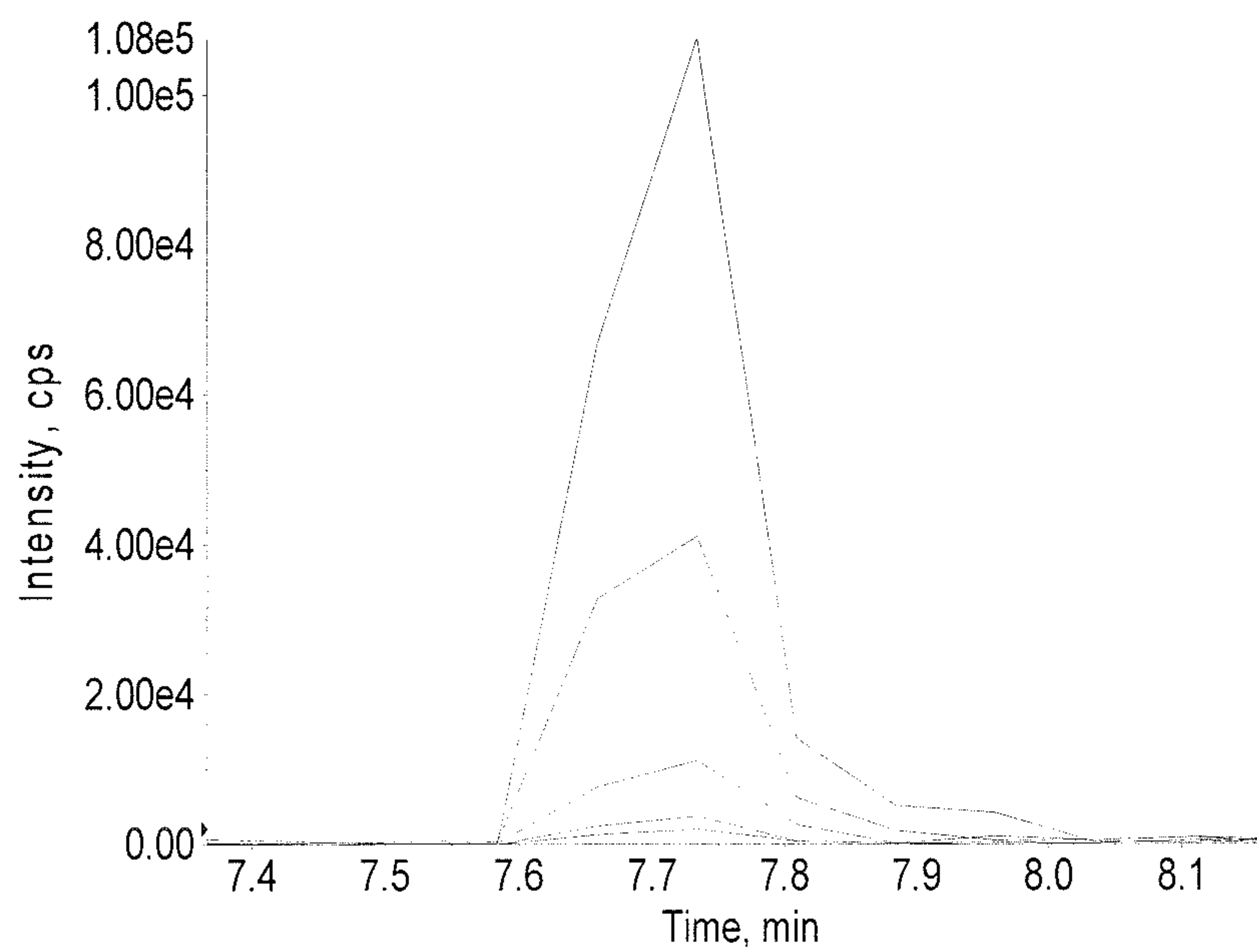
FIG. 8B

40/52



NVLDFNAITSIGK

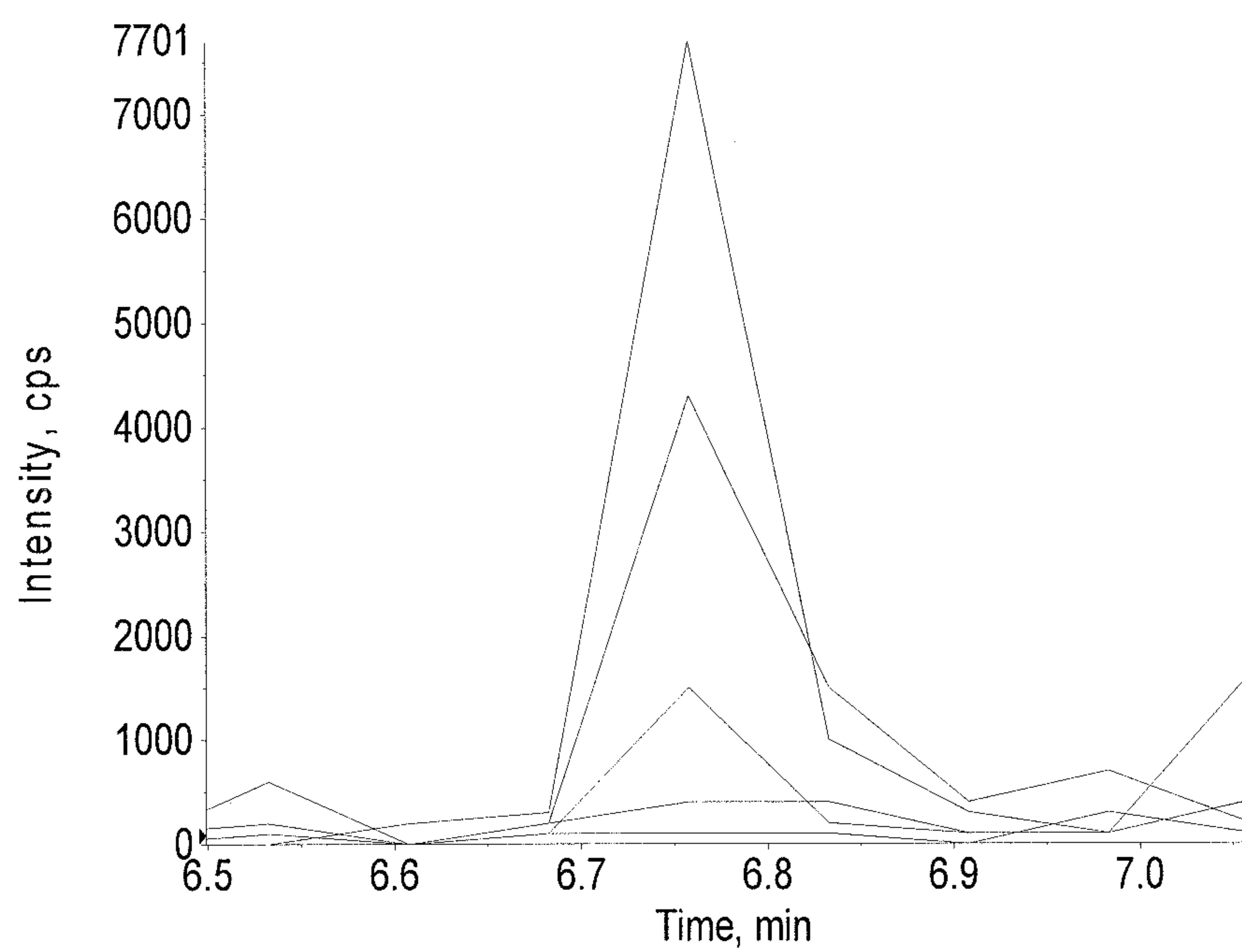
FIG. 8C



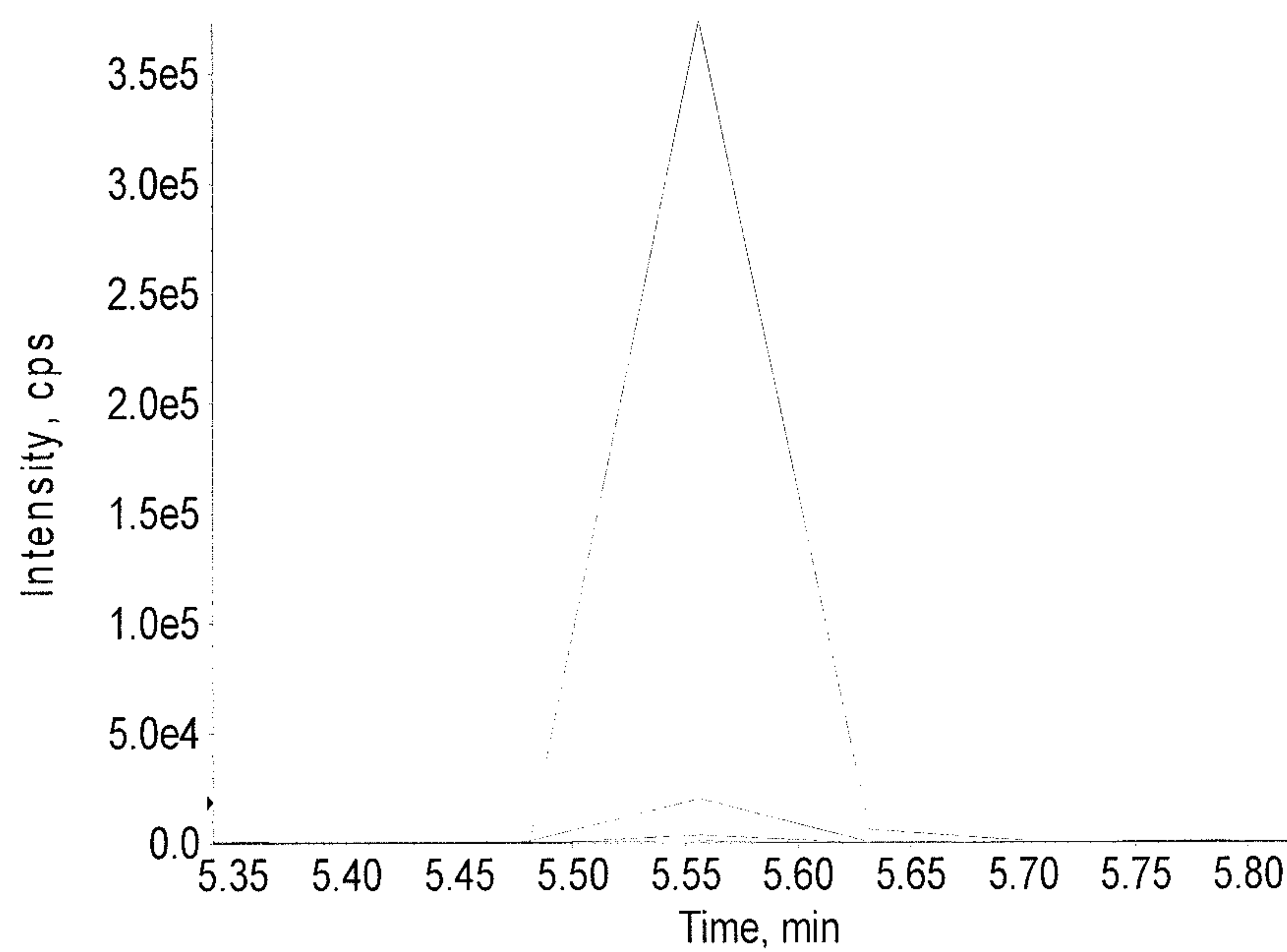
GGVIDTATGILGQGVSLVGGVIDTATSFLGR

FIG. 8D

41/52

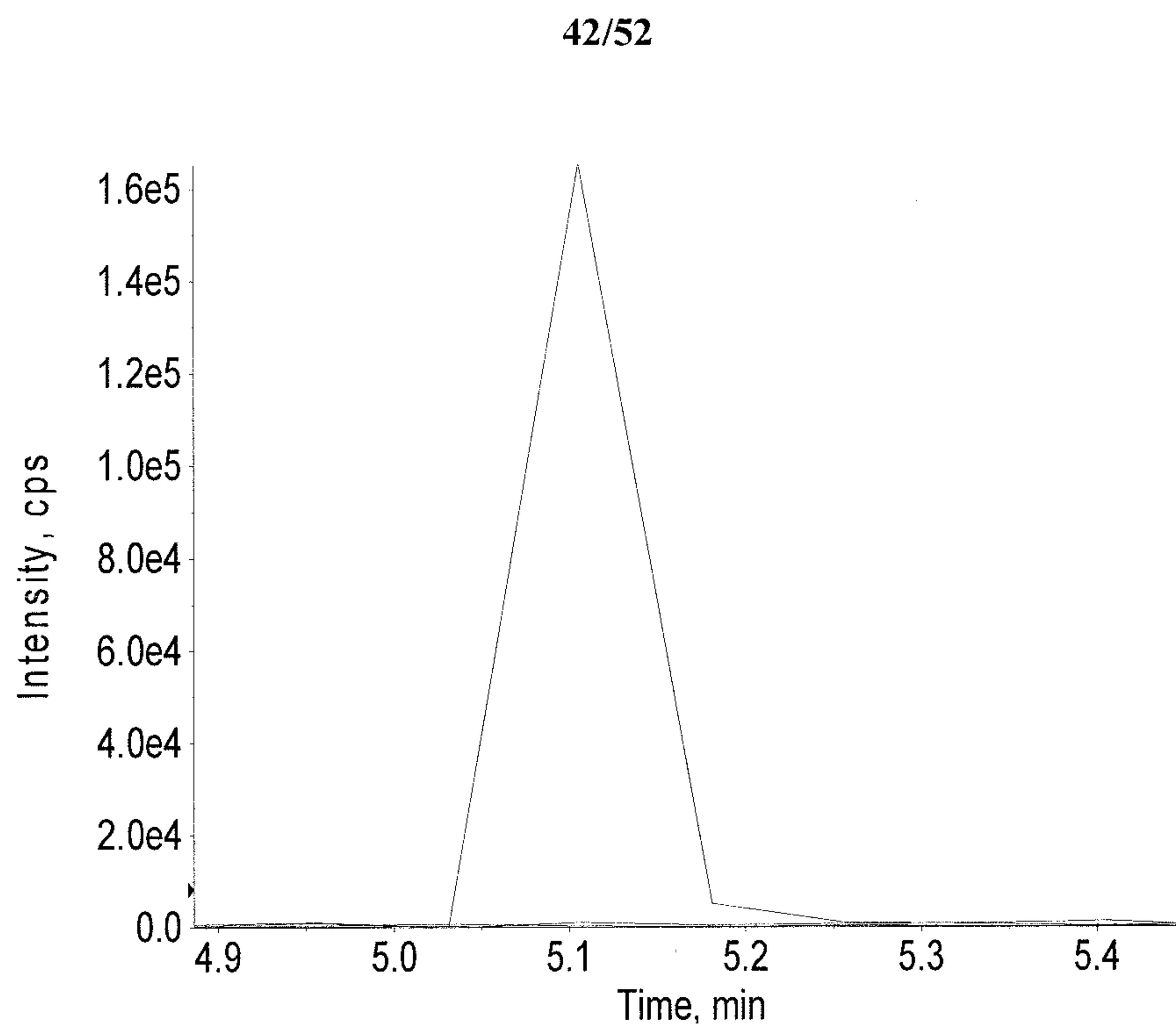


IFFVNDTYLPSATPAPLLK

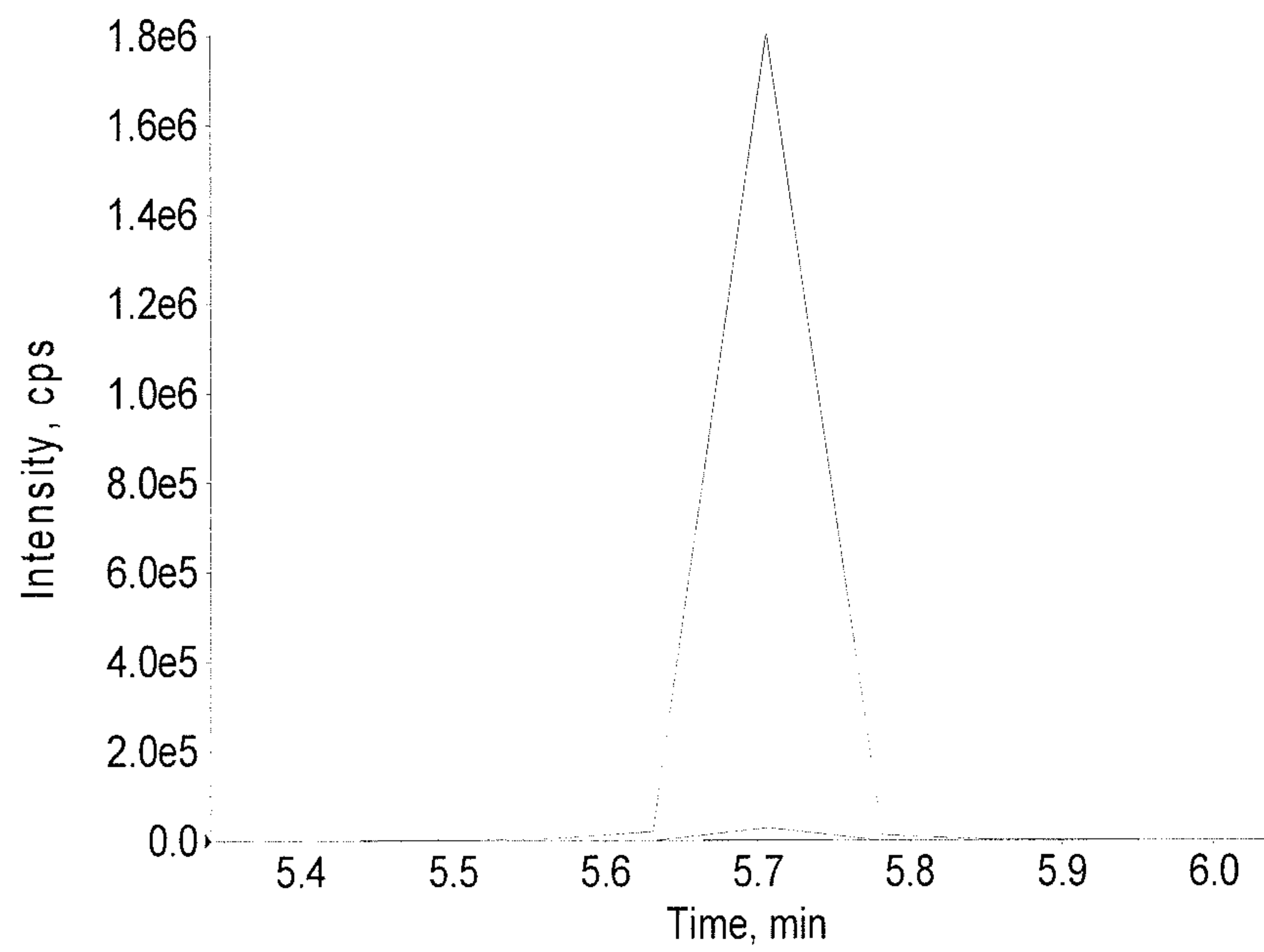
FIG. 8E

DENFGHLK

FIG. 8F

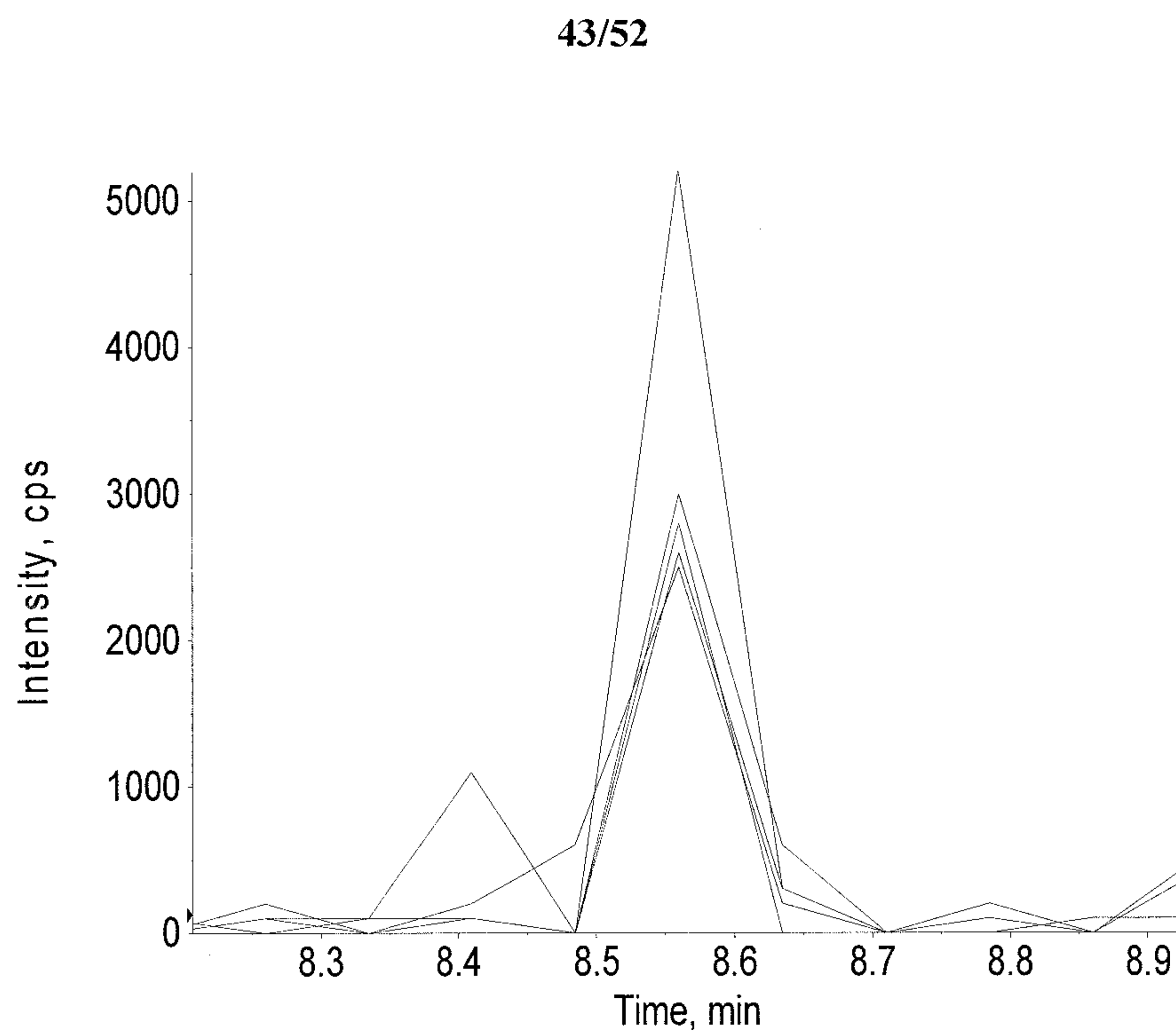


SLSHDVIPLFK

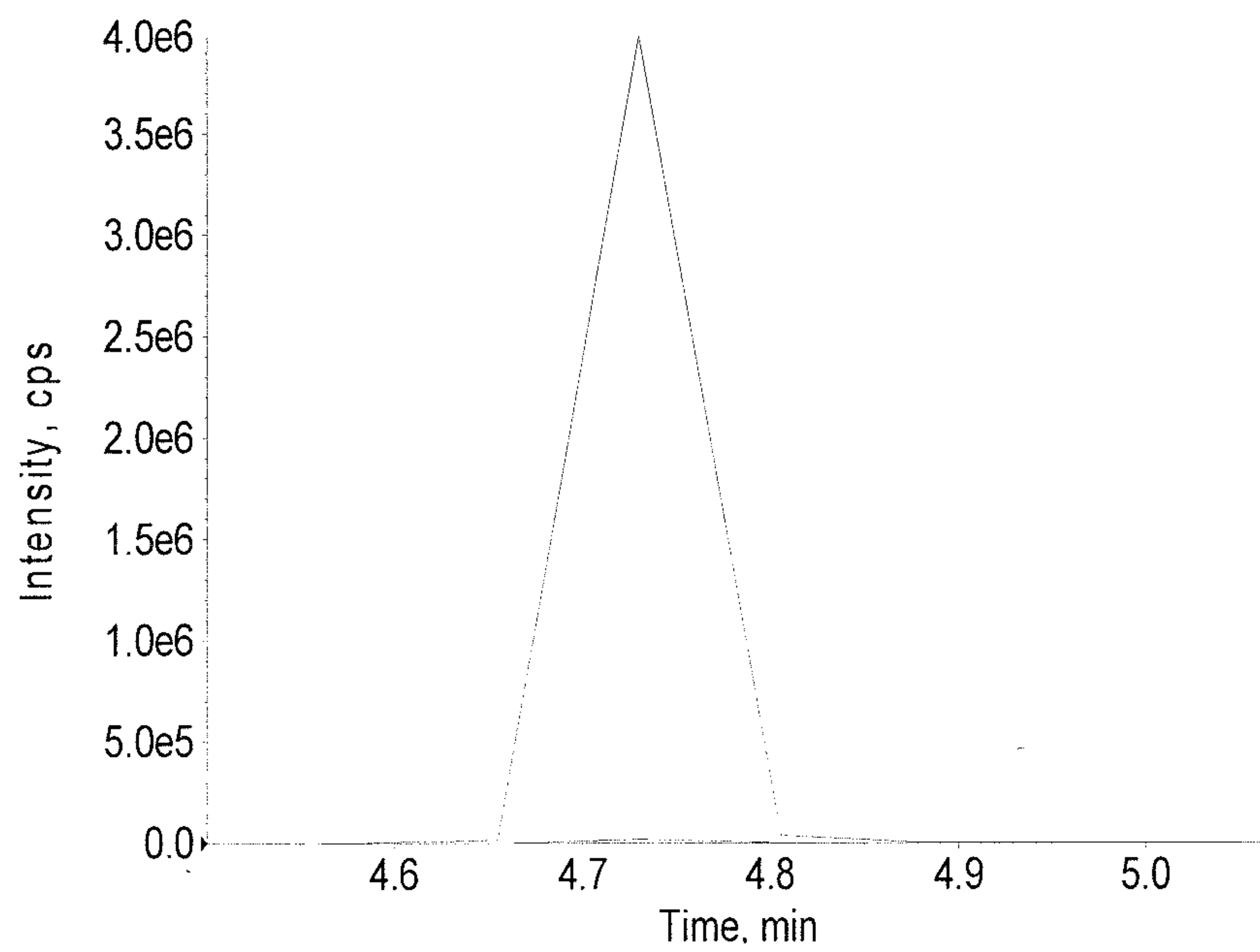
FIG. 8G

SLYEGGIK

FIG. 8H



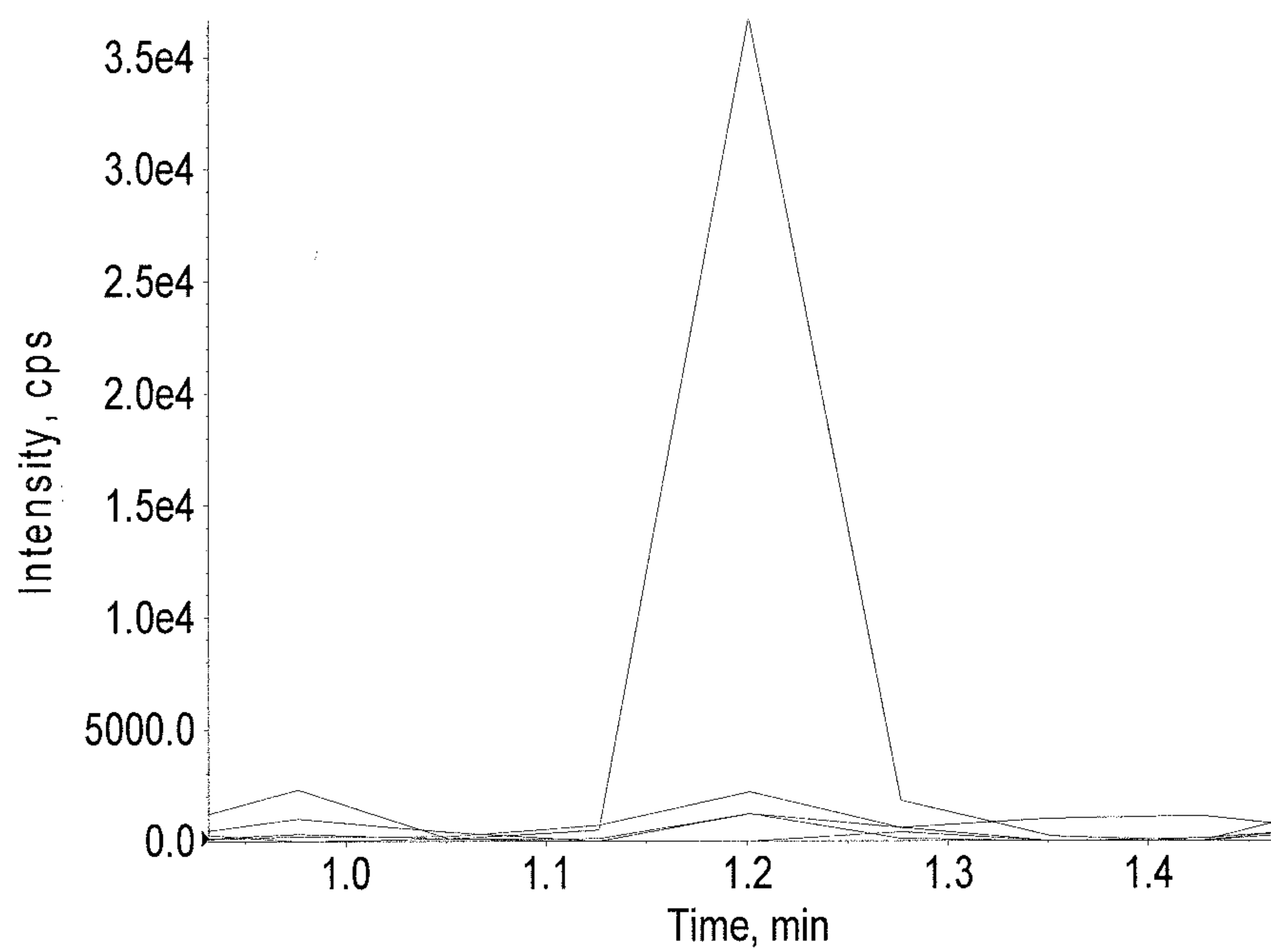
TDGENVLQFPPHVAK

FIG. 8I

INSLPTAK

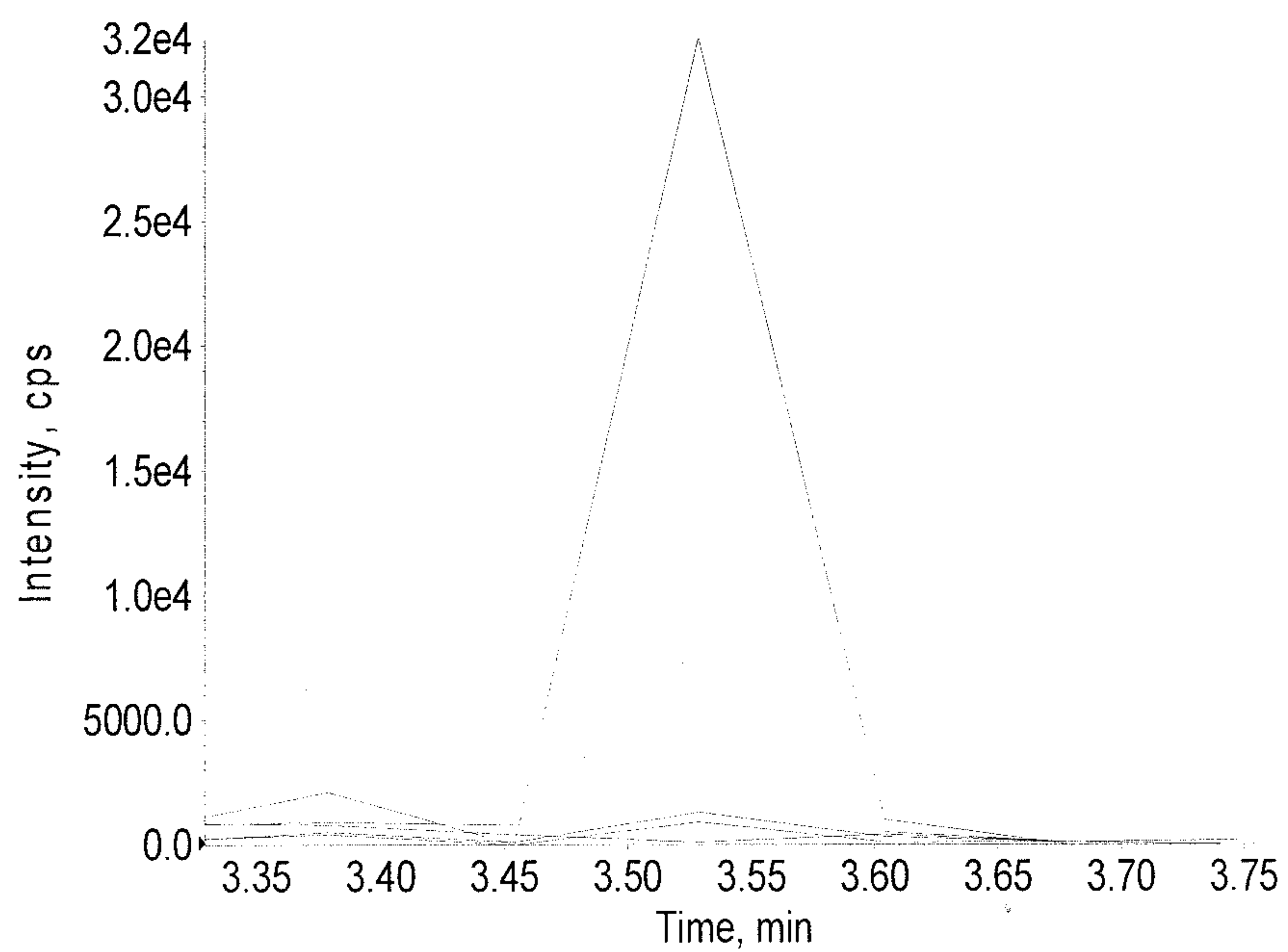
FIG. 8J

44/52



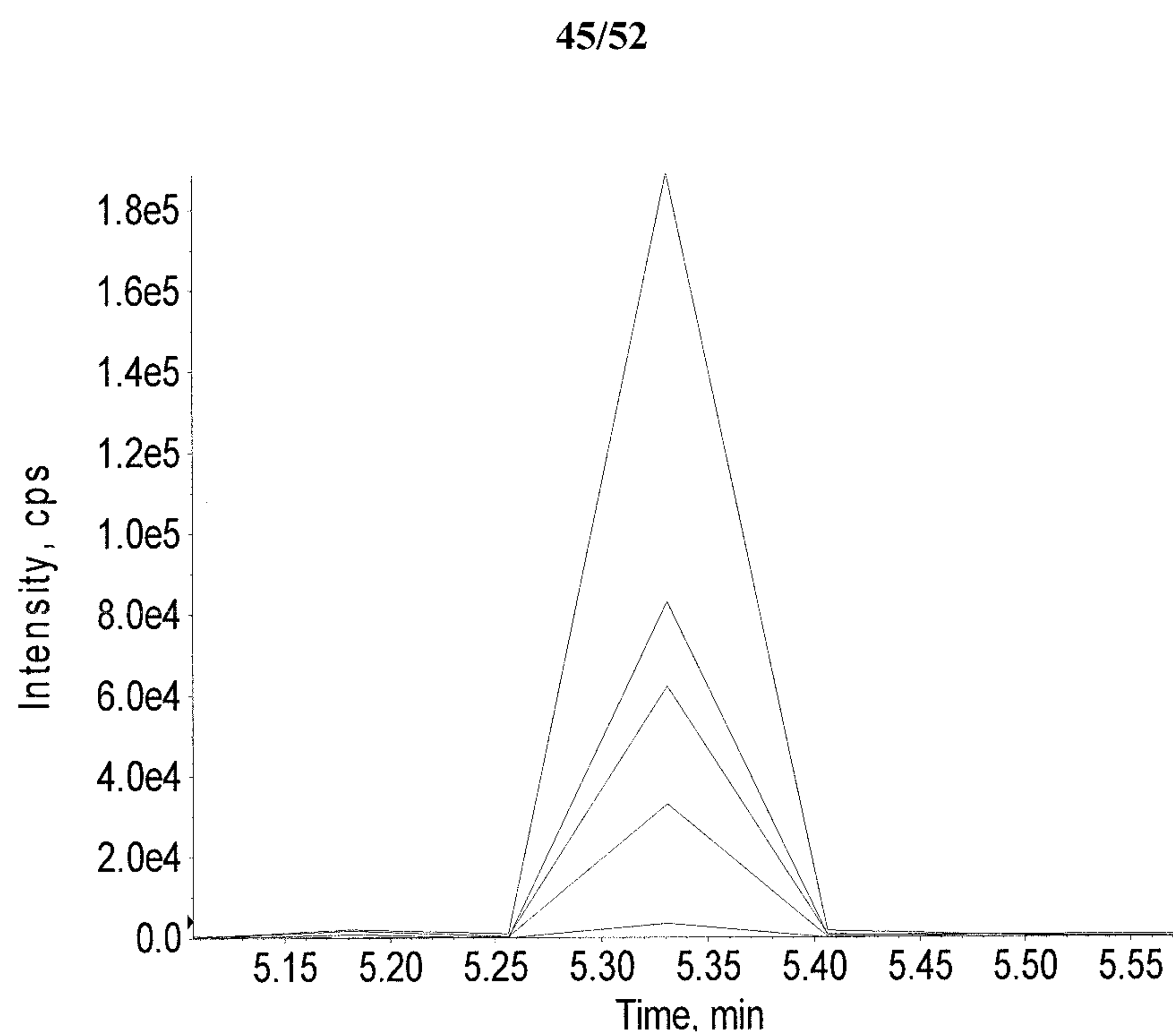
TILFLK

FIG. 8K

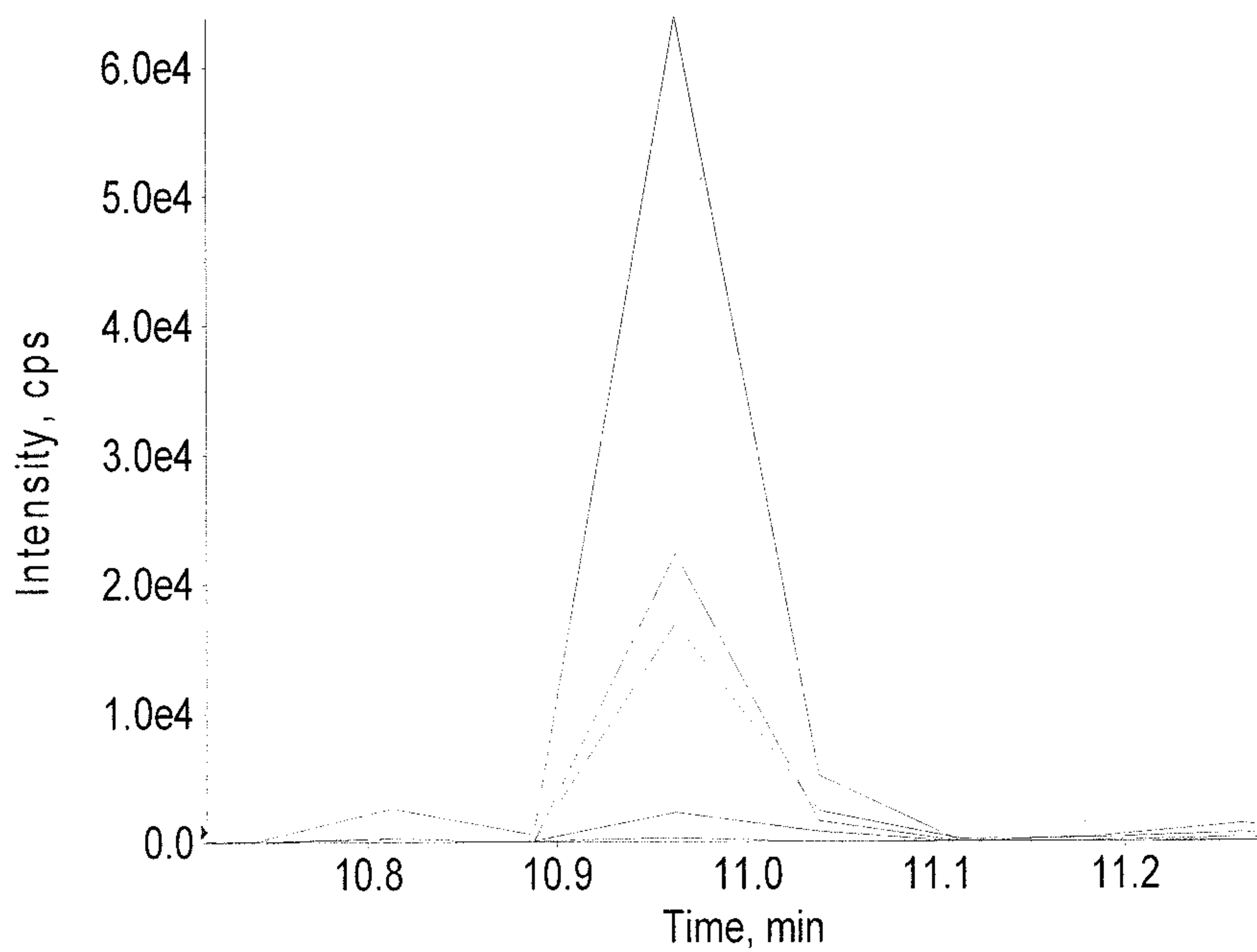


HLSVLHPIYK

FIG. 8L

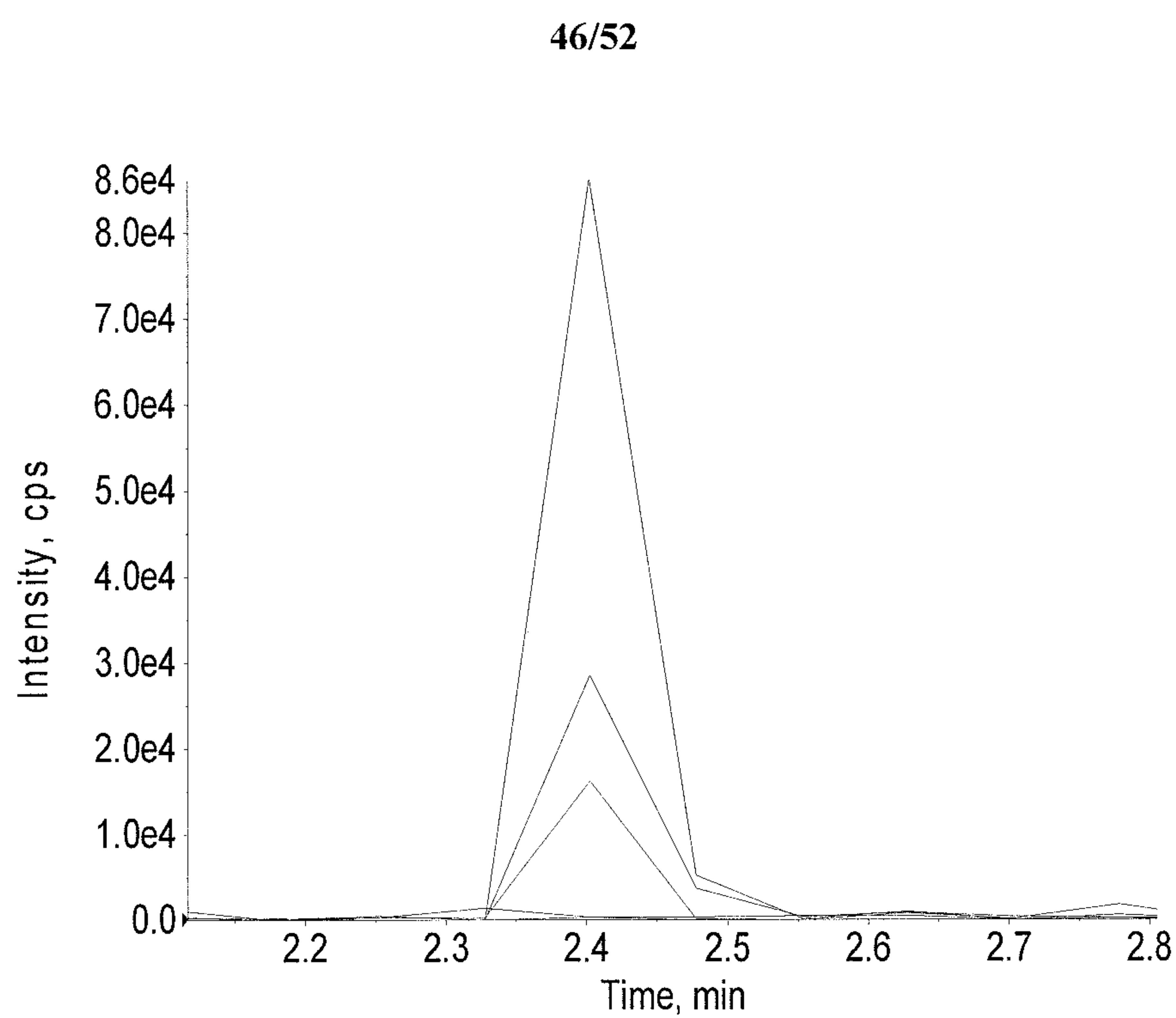


QSLINADGII EK

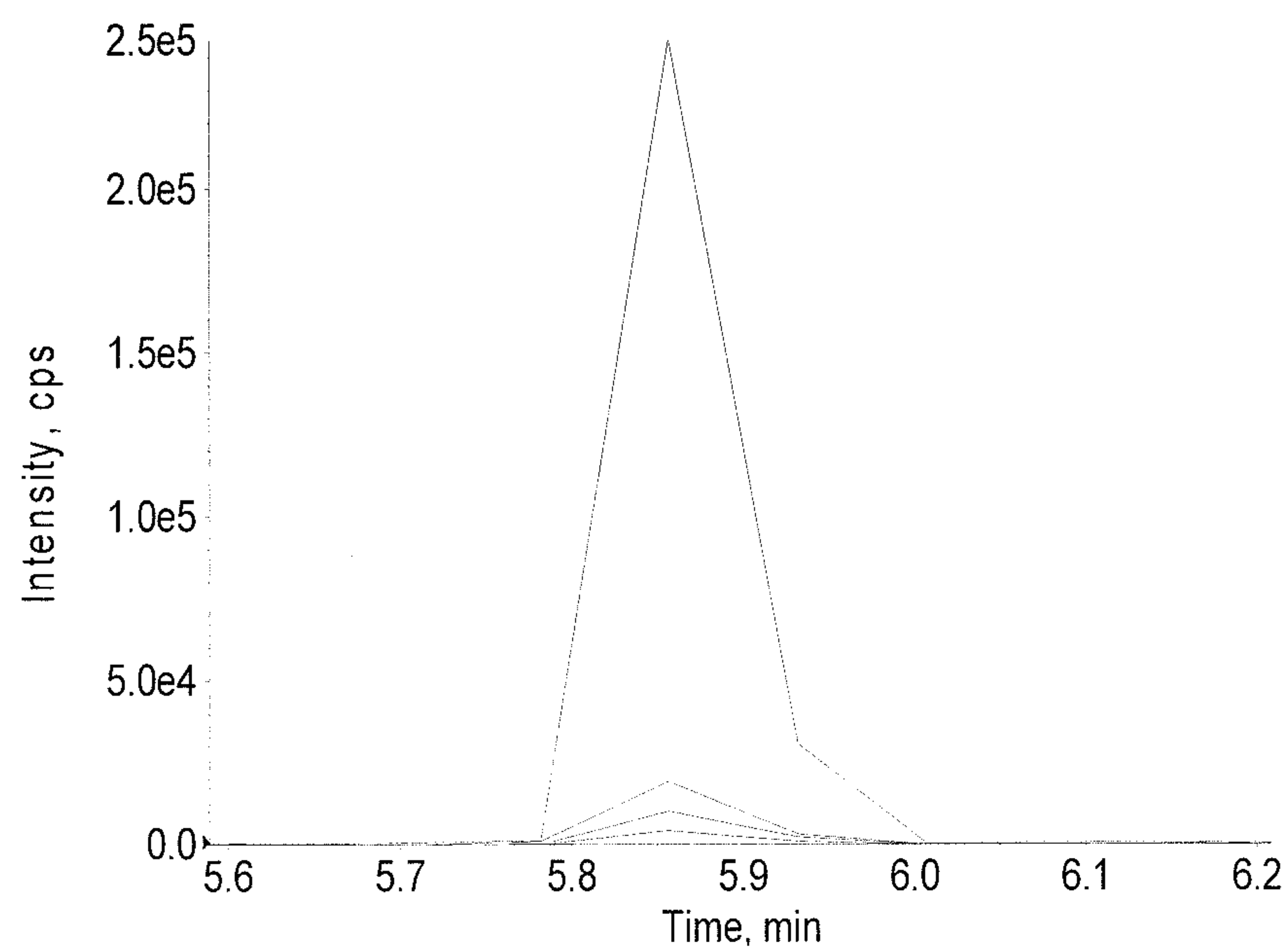
FIG. 8M

FIPAEGTPEYDEMVK

FIG. 8N

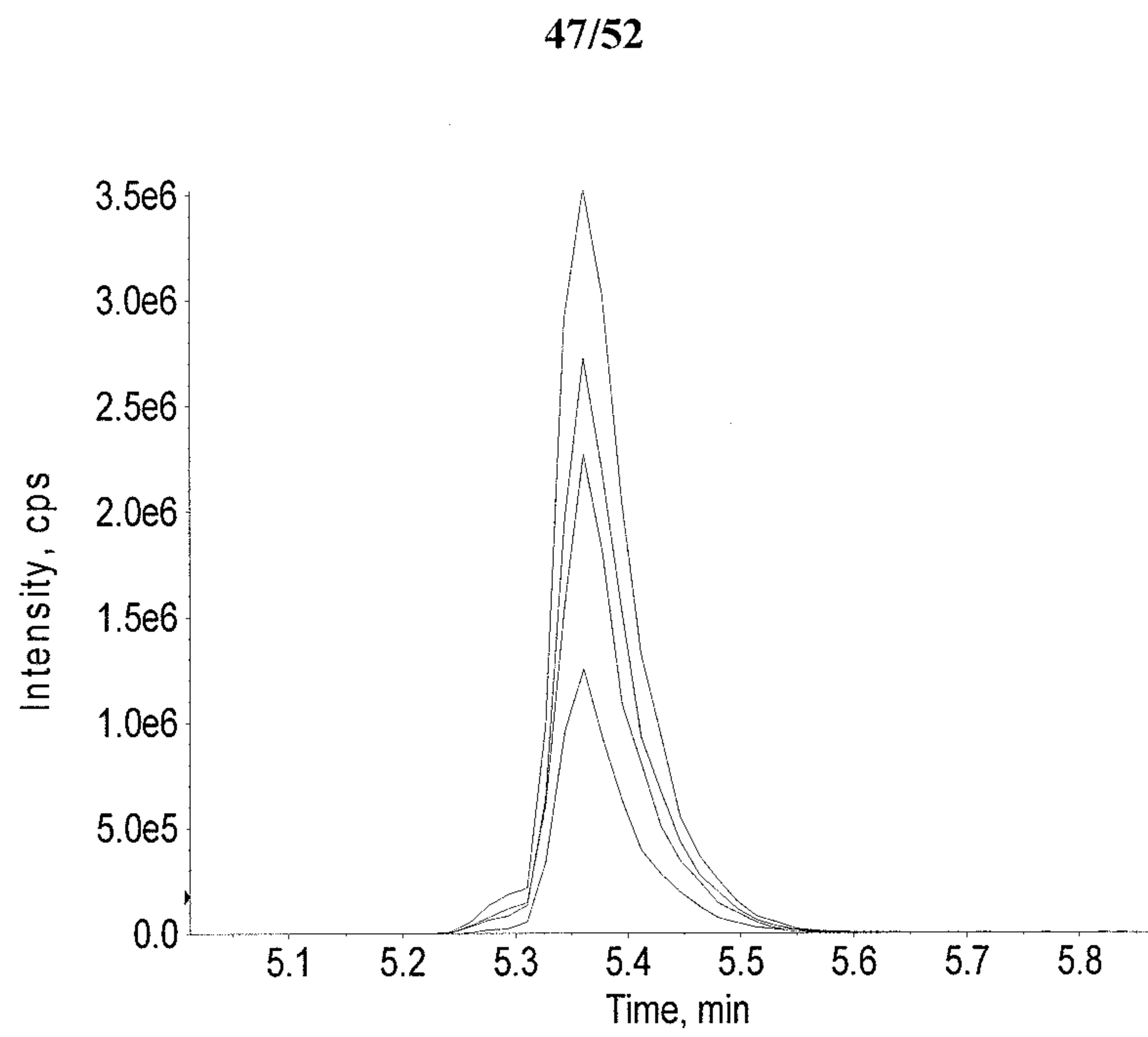


ALEAFK

FIG. 8O

GIPNSISI

FIG. 8P



SSDFLTYGK

FIG. 8Q

48/52

	1	50
Lipoxygenase 2IUK_A BLAST	(1) ----MFGIFDK-GQKIKGTVVLMFKNVLDNFNAITSIGKGGVIDTATGILG	
Lipoxygenase NP_001238676 BLAST	(1) ----MFGIFDK-GQKIKGTVVLMFKNVLDNFNAITSIGKGGVIDTATGILG	
Lipoxygenase CAA39604	(1) ----MFGIFDK-GQKIKGTVVLMFKNVLDNFNAITSIGKGGVIDTATGILG	
Lipoxygenase Glyma08g20220.1 BLAST	(1) ----MLGLFDK-SHKIKGTVVLMFKSVLDINDLNSVKNGGCVGVVSGIFG	
Lipoxygenase P24095 BLAST	(1) ----MFGIFDK-GQKIKGTVVLMFKNVLDNFNAITSIGKGGVIDTATGILG	
Lipoxygenase Glyma07g00900.1 BLAST	(1) ---MTGGMFGRKGQKIKGTVVLMFKNVLDNFNAITSVGKGSAKDTATDFLG	
Lipoxygenase NP_001235189 BLAST	(1) ---MTGGMFGRKGQKIKGTVVLMFKNVLDNFNAITSVGKGSAKDTATDFLG	
Lipoxygenase Glyma07g03910.1 BLAST	(1) ---MFGILGGNKGHKIKGTVVLMFKNVLDNFNEIVSTTQGGVLGAATGIFG	
Lipoxygenase Glyma07g03920.2 BLAST	(1) --MLIGSLNRR-PKIKGTVVLMFKNVLDNFNAITSGGPAAVAGGIFG	
Lipoxygenase Glyma08g20190.1 BLAST	(1) MYSGVKGLFNR-SQKVKGTVVLMRKNVLDINSITSVR-----G--LIG	
Consensus	(1) MGGIFDK GQKIKGTVVLMFKNVLDNFNAITSIGKGGVIDTATGILG	
	51	100
Lipoxygenase 2IUK_A BLAST	(46) QGVSLVGGVIDTATSFLGRNISMQLISATQTD-GSGNGKVGKEVYLEKHL	
Lipoxygenase NP_001238676 BLAST	(46) QGVSLVGGVIDTATSFLGRNISMQLISATQTD-GSGNGKVGKEVYLEKHL	
Lipoxygenase CAA39604	(46) QGVSLVGGVIDTATSFLGRNISMQLISATQTD-GSGNGKVGKEVYLEKHL	
Lipoxygenase Glyma08g20220.1 BLAST	(46) AVADVTGQIVDTATAIFSRNVSFKLISATSTD-AKNGKVGNETFLEKHL	
Lipoxygenase P24095 BLAST	(46) QGVSLVGGVIDTATSFLGRNISMQLISATQTD-GSGNGKVGKEVYLEKHL	
Lipoxygenase Glyma07g00900.1 BLAST	(48) KGLDALGHAVDALTAFAHGSISLQLISATQTD-GSGKGVGNEAYLEKHL	
Lipoxygenase NP_001235189 BLAST	(48) KGLDALGHAVDALTAFAHGSISLQLISATQTD-GSGKGVGNEAYLEKHL	
Lipoxygenase Glyma07g03910.1 BLAST	(48) AATGIVGGVVDGATAIFSRNIAIQLISATKTD-GLGNGKVGKQTYLEKHL	
Lipoxygenase Glyma07g03920.2 BLAST	(49) AAQDIVGGIVDGATAIFSRNIAIQLISATKSENALGHGKVGKLTYLEKHL	
Lipoxygenase Glyma08g20190.1 BLAST	(41) TGINIIGSTIDGLTSFLGRSVCLQLISATKAD-GNGNGVVGKKTYLEGII	
Consensus	(51) AGVSLVGGVIDTATAFLGRNISMQLISATQTD GSGNGKVGKEVYLEKHL	
	101	150
Lipoxygenase 2IUK_A BLAST	(95) PTLPTLGARQDAFSIFFEWDASFGIPGAFYIKNFMTDEFFLVSVKLEDIP	
Lipoxygenase NP_001238676 BLAST	(95) PTLPTLGARQDAFSIFFEWDASFGIPGAFYIKNFMTDEFFLVSVKLEDIP	
Lipoxygenase CAA39604	(95) PTLPTLGARQDAFSIFFEWDASFGIPGAFYIKNFMTDEFFLVSVKLEDIP	
Lipoxygenase Glyma08g20220.1 BLAST	(95) PTLPTLGDRRDAYDIHFEWDANFGIPGAFYIRNYTYDEFFLVSVTLEDIP	
Lipoxygenase P24095 BLAST	(95) PTLPTLGARQDAFSIFFEWDASFGIPGAFYIKNFMTDEFFLVSVKLEDIP	
Lipoxygenase Glyma07g00900.1 BLAST	(97) PTLPTLGARQEAFDINFEWDASFGIPGAFYIKNFMTDEFFLVSVKLEDIP	
Lipoxygenase NP_001235189 BLAST	(97) PTLPTLGARQEAFDINFEWDASFGIPGAFYIKNFMTDEFFLVSVKLEDIP	
Lipoxygenase Glyma07g03910.1 BLAST	(97) PSLPTLGDRQDAFSVYFEWDNDFGIPGAFYIKNFQSEFFLVSVTLEDIP	
Lipoxygenase Glyma07g03920.2 BLAST	(98) PSLPNLGDQRQDAFDVYFEWDESGIPGAFYIKNYMQSEFFLVSPKLEDVP	
Lipoxygenase Glyma08g20190.1 BLAST	(90) TSIPTLGAGQSAFTIHFEWDADMGIPGAFYIKNYMQVELFLVSLTLEDIP	
Consensus	(101) PTLPTLGARQDAFSIFFEWDASFGIPGAFYIKNFMTDEFFLVSVKLEDIP	
	151	200
Lipoxygenase 2IUK_A BLAST	(145) NHGTIEFVCNSWVYNFRSYKK-NRIFFVNDTYLPSATPAPLLKYRKEEFE	
Lipoxygenase NP_001238676 BLAST	(145) NHGTIEFVCNSWVYNFRSYKK-NRIFFVNDTYLPSATPAPLLKYRKEEFE	
Lipoxygenase CAA39604	(145) NHGTIEFVCNSWVYNFRSYKK-NRIFFVNDTYLPSATPAPLLKYRKEEFE	
Lipoxygenase Glyma08g20220.1 BLAST	(145) NHGTIEFVCNSWVYNFKDYDKKDRIFFANKTYLPSATPGPLVKYRKEELE	
Lipoxygenase P24095 BLAST	(145) NHGTIEFVCNSWVYNFRSYKK-NRIFFVNDTYLPSATPAPLLKYRKEEFE	
Lipoxygenase Glyma07g00900.1 BLAST	(147) NHGTIEFVCNSWVYNFKSYKK-NRIFFVNDTYLPSATPAPLLKYRKEEFE	
Lipoxygenase NP_001235189 BLAST	(147) NHGTIEFVCNSWVYNFKSYKK-NRIFFVNDTYLPSATPGPLVKYRQEELE	
Lipoxygenase Glyma07g03910.1 BLAST	(147) NHGTIEFVCNSWVYNKASYKK-DRIFFANKTYLNETPTPLVKYRKEEFE	
Lipoxygenase Glyma07g03920.2 BLAST	(148) NHGTILFACNSWVYNKLYKK-DRIFFANKAYLPNDTPTPLVKYRKEEFE	
Lipoxygenase Glyma08g20190.1 BLAST	(140) NQGSMHFVCNSWVYNKSYKK-DRIFFASETYVPSETPGPLVITYREAELO	
Consensus	(151) NHGTIEFVCNSWVYNFKSYKK NRIFFVNDTYLPSATPAPLLKYRKEEFE	

FIG. 8R

49/52

	201		250
Lipoxygenase 2IUK_A BLAST	(194)	VLRGDGTGKRKDFDRIYDYDVYNDLGNPDG--GDPRPILGGCSIYPYPLR	
Lipoxygenase NP_001238676 BLAST	(194)	VLRGDGTGKRKDFDRIYDYDVYNDLGNPDG--GDPRPILGGCSIYPYPLR	
Lipoxygenase CAA39604	(194)	VLRGDGTGKRKDFDRIYDYDVYNDLGNPDG--GDPRPILGGSSIYPYPRR	
Lipoxygenase Glyma08g20220.1 BLAST	(195)	ILRGDGTGERKEHERIYDYDVYNDLGNPDDEDVKLARPVLGGSSTYFYPRR	
Lipoxygenase P24095 BLAST	(194)	VLRGDGTGKRKDFDRIYDYDVYNDLGNPDG--GDPRPILGGSSIYPYPRR	
Lipoxygenase Glyma07g00900.1 BLAST	(196)	VLRGDGTGKRKDFDRIYDYDVYNDLGNPDG--GDPRPILGGSSIYPYPRR	
Lipoxygenase NP_001235189 BLAST	(196)	VLRGDGTGKRDFDRIYDYDIYNDLGNPDG--GDPRPIIGGSSNYFYPRR	
Lipoxygenase Glyma07g03910.1 BLAST	(196)	NLRGDGKGERKEYDRIYDYDVYNDLGNPDKSNDLARPVLGGSSAYFYPRR	
Lipoxygenase Glyma07g03920.2 BLAST	(197)	NLRGDGRGERKELDRIYDYDVYNDLGNPDENDDLARPILGGSSKHFYPRR	
Lipoxygenase Glyma08g20190.1 BLAST	(189)	ALRGNGTGKRKEWDRVYDYDVYNDLGNPDSENFARPVLGGSLSLTHFYPRR	
Consensus	(201)	VLRGDGTGKRKDFDRIYDYDVYNDLGNPDG	GDPRPILGGSSIYPYPRR
	251		300
Lipoxygenase 2IUK_A BLAST	(242)	VRTGRETRTRDPNSEKPG-EVYVPRDENFGHLKSSDFLTYGIKSLSHDVI	
Lipoxygenase NP_001238676 BLAST	(242)	VRTGRETRTRDPNSEKPG-EVYVPRDENFGHLKSSDFLTYGIKSLSHDVI	
Lipoxygenase CAA39604	(242)	VRTGRETRTRDPNSEKPG-EVYVPRDENFGHLKSSDFLTYGIKSLSHDVI	
Lipoxygenase Glyma08g20220.1 BLAST	(245)	VRTGRKATKKDPKSERPASELYMPRDEKFGHLKSSDFLTYGIKSLSQKLL	
Lipoxygenase P24095 BLAST	(242)	VRTGRETRTRDPNSEKPG-EVYVPRDENFGHLKSSDFLTYGIKSLSHDVI	
Lipoxygenase Glyma07g00900.1 BLAST	(244)	VRTGRETRTRDPNSEKPG-EVYVPRDENFGHLKSSDFLTYGIKSLSHDVI	
Lipoxygenase NP_001235189 BLAST	(244)	VRTGREKTRKDPNSEKPG-EIYVPRDENFGHLKSSDFLTYGIKSLSQNVI	
Lipoxygenase Glyma07g03910.1 BLAST	(246)	GRTGRKPTTKDSKSESPSSSTYIPRDNFGLKSSDFLTYGIKSLAQTVL	
Lipoxygenase Glyma07g03920.2 BLAST	(247)	GRTGRKPTKKDPRCERPTSSTYIPRDNFGLKSSDFLTYAIKSLTQNVL	
Lipoxygenase Glyma08g20190.1 BLAST	(239)	GRTGRKPTKKDPNSEKPG-EAYIPRDNFGLKSSDFLTYGLKSLTRSFL	
Consensus	(251)	VRTGRETRTRDPNSEKPG	EVYVPRDENFGHLKSSDFLTYGIKSLSHDVI
	301		350
Lipoxygenase 2IUK_A BLAST	(291)	PLFKSAIFQLRVTSSFEFESFEDVRSLYEGGIKLPDILSQISPLPALKEI	
Lipoxygenase NP_001238676 BLAST	(291)	PLFKSAIFQLRVTSSFEFESFEDVRSLYEGGIKLPDILSQISPLPALKEI	
Lipoxygenase CAA39604	(291)	PLFKSAIFQLRVTSSFEFESFEDVRSLYEGGIKLPDILSQISPLPALKEI	
Lipoxygenase Glyma08g20220.1 BLAST	(295)	FSLEN-VFDSDLTWNEFDSFEFVRDLYEGGIKVFTGVLSDISPIPIFKEI	
Lipoxygenase P24095 BLAST	(291)	PLFKSAIFQLRVTSSFEFESFEDVRSLYEGGIKLPDILSQISPLPALKEI	
Lipoxygenase Glyma07g00900.1 BLAST	(293)	PLFKSAIFQLRVTSSFEFESFEDVRSLYEGGIKLPDILSQISPLPALKEI	
Lipoxygenase NP_001235189 BLAST	(293)	PLFKSIILNLRVTSSFEFDSFDEVRLFEFGGIKLPDILSQISPLPALKEI	
Lipoxygenase Glyma07g03910.1 BLAST	(296)	PTFQS---AF-GLNAEFDRFDDVRGLFEGGIHLPTDALSISPLPALKEI	
Lipoxygenase Glyma07g03920.2 BLAST	(297)	PQENT---AF-GFNNEFDSFEDVRCLFDGGVYLPTDVLKISPIPLKEI	
Lipoxygenase Glyma08g20190.1 BLAST	(288)	PALKT-VFDINFTPNFDSFEFVRALCEGGIKLPDILSKISPLPALKEI	
Consensus	(301)	PLFKSAIFQLRVTSSFEFDSFEDVRSLYEGGIKLPDILSQISPLPALKEI	
	351		400
Lipoxygenase 2IUK_A BLAST	(341)	FRTDGENVLQFPPPHVAKVSKSGWMTDEEFAREVIAGVNPVIRRLQEFF	
Lipoxygenase NP_001238676 BLAST	(341)	FRTDGENVLQFPPPHVAKVSKSGWMTDEEFAREVIAGVNPVIRRLQEFF	
Lipoxygenase CAA39604	(341)	FRTDGENVLQFPPPHVAKVSKSGWMTDEEFAREVIAGVNPVIRRLQEFF	
Lipoxygenase Glyma08g20220.1 BLAST	(344)	FRTDGESVLQFPPPHVQVTKSAWMTDDEFAREMIAGVNPVIRLLKEFF	
Lipoxygenase P24095 BLAST	(341)	FRTDGENVLQFPPPHVAKVSKSGWMTDEEFAREVIAGVNPVIRRLQEFF	
Lipoxygenase Glyma07g00900.1 BLAST	(343)	FRTDGENVLQFPPPHVAKVSKSGWMTDEEFAREVIAGVNPVIRRLQEFF	
Lipoxygenase NP_001235189 BLAST	(343)	FRTDGENTLQFPPPHVIRVSKSGWMTDDEFAREMIAGVNPVIRRLQEFF	
Lipoxygenase Glyma07g03910.1 BLAST	(342)	FRTDGEQVLKFPFPHVIVKSKSAWMTDEEFGREMLAGVNPCLIECLQVFP	
Lipoxygenase Glyma07g03920.2 BLAST	(343)	FRTDGEQALKFPFPHVIVKRESEWMTDEEFGREMLAGVNPGLMIQRLQEFF	
Lipoxygenase Glyma08g20190.1 BLAST	(337)	FRTDGESVLKFSVPDLIKVSKSAWMTDEEFAREMIAGVNPVIRRLQEFF	
Consensus	(351)	FRTDGENVLQFPPPHVAKVSKSGWMTDEEFAREMIAGVNPVIRRLQEFF	

FIG. 8R - continued

50/52

	401	450
Lipoxygenase 2IUK_A BLAST	(391) PKSTLDPTLYGDQSTITKEQLEINMGGVTVEEALSTQRLFILDYQDAFI	
Lipoxygenase NP_001238676 BLAST	(391) PKSTLDPTLYGDQSTITKEQLEINMGGVTVEEALSTQRLFILDYQDAFI	
Lipoxygenase CAA39604	(391) PKSTLDPTLYGDQSTITKEQLEINMGGVTVEEALSTQRLFILDYQDAFI	
Lipoxygenase Glyma08g20220.1 BLAST	(394) PQSKLDPSLYGDQSSTITKEHLEINMDGVTVEEALNGQRLFILDYQDAFM	
Lipoxygenase P24095 BLAST	(391) PKSTLDPTLYGDQSTITKEQLEINMGGVTVEEALSTQRLFILDYQDAFI	
Lipoxygenase Glyma07g00900.1 BLAST	(393) PKSTLDPTLYGDQSTITKEQLEINMGGVTVEEALSTQRLFILDYQDAFI	
Lipoxygenase NP_001235189 BLAST	(393) PKSTLDPATYGDQSTITKQOLEINLGGVTVEEALSAHRLFILDYHDAFF	
Lipoxygenase Glyma07g03910.1 BLAST	(392) PKSKLDPTVYGDQSTITKEHLEINLGGLSVEQALSGNRLFILDHHDFAFI	
Lipoxygenase Glyma07g03920.2 BLAST	(393) PKSKLDPTFEGDQSTITKEHLEINLGGTVEQALKGNKLFILDHHDFAFI	
Lipoxygenase Glyma08g20190.1 BLAST	(387) PQSKLDPSVYGDQTSKMTIDHLEINLEGLTVDKAIKDQRLFILDHHDFTFM	
Consensus	(401) PKSTLDPTLYGDQSTITKEQLEINMGGVTVEEALSTQRLFILDYQDAFI	451
Lipoxygenase 2IUK_A BLAST	(441) PYLTRINSLPTAKAYATRTILFLKDDGTLKPLAIELSKPHPDGDNLGPES	500
Lipoxygenase NP_001238676 BLAST	(441) PYLTRINSLPTAKAYATRTILFLKDDGTLKPLAIELSKPHPDGDNLGPES	
Lipoxygenase CAA39604	(441) PYLTRINSLPTAKAYATRTILFLKDDGTLKPLAIELSKPHPDGDNLGPES	
Lipoxygenase Glyma08g20220.1 BLAST	(444) PYLTRINALPSAKAYATRTILLKDDGTLKPLAIELSKPHPSGDNLGAES	
Lipoxygenase P24095 BLAST	(441) PYLTRINSLPTAKAYATRTILFLKDDGTLKPLAIELSKPHPDGDNLGPES	
Lipoxygenase Glyma07g00900.1 BLAST	(443) PYLTRINSLPTAKAYATRTILFLKDDGTLKPLAIELSKPHPDGDNLGPES	
Lipoxygenase NP_001235189 BLAST	(443) PYLTKINSLPIAKAYATRTILFLKDDGSLKPLAIELSKP-----ATVS	
Lipoxygenase Glyma07g03910.1 BLAST	(442) AYLRKINDLPTAKSYATRTILFLKDDGTLKPLAIELSLPHPRGDEFGAVS	
Lipoxygenase Glyma07g03920.2 BLAST	(443) PFMNLINGLPTAKSYATRTILFLQDDGTLKPLAIELSLPHPRGHEFGADS	
Lipoxygenase Glyma08g20190.1 BLAST	(437) PFLRRIDESKSSKAYATRTILFLKDDGTLKPLAIELSLPHPGQQQLGAYS	
Consensus	(451) PYLTRINSLPTAKAYATRTILFLKDDGTLKPLAIELSKPHPDGDNLGPES	501
Lipoxygenase 2IUK_A BLAST	(491) IVVLP--ATEGVDSTIWLLAKAHVIVNDSGYHQLVSHWLNTHAVMEPFAI	550
Lipoxygenase NP_001238676 BLAST	(491) IVVLP--ATEGVDSTIWLLAKAHVIVNDSGYHQLVSHWLNTHAVMEPFAI	
Lipoxygenase CAA39604	(491) IVVLP--ATEGVDSTIWLLAKAHVIVNDSGYHQLVSHWLNTHAVMEPFAI	
Lipoxygenase Glyma08g20220.1 BLAST	(494) KVVLP--ADQGVESTIWLLAKAHVIVNDSGYHQLMSHWLNTHAVTEPFII	
Lipoxygenase P24095 BLAST	(491) IVVLP--ATEGVDSTIWLLAKAHVIVNDSGYHQLVSHWLNTHAVMEPFAI	
Lipoxygenase Glyma07g00900.1 BLAST	(493) IVVLP--ATEGVDSTIWLLAKAHVIVNDSGYHQLVSHWLNTHAVMEPFAI	
Lipoxygenase NP_001235189 BLAST	(486) KVVLP--ATEGVESTIWLLAKAHVIVNDSGYHQLISHWLNTHAVMEPFAI	
Lipoxygenase Glyma07g03910.1 BLAST	(492) RVVLP--ADQGAESTIWLIKAYVVVNDSGYHQLMSHWLNTHAVIEPFVI	
Lipoxygenase Glyma07g03920.2 BLAST	(493) RVVLPFAAVNSAEGTIWLIKAYVAVNDTGYHQLISHWLNTHATIEPFVI	
Lipoxygenase Glyma08g20190.1 BLAST	(487) KVILP--ANQGVESTIWLLAKAHVIVNDSGYHQLISHWLNTHAVIEPFVI	
Consensus	(501) IVVLP ATEGVDSTIWLLAKAHVIVNDSGYHQLVSHWLNTHAVMEPFAI	551
Lipoxygenase 2IUK_A BLAST	(539) ATNRHLSVLHPIYKLLYPHYRDTININGLARQSLINADGIIIEKSFLPGKY	600
Lipoxygenase NP_001238676 BLAST	(539) ATNRHLSVLHPIYKLLYPHYRDTININGLARQSLINADGIIIEKSFLPGKY	
Lipoxygenase CAA39604	(539) ATNRHLSVLHPIYKLLYPHYRDTININGLARQSLINADGIIIEKSFLPGKY	
Lipoxygenase Glyma08g20220.1 BLAST	(542) ATNRRLSVLHPIYKLLYPHYRDTININGLARNALINAGGVIEESFLPGRY	
Lipoxygenase P24095 BLAST	(539) ATNRHLSVLHPIYKLLYPHYRDTININGLARQSLINADGIIIEKSFLPGKY	
Lipoxygenase Glyma07g00900.1 BLAST	(541) ATNRHLSVLHPIYKLLYPHYRDTININGLARQSLINADGIIIEKSFLPGKY	
Lipoxygenase NP_001235189 BLAST	(534) ATNRHLSVLHPIYKLLYPHYRDTININGLARQSLINAGGIIIEQTFLPGKY	
Lipoxygenase Glyma07g03910.1 BLAST	(540) ATNRHLSVLHPIYKLLYPHYRDTMNINGLARQSLINAGGIIIEQSFLPGPF	
Lipoxygenase Glyma07g03920.2 BLAST	(543) ATNRHLSVLHPIHKLKLLYPHYRDTMNINALARQSLINADGVIERSEFLPGKY	
Lipoxygenase Glyma08g20190.1 BLAST	(535) ATNRNLSILHPIYKLLFPHYRDTMNINALARQSLINADGFIEKTFLLGGKY	
Consensus	(551) ATNRHLSVLHPIYKLLYPHYRDTININGLARQSLINADGIIIEKSFLPGKY	

FIG. 8R - continued

51/52

	601	650
Lipoxygenase 2IUK_A BLAST	(589)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase NP_001238676 BLAST	(589)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase CAA39604	(589)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase Glyma08g20220.1 BLAST	(592)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase P24095 BLAST	(589)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase Glyma07g00900.1 BLAST	(591)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase NP_001235189 BLAST	(584)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase Glyma07g03910.1 BLAST	(590)	AVEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase Glyma07g03920.2 BLAST	(593)	SLEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Lipoxygenase Glyma08g20190.1 BLAST	(585)	AVEISSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
Consensus	(601)	SIEMSSSVYKNWVFTHQALPADLVKRGGLAIEDPSAPHGLRLVIEDYPYAV
	651	700
Lipoxygenase 2IUK_A BLAST	(639)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase NP_001238676 BLAST	(639)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase CAA39604	(639)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase Glyma08g20220.1 BLAST	(642)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase P24095 BLAST	(639)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase Glyma07g00900.1 BLAST	(641)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase NP_001235189 BLAST	(634)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase Glyma07g03910.1 BLAST	(640)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase Glyma07g03920.2 BLAST	(643)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Lipoxygenase Glyma08g20190.1 BLAST	(635)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
Consensus	(651)	DGLEIWDIAIKTWVHEYVSLYYPTDAAVQQDTELQAWWKEAVEKGHGDLE
	701	750
Lipoxygenase 2IUK_A BLAST	(689)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase NP_001238676 BLAST	(689)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase CAA39604	(689)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase Glyma08g20220.1 BLAST	(692)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase P24095 BLAST	(689)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase Glyma07g00900.1 BLAST	(691)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase NP_001235189 BLAST	(684)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase Glyma07g03910.1 BLAST	(690)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase Glyma07g03920.2 BLAST	(693)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Lipoxygenase Glyma08g20190.1 BLAST	(685)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
Consensus	(701)	KPWWPCKQTTEDLIQSCSIIIVWTASALHAAVNFGQYPYGGILNRPDLAR
	751	800
Lipoxygenase 2IUK_A BLAST	(739)	RFIPAEGTPEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase NP_001238676 BLAST	(739)	RFIPAEGTPEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase CAA39604	(739)	RFIPAEGTPEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase Glyma08g20220.1 BLAST	(742)	RWIPEPGTKEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase P24095 BLAST	(739)	RFIPAEGTPEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase Glyma07g00900.1 BLAST	(741)	RFIPAEGTPEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase NP_001235189 BLAST	(734)	RFIPAEGTPEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase Glyma07g03910.1 BLAST	(740)	RLLPEPGTKEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase Glyma07g03920.2 BLAST	(743)	RFLPEPGTKEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Lipoxygenase Glyma08g20190.1 BLAST	(735)	RWIPEPGTKEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI
Consensus	(751)	RFIPAEGTPEYDEMVKNPQKAYLRTITPKFETLIDLSVIEILSRHASDEI

FIG. 8R - continued

52/52

	801	850
Lipoxygenase 2IUK_A BLAST	(789)	YLGERETPNWTTDDKKALEAFKKRFGSKLTGIEGKINARNSDPSLRNRTGPV
Lipoxygenase NP_001238676 BLAST	(789)	YLGERETPNWTTDDKKALEAFKKRFGSKLTGIEGKINARNSDPSLRNRTGPV
Lipoxygenase CAA39604	(789)	YLGERETPNWTTDDKKALEAFKKRFGSKLTGIEGKINARNSDPSLRNRTGPV
Lipoxygenase Glyma08g20220.1 BLAST	(792)	YLGERDNPNWTTSDSKALEAFKKRFGSKLAEIEGKITARNKDSNKKNRYPV
Lipoxygenase P24095 BLAST	(789)	YLGERETPNWTTDDKKALEAFKKRFGSKLTGIEGKINARNSDPSLRNRTGPV
Lipoxygenase Glyma07g00900.1 BLAST	(791)	YLGERETPNWTTDDKKALEAFKKRFGSKLTGIEGKINARNSDPSLRNRTGPV
Lipoxygenase NP_001235189 BLAST	(784)	YLGQRDNPNWTTDDSKALEAFKKFGNKLAEIEGKITQRNNDPSLKSRRHGPV
Lipoxygenase Glyma07g03910.1 BLAST	(790)	YLGQRDNPNWTTDDTKAIQAFKKFGNKLKEIEDKISGRNKNSSLRNRRNGPA
Lipoxygenase Glyma07g03920.2 BLAST	(793)	YLGKRDSDDWTTDDQKAIQAFKFGTKLKEIEAKINSRNKDSLSLRNRNGPV
Lipoxygenase Glyma08g20190.1 BLAST	(785)	YLGQRDNPNWTSNPKAIEAFKKFGKLAETKISERNHDPNLRNRTGPA
Consensus	(801)	YLGERDTPNWTTDDKKALEAFKKRFGSKLTGIEGKINARNSDPSLRNRTGPV
	851	876
Lipoxygenase 2IUK_A BLAST	(839)	QLPYTLLHRSSEEGLTFFKGIPNSISI
Lipoxygenase NP_001238676 BLAST	(839)	QLPYTLLHRSSEEGLTFFKGIPNSISI
Lipoxygenase CAA39604	(839)	QLPYTLLHRSSEEGLTFFKGIPNSISI
Lipoxygenase Glyma08g20220.1 BLAST	(842)	QLPYTLLLPSTSEEGLTFRGIPNSISI
Lipoxygenase P24095 BLAST	(839)	QLPYTLLHRSSEEGLTFFKGIPNSISI
Lipoxygenase Glyma07g00900.1 BLAST	(841)	QLPYTLLHRSSEEGLTFFKGIPNSISI
Lipoxygenase NP_001235189 BLAST	(834)	QLPYTLLHRSSEEGMSFKGIPNSISI
Lipoxygenase Glyma07g03910.1 BLAST	(840)	QMPYTVLLPSTSGEGLTFRGIPNSISI
Lipoxygenase Glyma07g03920.2 BLAST	(843)	QMPYTVLLPSTSEEGLTFRGIPNSISI
Lipoxygenase Glyma08g20190.1 BLAST	(835)	QLPYTVLLPSTSETGLTFRGIPNSISI
Consensus	(851)	QLPYTLLHRSSEEGLTFFKGIPNSISI

FIG. 8R - continued

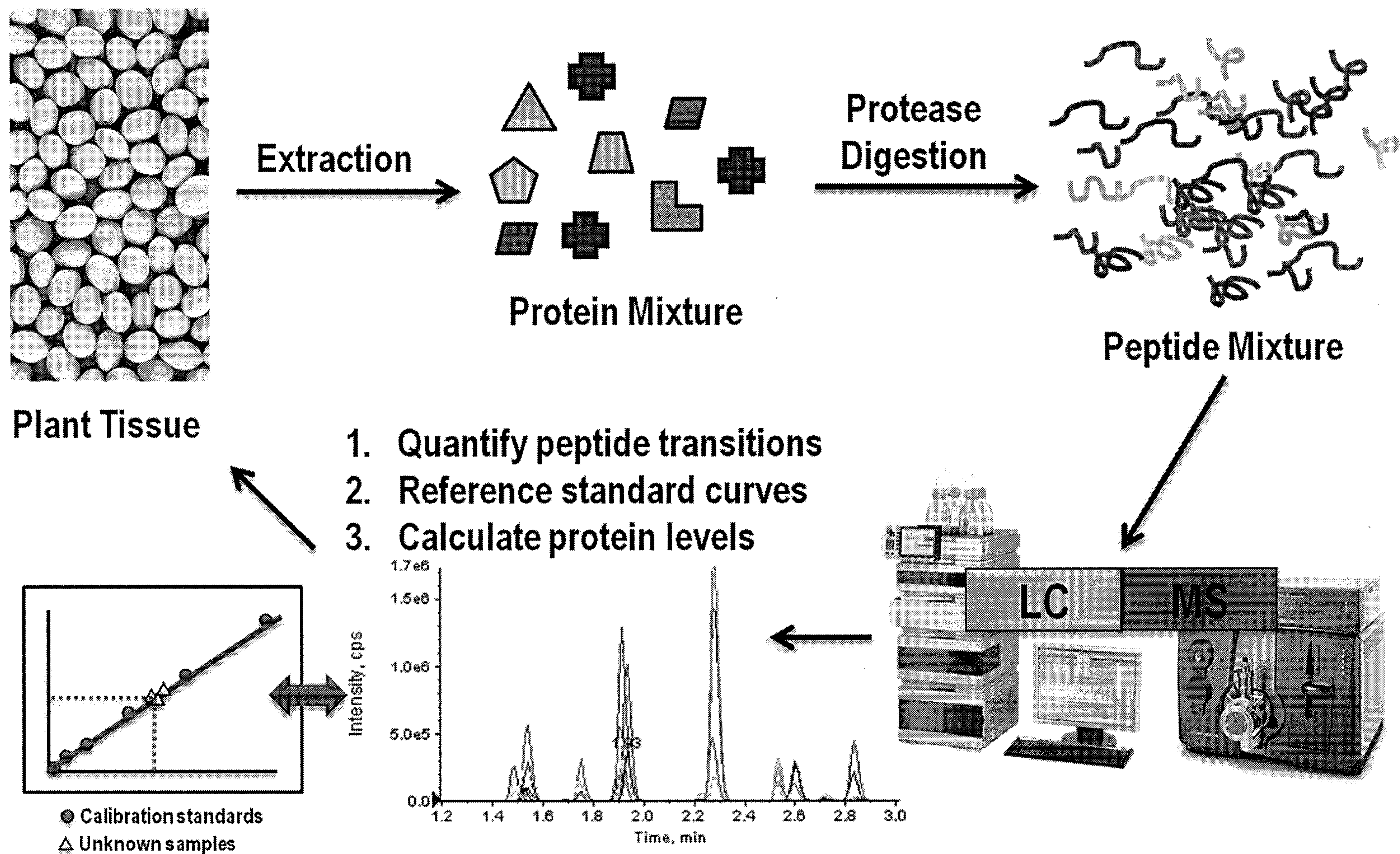


FIG. 1