ANALYSIS OF ELECTROPHILE-INDUCED NRF2 GENE ACTIVATION

by

Fei Hong

A Dissertation Submitted to the Faculty of

DEPARTMENT OF PHARMACEUTICAL SCIENCES

In Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2005

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation

prepared by Fei Hong

entitled ANALYISIS OF ELECTROPHILE-INDUCED NRF2 GENE ACTIVATION

and recommend that it be accepted as fulfilling the dissertation requirement for the

Degree of Doctor of Philosophy

Myron K. Jacobson; Daniel C. Liebler Date: 07/25/2005

Bernard W. Futscher Date: 07/25/2005

Danzhou Yang Date: 07/25/2005

Ann E. Cress Date: 07/25/2005

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Myron K. Jacobson; Daniel C. Liebler Date: 07/25/2005
Dissertation Director

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirement for an advanced degree at the University of Arizona and is deposited into the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Fei Hong

ACKNOWLEDGEMENTS

The author would like to acknowledge the contributions of several individuals who contributed significantly to the research described in this dissertation. Instrumentation for this research was made possible by the Proteomic Laboratory of the Mass Spectrometry Research Center at the Vanderbilt University School of Medicine and Thermo Electron. The development of the protocols described in Chapter 2 was a collaborative effort between Dr. Liebler's and Dr. Michael Freeman's groups. I would like to thank Dr. Freeman for his support, especially for my work described in Chapter 2. The concept for detecting Keap1 adducts via mass spectrometry was the "brainchild" of Dr. Daniel Liebler and the author would like to thank Dr. Liebler for his patience, support and guidance that went into making this dissertation possible. Finally, I would like to thank my mother and my brother, especially my husband Hao for their support and providing me the strength to complete this Ph.D. Funding for this research was provided by the following grants and is gratefully acknowledged: NIH ES010056, CA104590 & ES000267.

TABLE OF CONTENTS

LIST OF FIGURES AND SCHEMES 7

LIST OF TABLES 9

LIST OF ABBREVIATIONS 10

ABSTRACT 14

CHAPTER ONE-INTRODUCTION 16

- **Chemoprevention of Cancer** 16
- **Chemoprevention by Isothiocyanates** 17
- **Induction of Phase II Enzymes through the ARE Regulated by Nrf2** 19
- **Role of Nrf2 in Chemoprevention** 22
- **Keap1 as a Negative Regulator of Nrf2** 24
- **Other Mechanisms Involved in Nrf2 Activation** 28
- **Typical Phase II Inducers and Their Thiol Reactivity** 31
- **Analysis of Chemical Adducts and Posttranslational Modification Sites on Proteins** 34
- **Protein Identification and Modification Site Mapping from MS Data and MS-MS Data** 37
- **Sample Preparation for LC-MS-MS** 42
- **Sequence Coverage is an Issue for Modification Identification** 43
- **Strategies to Improve Detection of Modified Peptides** 43
- **Statement of the Problem** 45

CHAPTER TWO-SPECIFIC PATTERNS OF ELECTROPHILE ADDUCTION TRIGGER KEAP1 UBIQUITINAION AND NRF2 ACTVIATION 47

- **Introduction** 47
- **Experimental Procedures** 50
  - ***Cell culture, transfection and treatment*** 50
  - ***Cell fractionation and immunoblot analyses*** 50
  - ***Confocal microscopic analysis*** 51
  - ***Covalent adduction of model electrophiles to His6-Keap1 in vitro*** 52
  - ***Affinity capture and digestion of FLAG-Keap1*** 53
  - ***In-gel digestion of purified Keap1 separated through SDS-PAGE*** 54
  - ***LC-MS-MS analyses*** 54
- **Results** 56
  - ***Identification of Keap1 cysteine residues modified by eletrophiles in vitro*** 56
  - ***IAB and BMCC produce differential Nrf2 activation in vivo*** 59

TABLE OF CONTENTS – *Continued*

Differential effects of IAB and BMCC on formation of high molecular weight Keap1 forms in vivo 62
Characterization of ubiquitination and IAB adduction sites on Keap1 in vivo 70
Nrf2 stabilization coincides with electrophile-specific adduction and ubiquitination of Keap1 in vivo 72
Discussion 75

CHAPTER THREE-IDENTIFICATION OF SENSOR CYSTEINES IN HUMAN KEAP1 MODIFIED BY THE CHEMOPREVENTIVE AGENT SULFORAPHANE 81
Introduction 81
Experimental Procedures 85
Cell culture, transfection and treatment 85
Cell fractionation and immunoblot analyses 85
Flow cytometry analysis of GFP expression induced by sulforaphane 86
Modification of TpepC by sulforaphane 87
Covalent adduction of sulforaphane to His6-Keap1 in vitro 87
LC-MS-MS analyses 88
Results 90
Induction of ARE-regulated gene expression by sulforaphane 90
Sulforaphane-induced stabilization and nuclear accumulation of Nrf2 90
Sulforaphane does not trigger the ubiquitination of Keap1 as tBHQ does 95
Stability of sulforaphane cysteine adducts 97
Concentration-dependent formation of sulforaphane cysteine adducts on Keap1 104
Discussion 106

CHAPTER FOUR-DISCUSSION 111

APPENDIX 1 118

APPENDIX 2 142

REFERENCES 161

LIST OF FIGURES AND SCHEMES

Figure 1-1. Sequence of Keap1 protein and its five domains 26

Figure 1-2. Keap1-Nrf2 signaling pathway under basal and under stress conditions 29

Figure 1-3. Structures of typical phase II inducers 33

Figure 1-4. Analytical proteomics overview 36

Figure 1-5. Spectra of TpepC and iodoacetamide modified TpepC 41

Figure 2-1. Structures of IAB and BMCC 57

Figure 2-2. Activation of Nrf2 and ARE gene expression induced by IAB in HEK293 cells. 61

Figure 2-3. Transfected HEK 293 cells stably overexpress Flag-Keap1 63

Figure 2-4. Effect of IAB, BMCC and tBHQ on nuclear Nrf2 and cytosolic HO-1 protein levels 64

Figure 2-5. Formation of HMW keap1 protein forms in HEK293 cells stably expressing FLAG-Keap1 treated with electrophiles 65

Figure 2-6. IAB itself could not trigger the formation of HMW keap1 in vitro 67

Figure 2-7. Concentration-dependent formation of HMW Keap1 by electrophiles 68

Figure 2-8. Distribution of Keap1 and ubiquitin peptide identifications in SDS-PAGE gel fractions from control and IAB-treated transfected cells 69

Figure 2-9. MS-MS spectrum of Gly-Gly-modified ubiquitin peptide containing the modification site on Lys48 71

LIST OF FIGURES AND SCHEMES – *Continued*

Figure 2-10. MS-MS spectrum of Gly-Gly-modified Keap1 peptide containing the modification site on Lys298 73

Figure 2-11. Nrf2 stabilization coincides with electrophile-specific adduction and ubiquitination of Keap1 in vivo 74

Figure 2-12. Proposed mechanism of Nrf2 activation by electrophilic inducers of Nrf2-dependent gene expression 80

Figure 3-1. Activation of ARE gene expression induced by sulforaphane in HepG2 cells 91

Figure 3-2. Sulforaphane caused a concentration and time dependent stabilization of Nrf2 93

Figure 3-3. Sulforaphane induced expression of HO-1 and sulforaphane inhibited degradation of Nrf2 was enhanced by proteasome inhibitor MG-132 94

Figure 3-4. tBHQ, but not sulforaphane triggers the ubiquitination of Keap1 96

Figure 3-5. Longer incubation and higher temperature facilitate the reversion of sulphoraphane cysteine adduct to sulforaphane and free cysteine 100

Figure 3-6. MS-MS spectrum of sulforaphane modified TpepC 101

Scheme 3-1. Sulforaphane cysteine adduct formation and hydrolysis of the adduct to sulforaphane and free cysteine. 98

Scheme 3-2. Experimental procedures utilized to prepare samples for LC-MS-MS analyses 99

LIST OF TABLES

Table 2-1. Sites of modification of Keap1 by electrophiles *in vitro* 58

Table 2-2. Modification of Keap1 by IAB as a function of time *in vitro* 60

Table 3-1. Procedures to optimize workup conditions to identify SFP adduct sites by LC-MS-MS 103

Table 3-2. Modification of Keap1 by sulforaphane as a function of concentration *in vitro* 105

LIST OF ABBREVIATIONS

ACF	aberrant crypt foci
ACN	acetonitrile
AMBIC	ammonium bicarbonate
ARE/EpRE	antioxidant response element/electrophile response element
tBHA	*tert*-butylhydroxyanisole
tBHQ	*tert*-butylhydroquinone
BMCC	1-biotinamido-4-(4'-[maleimidoethyl-cyclohexane]-carboxamido)butane
CBP	CREB binding protein
CID	collision induced dissociation
CNC	cap'n'collar
COX	cyclooxygenase
cDNA	complementary DNA
Da	Dalton
Dex-mes	dexamethasone-21-mesylate
DMSO	dimethylsulfoxide
D3T	3H-1,2-dithiole-3-thione
DTT	dithiothreitol
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization

LIST OF ABBREVIATIONS –*Continued*

FBS	fetal bovine serum
g	force of gravity (e.g., 10000 x g)
GCS	γ-Glutamylcysteinyl synthetase
GFP	green fluorescent protein
GSH	glutathione
GST	glutathione-S-transferase
HEK293	human embryonic kidney cell line
His	histidine
HMW	high molecular weight
HO-1	heme oxygenase 1
HPLC	high performance liquid chromatography
IAB	N-iodoacetyl-N-biotinylhexylenediamine
Keap1	Kelch-like ECH-associated protein 1
LC-MS-MS	liquid chromatography-tandem mass spectrometry
IgG	immunoglobulin
kD	kilodalton
MALDI	matrix-assisted laser disorption ionization
MWCO	molecular weight cutoff
MS	mass spectrometry
MS/MS	tandem mass spectrometry

LIST OF ABBREVIATIONS –*Continued*

m/z	mass to charge ratio
NES	nuclear export signal
NQO1	NAD(P)H:quinone oxidoreductase 1
2D-PAGE	two dimensional polyacrylamide gel electrophoresis
PAGE	polyacrylamide gel electrophoresis
PAR	4-(2-pyridylazo)resorcinol
PBS	phosphate buffered saline
15d-PGJ_2	15-Deoxy-$\Delta^{12,14}$-prostaglanin J_2
PI3K	phosphaidylinositol 3-kinase
PKC	protein kinase C
PMA	phorbol 12-myristate 13-acetate
Q-TOF	quadrupole-time of flight
ROS	reactive oxygen species
SCF	Skp1p-cullin-F-Box protein
SCX	strong cation exchange
SDS-PAGE	sodium dodecyl sulfate-PAGE
SFP	sulforaphane
TCEP	tris-(2-carboxyethyl)phosphine hydrochloride
TOF	time of flight
TpepC	peptide, sequence AVAGCAGAR

LIST OF ABBREVIATIONS –*Continued*

TRE	12-O-tetradecanoylphorol-13-acetate response element
bZIP	basic-leucine zipper

ABSTRACT

Activation of the transcription factor Nrf2 regulates expression of phase II enzymes and other adaptive responses to electrophile and oxidant stress. Nrf2 concentrations are regulated by the thiol-rich sensor protein Keap1, which is an adaptor protein for Cul3-dependent ubiquitination and degradation of Nrf2. However, the links between site-specificity of Keap1 modification by electrophiles and mechanisms of Nrf2 activation are poorly understood. We studied the actions of the prototypical Nrf2 inducer *tert*-butylhydroquinone (tBHQ) and two biotin-tagged, thiol-reactive electrophiles N-iodoacetyl-N-biotinylhexylenediamine (IAB) and 1-biotinamido-4-(4'-[maleimidoethyl-cyclohexane]-carboxamido)butane (BMCC). Both IAB and tBHQ induce expression of ARE-directed GFP expression in ARE/TK-GFP HepG2 cells and both initiatednuclear Nrf2 accumulation and induction of heme oxygenase 1 in HEK293 cells. In contrast, BMCC produced none of these effects. Liquid chromatography tandem mass spectrometry (LC-MS-MS) analysis of human Keap1 modified by IAB or BMCC *in vitro* indicated that IAB adduction occurred primarily in the central linker domain, whereas BMCC modified other Keap1 domains. Treatment of FLAG-Keap1 transfected HEK293 with the Nrf2-activating compounds IAB and tBHQ generated high molecular weight Keap1 forms, which were identified as K-48-linked polyubiquitin-conjugates by immunoblotting and LC-MS-MS. Keap1 polyubiquitination coincided with Nrf2 stabilization and nuclear accumulation. In contrast, BMCC did not induce Keap1 polyubiquitination. Our results suggest that Nrf2 activation is regulated through the

polyubiquitination of Keap1, which in turn is triggered by specific patterns of electrophile modification of the Keap1 central linker domain. These results suggest that Keap1 adduction triggers a switching of Cul3-dependent ubiquitination from Nrf2 to Keap1, leading to Nrf2 activation.

The chemopreventive agent sulforaphane is an isothiocyanate, which was isolated from broccoli. Sulforaphane was demonstrated to induce ARE-regulated genes by stimulating the Keap1-Nrf2 system. This agent is a powerful electrophile that can react with thiols to form thionoacyl adducts. A specific sulforaphane modification pattern on Keap1 may trigger the activation of Nrf2. However, thionoacyl adducts are labile to hydrolysis and transacylation reactions, which prevent the identification of the sulforaphane modification patten on Keap1. In this study, we have developed a LC-MS-MS method to map sulforaphane modification sites formed on Keap1 *in vitro*. Our studies indicate that sulforaphane displays a different pattern of Keap1 modification than ARE/ERE inducers that modify Keap1 by alkylation. Moreover, the modification of Keap1 *in vivo* by sulforaphane does not trigger the ubiquitination of Keap1, which suggests a novel mechanism for Nrf2 stabilization by sulforaphane thionoacyl adduct formation.

CHAPTER ONE – INTRODUCTION

Chemoprevention of cancer

Numerous studies have demonstrated that environmental chemicals cause a variety of cancers and suggested that risk of cancers could be dramatically decreased by avoiding exposure to these carcinogens. However, it is not possible and practical for humans to escape from daily exposure to diverse carcinogens. Therefore, chemoprevention, a cancer-preventive strategy, was proposed as an effective approach to reduce cancer risk. Cancer chemoprevention was defined as pharmacological intervention with synthetic or natural compounds that may prevent, inhibit, or reverse carcinogenesis, or prevent the development of cancer (*1*). Plenty of evidence has suggested that there is a reverse relationship between cancer risk and a diet rich in fruits and vegetables (*2*). Generally, diets high in vegetables and fruits (400g/day) may prevent more that 20% of all cancers and vegetables appear to be more effective that fruits (*3*). For example, crucifer intake of around 3 servings per week decreases prostate cancer risk by 41% (*4*).

A number of natural products with inhibitory effects on carcinogenesis have been identified from sources of our diet. These compounds include isothiocyanates from cruciferous vegetables, catechins from green tea, resveratrol from grape seeds, lycopene from tomato, procyanidins from various fruits and nuts, isoflavones from soybean, and antioxidant vitamins in various food (*5;6*). Based on concepts of mechanisms leading to cancer at cellular and molecular levels, synthetic chemopreventive agents have been developed for cancer prevention based on selective action on specific molecular targets, such as estrogen and androgen receptors and inducible cyclooxygenase (*1;7*). These

natural or synthetic agents are able to retard or reverse carcinogenetic processes by either interrupting interaction between carcinogens and normal cells, by slowing the proliferation rate or by inducing apoptosis in cancer cells (*8*)

One of the successful strategies for cancer prevention is to protect DNA and other important cellular molecules by enhancing the detoxication of chemical carcinogens and oxidative stressors. Many carcinogens require activation by phase I reactions to be electrophilic intermediates to attack macromolecules within cells (*9*). In contrast to phase I enzymes, phase II enzymes often detoxify these reactive electrophiles by conjugating them with hydrophilic endogenous ligands, such as glutathione and glucuronic acid, so that these carcinogenic intermediates could be easily eliminated (*10*). Many drugs and natural products increase the expression of phase II enzymes and these agents have been classified into nine structurally diverse classes: isothiocyanates, dithiolethiones, oxidizable hydroquinones, Michael reaction acceptors, trivalent arsenicals, heavy metals, hydroperoxides, vicinal dimercaptans, and caroteinoids (*11-13*). Even though these chemicals are structurally different, these chemicals share common properties of electrophilicity (or the ability to be converted to reactive electrophiles) and the capacity to modify sulfhydryl groups.

Chemoprevention by isothiocyanates

Of identified naturally occurring chemopreventive compounds, isothiocyanates have attracted significant attention. Chemoprotective effects of isothiocyanates were first observed in experiments where Wistar rats were fed α-naphthyl-isothiocyanate decades

ago (*14*). Since then, many studies have demonstrated that these compounds exert potent anticarcinogenic properties in carcinogen-induced tumorigenesis in a variety of experimental animal models (*15-17*). Isothiocyanates are found in cruciferous vegetables such as broccoli, Brussels sprouts, cabbage and cauliflower. All cruciferous plants contain high amount of glucosinolates, which exist as an N-hydroxysulfate with a sufur linked β-glucose and a variable side chain containing alkyl, alkenyl, aromatic, indolyl, or other moieties (*18*). The sugar can be hydrolyzed by myrosinase, a plant specific enzyme that is released under physical stress. The remaining intermediate is unstable and converted into different biologically active isothiocyanates such as sulforaphane [1-isothiocyanato-4-(methylsulfinyl)-butane].

Sulforaphane was demonstrated to have powerful anticancer effects. Epidemiological studies have shown a connection between consumption of broccoli and reduced risk of colon cancer in individuals with the glutathione S-transferase M1 (GSTM1) null genotype(*19*). Sequencial studies reported that sulforaphane significantly reduced the formation of colonic aberrant crypt foci (ACF) induced by azoxymethane in F344 rats (*20*). Cell growth arrest and apoptosis were observed when androgen-dependent human prostate cancer LNCaP cells line was treated with sulforaphane (*21*). Sulforaphane can also brock benzo[α]pyrene-evoked forestomach tumors in ICR mice by inducing phase II detoxication and antioxidant enzymes (*22*). Cancer prevention by sulforaphane was associated with its ability to inhibit phase I gene expression and to induce apoptosis and expression of phase II enzymes. Although sulforaphane is not a direct antioxidant like GSH, it can elevate GSH levels in various cultured cells and

animal models, and activate transcription of phase II genes including GST, NAD(P)H:quinone oxidoreductase (NQO1), UDP-glucuronosyltransferases, glutamylcysteine ligase, epoxide hydrolase, heme oxygenase 1 (HO-1) (*23;24*). These enzymes protect cells against carcinogen by scavenging and detoxifying reactive oxygen species (ROS) and harmful byproducts of oxidative stress, such as lipid and DNA base hydroperoxides.

Induction of phase II enzymes through the ARE regulated by Nrf2

Transcription of many phase II genes depends on activation of a *cis*-acting enhancer sequence called the antioxidant response element (ARE). To understand the mechanism by which the Ya subunit of rat GST (GST A2) is induced, studies seeking to identify the regulatory element controlling expression of this gene were conducted (*25-27*). An element termed as ARE in the 5'-flanking region of the rat GST A2 gene was determined to be essential for gene expression and inducibility. A number of subsequent studies detected ARE sequences in the promoters of various genes that encode phase II enzymes, enzymes involved in glutathione biosynthesis, and genes that encode other proteins that protect cells against oxidative stress (*28-32*). Later the consensus sequence of the ARE was proposed by Wasserman and Fahl (*33*) to be (A/G)TGA(C/T)nnnGC(A/G) and this sequence has been widely accepted as the core sequence of ARE in many studies. However, not all ARE-containing genes respond to typical ARE-mediated inducers, such as human GSTP1 (*34*). A comprehensive mutational analysis of the ARE of murine *nqo1* gene indicates that there is a greater

variability in the core sequence (*35*). These results indicate that more factors contribute to the functional AREs. Since AREs plays an important role in regulating many phane II genes, many studies were carried out to identify the transcriptional factors binding to the ARE consensus sequence. Since ARE sequences in genes such as *nqo1* (TGACTCAGC) and *gstp* (TGATTCAGC) are similar to the TRE (12-O-tetradecanoylphorbol-13-acetate response element) (TGAC/GTCA), the AP-1 complex that contains c-Fos and Jun-D was proposed as the ARE-binding protein complex in early studies (*36;37*). However, instead of activating ARE, binding of AP-1 factors to ARE apparently prevents the binding of other transcription factors to the same site. Subsequent studies indicated that overexpression of AP-1 factors represses the expression of an ARE reporter gene in human hepatoma HepG2 cells (*38*) and deletion of the c-fos gene in mice significantly increases NQO1 and GST activities in several murine tissues (*39*). At the same time, other novel trans-acting proteins were found to bind to the ARE sequence. Studies by Venugopal have shown that Nrf1 and Nrf2 could bind to ARE from human *nqo1* and overexpression of these factors transactivate ARE-regulated reporter activity in HepG2 cells and expression of the reporter were inducible by β-naphthoflavone and *tert*-butyl hydroquinone (*38*). Nrf1 and Nrf2, 66 to 68 kDa proteins, were originally cloned using a yeast complementation assay (*40;41*). These transcription factors are significantly homologous to NF-E2 and are ubiquitously expressed. Homologous disruptiona of the *nrf1* gene in mice caused fetal death because of severe anemia, whereas *nrf2*$^{-/-}$ mice were viable without phenotypic changes (*42*). These results indicate that Nrf2 is not an essential transcription factor for murine erythropoiesis, growth and development.

Recently, Nrf3 was cloned and sequenced (*43*). Nrfs are members of the basic-leucine zipper (bZIP) family of transcription factors and a cap'n'collar (CNC) region is highly conserved among the Nrfs. By aligning human Nrf2 with chicken ECH proteins, six highly conserved regions were identified, which were named as Neh1 to Neh6 (Nrf2-ECH homology) (*44*). In this study, it was found that Neh1 corresponds to the bZIP domain and the Neh2 domain serves as a negative regulatory domain of Nrf2 function by binding to the Keap1 protein, which will be discussed later. It was observed by the Katoh group that a basic transcription machinery recruiter, CREB binding protein (CBP), can bind to Nrf2 through Neh4 and Neh5 domains (*45*).

Yamamoto and colleagues noticed that, through comparison of consensus binding sequences with AREs, Nrf2 might be an important regulatory factor to activate ARE-dependent genes (*46;47*). They also found that Nrf2 bound to ARE with high affinity as a heterdimer with a small muscle aponeurotic fibrosarcoma (Maf) protein. At the same time, they demonstrated that regulation of phase II genes through ARE is Nrf2 dependent. *Nrf2* knockout mice significantly abrogated the inducible expression of GST and NQO1 by antioxidant tert-butylhydroxyanisole (t-BHA) in liver and intestine (*47*). This study originally provided evidence that Nrf2 is a critical factor in the regulation of phase II genes. The Mulcahy group has demonstrated that γ-glutamylcysteine synthetase subunit gene is regulated by the binding of Nrf2 to an ARE sequence and that upregulation of this gene is inducible by β-naphthoflavone (*28;48;49*). Overexpression of Nrf2 increased the activity of reporter proteins constructed downstream of ARE from HO-1, indicating this gene is regulated by Nrf2 (*50*).

Role of Nrf2 in chemoprevention

Since Nrf2 was found to regulate ARE-driven phase II genes, many studies were carried out to investigate role of Nrf2 in cancer chemoprevention. Additional genes were identified to be regulated by Nrf2 using $Nrf2^{-/-}$ mice treated with phase II enzyme inducers. It was demonstrated that expression of GSTs (GSTA, GSTM, and GSTP) and NQO1 were lower in small intestine isolated from $Nrf2^{-/-}$ mice than from wild-type mice. γ-Glutamylcysteinyl synthetase (GCS(h)) was found to be increased dramatically in the small intestine of wild type mice feed with butylated hydroxyanisole (BHA) or oltipraz (4-methyl-5-[2-pyrazinyl]-1,2-dithiole-3-thione), while inducibility of this gene was lost in $Nrf2^{-/-}$ mice (*51*). By using the synthetic chemopreventive agent D3T (3H-1,2-dithiole 3-thione) to treat $Nrf2^{-/-}$ mice, inducible expression of GSTA2, NQO1, epoxide hydrolase, γ-glutamate cysteine ligase regulatory chain, UDP-glucuronosyl-transferase 1A6 and HO-1 were decreased or blocked in the liver (*52*). Constituent transcription of GST A1 and GST P1, epoxide hydrolase and ferritin heavy chain were significantly less in livers from $Nrf2^{-/-}$ mice than that from wild-type mice (*52*). A similar study has shown that constitutive activity of GST and NQO1 were reduced by 50-80% in the liver and intestine from $Nrf2^{-/-}$ mice than those from wild-type mice. In the same study, these enzymes could be induced up to 2-5 fold by oltipraz in wild-type mice, whereas inductibility was abrogated in $Nrf2^{-/-}$ mice (*53*). In a study by Thimmulappa et al.(*54*), sulforaphane was used to treat wild-type and $Nrf2^{-/-}$ mice for seven consecutive days and gene expression patterns in the small intestine were analyzed by microarray. They found that sulforaphane increased fifty genes in wild-type mice compared in $Nrf2^{-/-}$ mice and most

of the genes are related to metabolism of electrophiles and free radicals. Because of decreased expression of detoxifying and antioxidant enzymes, Nrf2$^{-/-}$ mice are dramatically more susceptible to toxic environmental chemicals and stress. Nrf2$^{-/-}$ mice had significantly more gastric tumors after treatment with benzo-[a]pyrene than wild-type mice. Moreover, oltipraz dramatically reduced gastric tumors in wild-type mice by 55%, but had no effect in Nrf2$^{-/-}$ mice (*53*). Nrf2$^{-/-}$ mice were susceptible, showing acute respiratory distress syndrome, to the administration of antioxidant butylated hydroxytoluene (*55*). Also, Nrf2$^{-/-}$ mice were more sensitive to hepatotoxicity induced by acetaminophen than wild-type mice (*56;57*). In the lung of Nrf2$^{-/-}$ mice, hyperpermeability, macrophage inflammation and epithelial injury due to hyperoxic stress were greater than in wild-type mice (*58;59*). DNA adduct formation was significantly greater in Nrf2$^{-/-}$ mice than in wild-type mice when these animals were exposed to benzo[a]pyrene and diesel exhaust (*60;61*). Another study with typical phase II enzyme inducer, sulforaphane, also has shown that this chemopreventive agent could prevent benzo[a]pyrene-induced stomach cancer (*22*). In summary, these studies strongly indicate that Nrf2 is a critical regulator for the expression of phase II detoxifying and antioxidant enzymes and that induction of these enzymes by chemopreventive chemicals is mediated though Nrf2.

In addition to phase II detoxifying enzymes, other genes associated with protection against environmental electrophiles and oxidative stress have been identified to be regulated by Nrf2 (*54;62-64*). Thioredoxin, type 1 peroxiredoxin and thioredoxin reductase 1 were increased in the liver from mice treated with D3T and in cortical

astrocytes treated with tBHQ. Through the Nrf2 pathway, catalase-1 and Cu/Zn superoxide dismutase were upregulated by tBHQ in astrocytes. Expression of the antioxidative genes HO-1 and ferritin was increased in small intestine from $Nrf2^{-/-}$ mice fed with sulforaphane. Other notable Nf2 regulated genes associated with enhanced reducing potential were NADPH-generating enzymes such as glucose 6-phosphate dehydrogenase, and other enzymes in this pathway such as transaldolase, transketolase etc. Increased NADPH can enhance the catalytic activity of phase II detoxifying enzymes.

Keap1 as a negative regulator of Nrf2

Keap1 (Kelch-like ECH-associated protein 1) was isolated from mouse cDNA and identified as an inhibitory binding partner of Nrf2 through binding to Neh2 domain of Nrf2 (*44*). Keap1 is structurally similar to Drosophila actin-binding protein Kelch (*65*). Keap1 proteins are highly conserved between mouse, rat and human. Keap1 contains five domains: an N-terminal domain, a BTB domain, a central linker domain, a Kelch repeat domain and a C-terminal domain (Figure 1-1). The Kelch repeat domain is an actin-binding domain and binding of Keap1 to Nrf2 is through this domain (*44;66*). The BTB domain was identified to be responsible for the dimerization of Keap1 based on the observation that mutation of Ser104 in BTB domain disrupted Keap1 dimerization (*67*).

A critical role of Keap1 in regulating Nrf2 dependent gene expression was recognized in a study by Yamamoto group using $keap1^{-/-}$ mice (*68*). These mice were susceptible to postnatal death due to malnutrition resulting from hyperkeratosis in the esophagus and forestomach and this phenotype could be overcome by breeding $keap1^{-/-}$

mice to Nrf2$^{-/-}$ mice. Expression of phase II enzymes such as GSTs and NQO1 in livers of keap1$^{-/-}$ mice was much higher than that in wild-type mice. These results clearly indicate that Keap1 plays an inhibitory role in controlling Nrf2 regulated genes.

Keap1 is a cysteine-rich protein and human Keap1 contains 27 cysteines among 624 amino acids (*44*). This characteriatic of Keap1, together with the fact that phase II enzyme inducers are electrophiles, lead to a hypothesis that thiol modification on some specific cysteines in Keap1 might change Keap1 conformation and thereafter cause the release of Nrf2. To examine this hypothesis, the model electrophile dexamethasone 21-mesylate (Dex-mes) was used to identify the most reactive cysteines in mouse Keap1. Four cysteines in the central linker domain, Cys257, Cys273, Cys288 and Cys297, were found by mass spectrometry analysis to be the most reactive residues toward Dex-mes adduction (*69*). Subsequent studies by Wakabayashi et al. from the same group have demonstrated that Cys273 and Cys288 of Keap1 are the critical sensors to electrophilic inducers (*70*). Point mutations on these two residues abrogated inhibitory effects of Keap1 on Nrf2 activation, indicating Cys273 and Cys288 play an important role in controlling Keap1-Nrf2 interaction. Analysis by 2D SDS PAGE indicated apparent formation of a Keap1 protein with a molecular weight of a Keap1 dimer in cells treated with phase II inducers such as D3T or sulforaphane. Based on these results, the authors proposed that, following treatment with inducers, the reactive residues Cys273 and Cys288 form intermolecular disulfide bonds, leading to conformational change of Keap1 and thereafter dissociation of Nrf2 from Keap1. However, the exact mechanism by

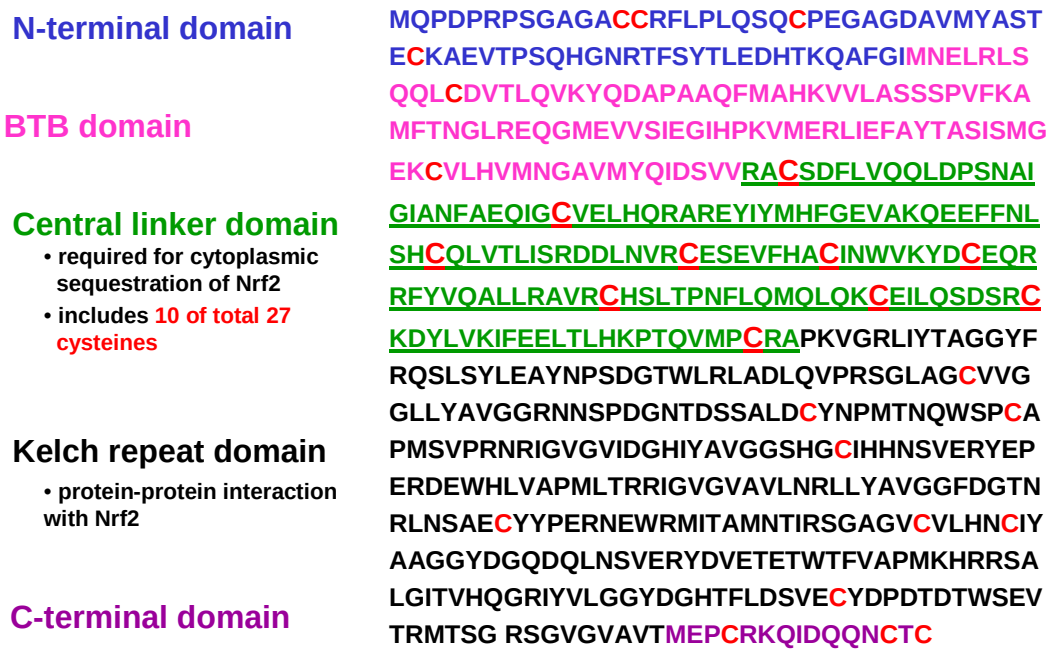


Figure 1-1. Sequence of Keap1 protein and its five domains: N-terminal domain (blue), BTB domain (pink), Central linker domain (green), Kelch repeat domain (black) and C-terminal domain (purple). Cysteine residues are highlighted in red.

which phase II inducers activate Nrf2 by modifying Keap1 is still unknown. Further studies need to be done to address this problem.

It was proposed by Sekhar that Nrf2 could undergo proteasomal degradation (*71*). In this study, inhibition of proteasome activity increased expression of the catalytic subunit of γ-glutamate cysteine ligase in HepG2 cells. This effect coincided with an increased amount of Nrf2 bound to the ARE sequence of this gene. The Kensler group later reported that levels of Nrf2 were increased in whole cell lysates from murine keratinocytes treated with protease inhibitor MG-132 (*72*). This result indicates that Nrf2 could be degraded by 26S proteasome and it was suggested that Keap1 is associated with Nrf2 degradation. This hypothesis was confirmed by the studies from the Yamamoto group (*44*), who demonstrated that the Phase II inducer diethylmaleate could dramatically extend the half-life of Nrf2 in cells compared to vehicle control. Moreover, compared to control peritoneal macrophages, there was a nuclear accumulation of Nrf2 in keap1$^{-/-}$ cells. Thus, it was hypothesized that association of Nrf2 with Keap1 causes proteasomal degradation of Nrf2 in the cytoplasm. In a sequential study by MacMahon (*73*), it has been shown that, by interacting with Keap1 through the Neh2 domain, Nrf2 undergoes ubiquitination and proteasomal degradation (Figure 1-2). Further studies by Zhang and Hannink (*74*) demonstrated that point mutation on Cys273 or Cys288 in Keap1 abrogates ubiquitination and degradation of Nrf2. Nrf2 also could be stabilized in cells treated with sulforaphane and tBHQ, leading to the hypothesis that modification of Keap1 by electrophilic phase II inducers inhibits ubiquitination of Nrf2, so that the levels of Nrf2 in cells increase (Figure 1-2). In the same study, another cysteine residue, Cys151, in the

BTB domain of Keap1 was identified to play an important role in regulating Nrf2 stability. Point mutation on Cys151 abrogated sulforaphane- or tBHQ-induced inhibition of Nrf2 degradation. Later studies by several groups including the Hannink group have found that Keap1 functions as an adaptor for Nrf2 ubiquitination and degradation by the Cullin 3-Roc1 ligase (*75-78*). In the studies by Zhang (*75*), Keap1 was found to be converted to a high mass form upon treatment with phase II inducers, which was coincident with triggered inhibition of Nrf2 degradation by these agents. Keap1 with mutation at Cys151 is dramatically resistant to this inhibition and the same mutation prevents inducible formation of HMW keap1, which suggests this residue is critical for the response of the Keap1-Nrf2 system to inducers.

Other mechanisms involved in Nrf2 activation

Other signaling pathways have been identified as interacting with Keap1-Nrf2 for the activation of Nrf2 dependent genes. The Kong group first found that the MAKP cascade is involved in Nrf2-ARE pathway. Typical phase II enzyme inducers, tBHA and its metabolite tBHQ, stimulate ERK2 and JNK1 activities in different murine cells (*79*). Upregulation of an ARE-driven reporter gene triggered by sulforaphane and tBHQ can be weakened by treatment with the MEK1 inhibitor and by overexpression of dominant negative mutant ERK2 (*80;81*), which suggests that a MAP kinase cascade in involved in the positive regulation ARE-driven genes in response to phase II inducers.

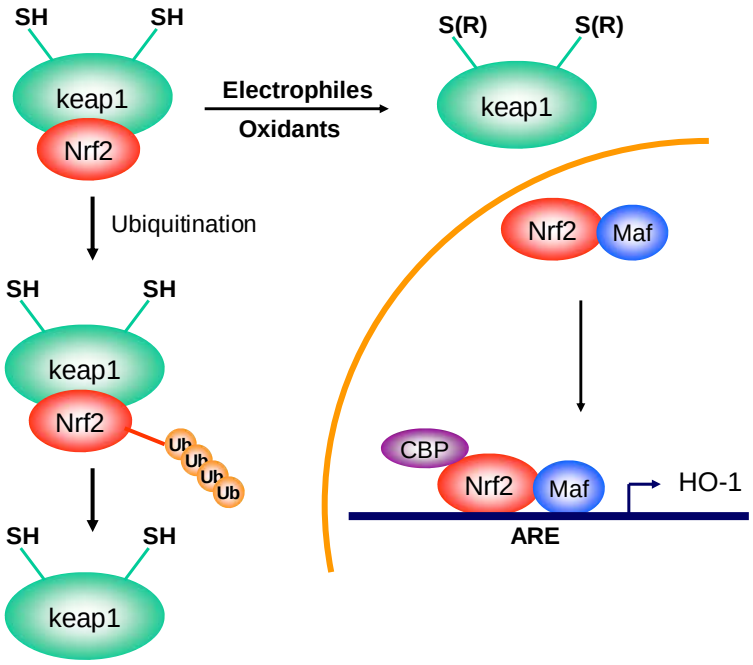


Figure 1-2. Keap1-Nrf2 signaling pathway under basal and under stress conditions.

Phosphatidylinositol 3-kinase (PI3K) was also found to affect ARE-regulated gene expression. tBHQ treatment of rat hepatoma H4IIE cells stimulated Nrf2 translocation into nuclei, leading to the induction of GST A2, which was prevented by pretreating the cells with PI3-kinase inhibitors, wortmannin or LY294002 (*82-84*). Sequential studies by Johnson group have shown the same result in another model with human neuroblastoma IMR32 cells (*85*). Using microarray analysis, this group identified a cluster of genes affected by PI3K in IMR32 cells. Expression of 49 out of 63 tBHQ-inducible genes was significantly blocked by inhibition of PI3K (*86*).

Another signaling pathway identified to influence Keap1-Nrf2 system is protein kinase C (PKC). It was found that transcription of ARE-regulated reporter gene was activated by phorbol 12-myristate 13-acetate (PMA) in human and rat hepatoma cells. However this activation was suppressed by inhibitors of PKC, including staurosporine and Ro-32-0432 (*87*). Subsequent studies by the same group (*88*) showed that Nrf2 was phosphorylated upon tBHQ treatment and that Ser40 was identified to be the phosphorylation site. Phosphorylation of wild-type Nrf2 by PKC promoted Keap1-Nrf2 dissociation, whereas point mutation of Ser40 in the Neh2 domain of Nrf2 blocked this effect on the Keap1-Nrf2 interaction. Phosporylation of Ser40 was confirmed in a study by the Jaiswal group (*89*) and they also demonstrated that phosphorylation of this residue is necessary for the escape of Nrf2 from Keap1, but is not required for Nrf2 accumulation in the nuclei.

Typical phase II inducers and their thiol reactivity

The role of sulforaphane as a chemopreventive agent was introduced above. Studies by Thimulappa (*54*) have identified Nrf2-regulated genes induced by sulforaphane (Figure 1-3) by comparing transcriptional profiles of small intestine from wild-type and Nrf2 knockout mice treated with vehicle or sulforaphane. Nrf2-regulted genes that can be induced by sulforaphane include NQO1, GST, γ-glutamylcysteine synthetase, UDP-glucuronosyltransferases, epoxide hydrolase, as well as many genes not previously associated with phase II metabolism.

Dithiolethiones such as 3H-1,2-dithiole-3-thione (D3T), anethole dithiolethione (5-[p-methooxyphenyl]-1,2-dithiole-3-thione) and oltipraz (4-methyl-5-[2-pyrazinyl]-1,2-dithiole-3-thione) constitute another group of chemopreventive agents (Figure 1-3). These chemicals were demonstrated to inhibit carcinogenicity of many chemicals in multiple target organs and are undergoing preclinical and clinical evaluations for cancer prevention (*90-93*). Cancer prevention by oltipraz was studied in Qidong, China, where the residents are exposed to dietary aflatoxins and are at high risk for the development of liver cancer. Oltipraz significantly increased excretion of phase II conjugated aflatoxins and decreased its oxidative metabolites (*94*). This study confirmed that protective effects of dithiolethiones are associated with the induction of phase II and antioxidant genes including NQO1, GST A1 and γ-glutamylcysteine synthetase (*28;31;95*). Studies with $Nrf2^{-/-}$ mice indicated that induction of phase II genes by dithiolethiones is through the function of Nrf2 (*94*).

Phenolic antioxidant compounds such as 2(3)-tert-butyl-4-dydroxylanisole (BHA) and its active metabolite tBHQ (Figure 1-3) and 3-5-di-tert-bytyl-4-hydroxytoluene (BHT) were first validated as potential chemopreventive agents since they can upregulate carcinogen detoxifying enzymes in animal models (*96;97*). Comparison of responses from full length and deletion constructs of the regulatory region of the GST Ya gene showed that the induction effect on gene expression of these phenolic compounds was controlled by the ARE sequence (*95*).

The compound classes introduced above are typical inducers of phase II enzymes through Nrf2 regulated ARE sequence. These chemicals share the common characteristic of thiol-reactivity. Alkylation of cysteines by sulforaphane is through the electrophilic isothiocyanate group –N=C=S. D3T and other dithiolthiones are disulfides, whereas tBHQ undergoes oxidation to a quinone, which could modify cysteine residues by the Michael addition. Keap1, the cysteine-rich protein, was demonstrated to be a sensor for Nrf2 activation and several cysteines were identified to play an essential role in regulating Keap-Nrf2 association through their redox conditions. Thus it was hypothesized that these phase II inducers interrupt Keap1-Nrf2 association through modifying specific cysteine residues on Keap1. However, there is no study reporting that Keap1 could be modified by the phase II inducers mentioned above. To explore the mechanism of Keap1-Nrf2 dissociation, sites of formation of adducts on Keap1 by these electrophiles are important targets for study.

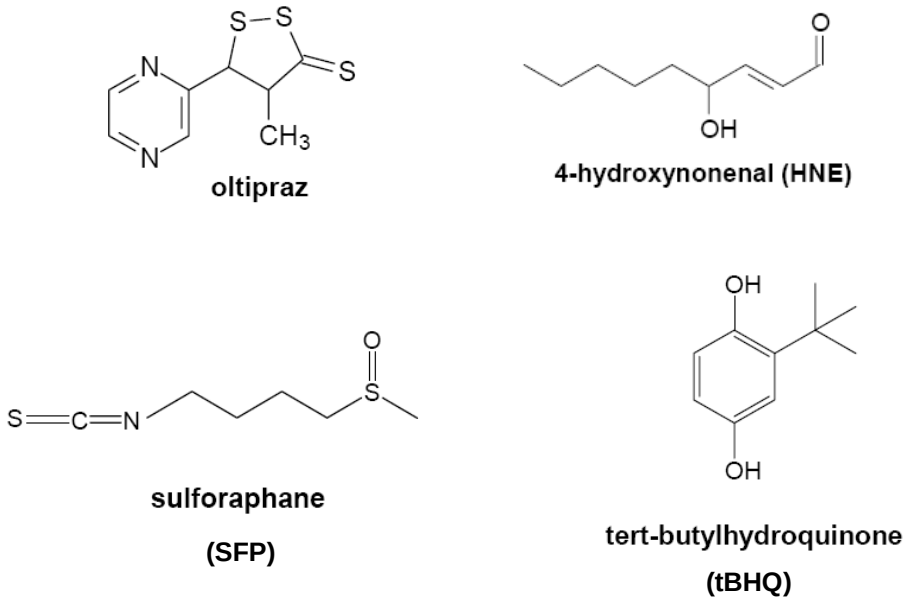


Figure 1-3. Structures of typical phase II inducers.

Analysis of chemical adducts and posttranslational modification sites on proteins

Reactive electrophiles that are generated metabolically from chemical carcinogens can modify cellular macromolecules such as nucleic acids and proteins. The Millers made the first observation that a metabolically activated carcinogenic chemicals bound to a cellular macromolecule by showing that the hepatic carcinogen p-dimethylaminoazabenzene bound covalently to rat liver protein (*98*). Studies of protein modification provides relatively precise dosimetry of chemical carcinogens (*99*).

Previously, measurements of macromolecular adduct required administration of a radiolabeled carcinogen (*100;101*). Later immunological approaches were used to detect specific chemical adduct formation such as HNE or posttranslational modification such as phosphorylation and ubiquitination on proteins (*102*). Combined with site-specific mutagenesis, these tools could be used to identify not only whether specific proteins are modified, but also where the proteins are modified at the amino acid level. However, to use this approach, very specific monoclonal antibodies against reactive chemicals, ubiquitin or phosphorylated residues as serine, threonine and tyrosine should be produced, which takes much time and energy. Moreover, creation of point-mutated proteins is tedious too. Another shortcoming of this method to map modification sites is that amino acid substitutions in mutated proteins may affect protein structure and association of the proteins with others in a protein complex.

Beginning in the 1990's, a number of GC-MS methods were applied to detect and quantify protein adducts formed by methylmethane sulfonate, dimethylnitrosamine, propylene oxide (*103*), benzo[*a*]pyrene (*104*) and styrene (*105*). These methods are

powerful because of their ability to determine the specific chemical structure of the purified protein adduct. However, these methods have their pitfalls too. For detection of hemoglobin or albumin adducts, adducts must be released from the protein, either chemically or enzymatically, before analysis.

Proteomics is a new approach for analyzing proteins and proteomes. Major applications of proteomics in biological studies include mining proteomes, protein expression profiling, identifying protein-protein interactions and protein complex and mapping protein modifications (*106*). Of all tools for proteomic analysis, mass spectrometry (MS) is the most powerful. The general procedure for proteome analysis using mass spectrometry is shown in Figure 1-4. Sample proteins are digested to generate small peptides, which then are analyzed by liquid chromatography-tandem MS (LC-MS-MS). Finally, the MS-MS data are analyzed with bioinformatics algorithms and software. MS provides information about molecular weight, sequences of peptides and modifications on the peptides, and bioinformatics software identifies proteins by assigning raw mass spectrometry data to database sequences. The following discussion is a more detailed discussion of MS-based proteomics approaches for protein identification and protein modification site mapping.

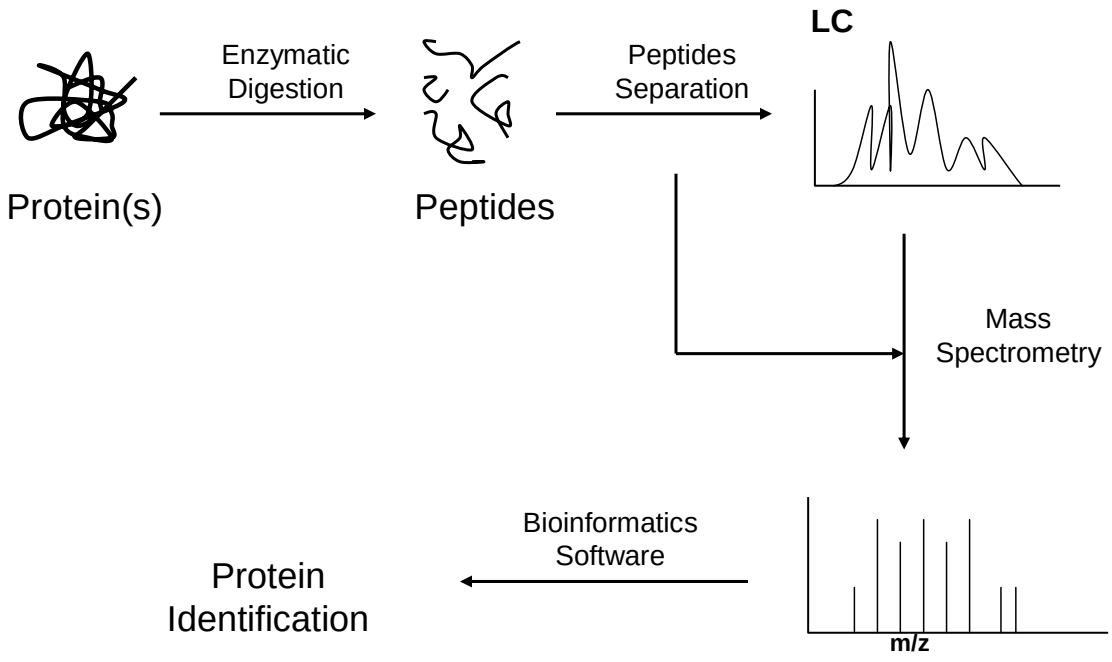


Figure 1-4. Analytical proteomics overview

Protein identification and modification site mapping from MS data and MS-MS data

Mass spectrometry instruments consist of two main parts, a source to generate ions and a mass analyzer to measure mass/charge ratio (*m/z*). Two ionization techniques developed in 1980's, matrix assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) are suitable to analyze large molecules such as proteins and peptides (*107;108*).

To prepare samples for MALDI analysis, the analyte is mixed with a UV-absorbing solid matrix and spotted on a plate. Proteins or peptides are ionized when laser light energy transfers from the matrix to the analyte. MALDI is typically coupled with a time of flight mass analyzer and primarily produces singly charged ions, which are suitable for protein identification by peptide mass fingerprinting (see below).

ESI analyzes peptides or proteins in liquid phase and generally produces multiply charged ions. Ionization is accomplished by placing a large potential difference between a narrow capillary and the mass analyzer inlet to produce a fine spray from an analyte solution. Solvent in the sample evaporates leaving an excess of charge on the shrinking droplet. This process continues until the analyte is completely desolvated but left with a net charge (*109*). ESI is commonly combined with tandem mass analyzers, such as triple quadrupole, ion trap and the quadrupole-time of flight (Q-TOF) analyzers. These mass analyzers can select a single precursor ion for collision-induced dissociation (CID), which produces fragment ions of the precursor ion. Ion traps tend to induce a much more complete fragmentation of the precursor ion than do quadrupoles, which means that more

of the precursor ions are converted more efficiently to product ions in ion traps, giving more sequence information. The fragment ions are analyzed by *m/z* and MS-MS spectrum is produced, which provides peptide sequence and protein identification with database search algorithms or by direct (*de novo*) interpretation of sequence from spectra (*110*).

Since different proteins can share the same nominal mass, identification of intact proteins cannot be accomplished with MS. Identification of proteins depends on analysis of peptides generated by digestion of proteins. The two principal methods for protein identification are peptide mass fingerprinting from full scan MS data and peptide identification by database searching using MS-MS data. With peptide mass fingerprinting, a protein is identified by matching the measured masses of tryptic peptides to corresponding theoretical peptide masses from a database. Within a certain mass tolerance, different peptides may share the same mass. Thus, more accurate *m/z* measurements provide more useful data for peptide mass fingerprinting. To reduce false positive hits, multiple peptide *m/z* values from one protein should be used to search a database for protein identification. There are several software tools available for protein identification by peptide mass fingerprinting, such as PepSea, Mascot, MS-FIT (*111-113*). The general concept of protein identification by these algorithms is similar, which is to assign the highest score to proteins whose predicted peptides match the greatest number of *m/z* values in the MS data. This approach can also be used to identify modified peptides, whose *m/z* would have a mass shift compared to the *m/z* of the unmodified peptide. Peptide mass fingerprinting search tools can accommodate common

modifications such as S-carboxamidomethylation (from protein alkylation by iodoacetamide), phosphorylation, sulfation, glycosylation, etc. The net mass of these molecules on protein can be calculated and the residues these modifications affect can be predicted.

MS-MS data of peptides provides information on peptide sequences. When peptide ions are subjected to collision induced dissociation in a triple quadrupole, a Q-TOF or in an ion trap mass analyzer, they tend to fragment mainly along the peptide backbone (*114*), leading to the formation primarily of y-ions and b-ions. Y-ion fragments contain the C-terminus and b-ion fragments contain the N-terminus of the original peptide ion. Since sixteen of the twenty amino acids have their unique residue mass, peptide sequences can be determined directly from the fragmentation patterns and then protein could be identified based on the sequenced peptides. For example, the synthetic peptide TpepC (AVAGCAGAR), has a predicted b- and y- ion fragmentations illustrated in Figure 1-5. In the same figure, the actual MS-MS spectrum of this peptide is shown, which contains most of the predicted b- and y- ions. When a peptide contains a modified residue,the MS-MS spectrum provides information of both peptide sequence and location of the modification on the peptide. For example, the peptide TpepC is modified by iodoacetamide on the cysteine residue in the middle (Fig. 1-5). Its MS-MS spectrum contains b- and y- ion peaks, which provide information about the sequence of TpepC. This spectrum also shows a mass shift of 57 in the y-ion peaks beginning with the y_5 ion and in the b-ion peaks beginning with the b_5 ion, which indicates that the modification site on TpepC is on the cysteine.

Algorithms and software tools have been developed to identify the proteins from uninterpreted MS-MS spectra of peptides by comparing the actual MS-MS data to peptides sequences in databases. Sequest, which was introduced by John Yates in 1995 (*115-117*) is the first algorithm to identify proteins by matching MS-MS data to theoretical MS-MS spectra generated from nucleotide and protein sequences in a database. A correct identification of proteins with Sequest depends on the quality of the MS-MS data obtained and the completeness and accuracy of the database. Peptide modifications can also be identified and mapped by Sequest, if the adduct mass and amino acid specificity of the modification are specified prior to the database search. Sequest is not able to correctly match spectra of peptides containing unanticipated modifications. To solve this problem, our laboratory under the guidance by Dr. Liebler developed SALSA (Scoring Algorithm for Spectral Analysis) and P-Mod as tools to map unanticipated protein modifications from MS-MS data (*118-120*). SALSA screens tendem-MS spectra for special characteristics, which include product ions, neutral losses, charged losses and ion pairs that are indicative of specific peptide modifications. SALSA was applied for the detection of peptide adducts of electrophiles such as benzoquinone (*121*). As to P-mod (*120*), it facilitates discovery and sequence mapping of modifications to target proteins known to be in the analysis or identified by Sequest. P-Mod screens data files for MS-MS spectra corresponding to peptide sequences in a search list. When a precursor ion mass differs from the expected peptide mass, P-Mod assumes the spectrum corresponds to a modified version of the original sequence. The program then generates an array of search criteria corresponding to the expected MS-MS ions and incorporating

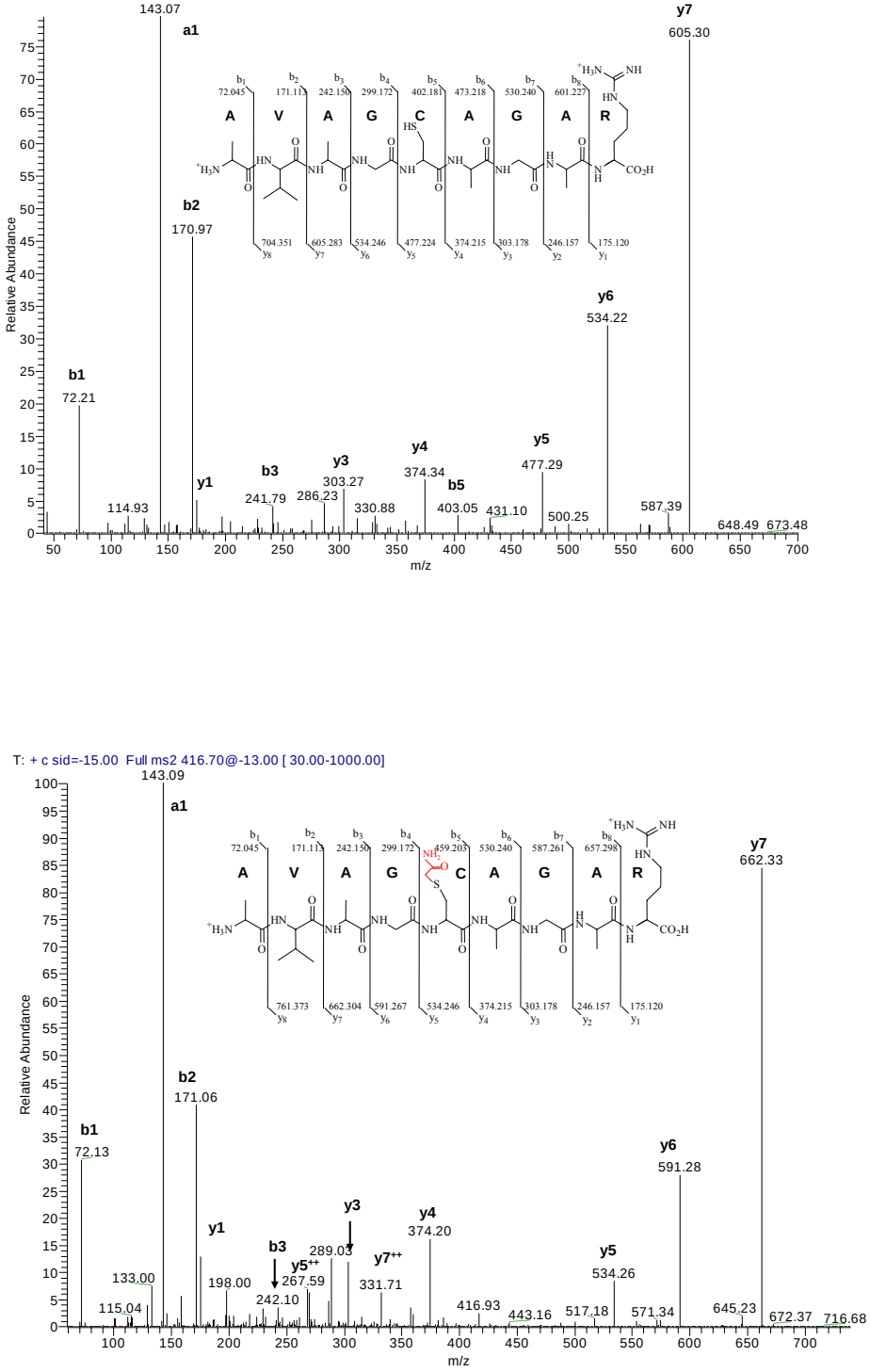


Figure 1-5. Spectra of TpepC and iodoacetamide modified TpepC.

the observed mass shift at all possible sequence positions. By matching each spectrum in the data file to sequences in the search list, P-Mod can localize the modification to a specific amino acid position.

Sample preparation for LC-MS-MS

Different labs use different methods to prepare samples for LC-MS-MS analysis. However, they share the same general approach, which involves denaturation and reduction of the protein, digestion (typically with trypsin) and analysis of the tryptic peptides. The typical procedure used in our lab is: 1) Dissolve proteins to be analyzed in 0.1 M ammonium bicarbonate; 2) Denature the proteins by incubating the proteins at 50^{o}C for 15 min; 3) Reduce disulfide bonds with TCEP or/and DTT at the same time to denature the proteins; 4) Add iodocetamide to alkylate free cysteines; 5) Digest proteins with protease; 6) Add formic acid to acidify peptides.

Digestion of proteins to obtain peptides is necessary for protein identification by MS. Enzymes commonly used to cleave proteins include trypsin (*122*), chymotrypsin, endoproteinase Glu-C (*123-126*), endoproteinase Lys-C (*126*), endoproteinase Asp-N (*127;128*), endoproteinase Arg-C (*129;130*) and carboxypeptidase Y (*131;132*). Of all these enzymes, trypsin is the most commonly used protease, because it is specific to hydrolyze peptide bonds C-terminal to lysine and arginine residues except those right before proline (*122*). Trypsin produces peptides with an average length of about nine residues, which is ideal for generating good MS-MS data.

Sequence coverage is an issue for modification identification

Sequence coverage is referred to as the extent to which the entire protein sequence is represented by MS data. Low coverage would not prevent protein identification. With peptide mass fingerprinting, 2-3 peptides that matched to database are usually sufficient for protein identification. With LC-MS-MS, as few as two good MS-MS spectra of middle-sized peptides are enough to correctly identify a protein through database searching. However, to detect all the modifications on a specific protein, 100% coverage is theoretically required, as modifications could occur anywhere on the protein.

Strategies to improve detection of modified peptides

Protein modification mapping is one of the major applications of proteomics on molecular biology. Characterization of exogenous modifications of target proteins by reactive chemicals assists us to understand the mechanisms of cytotoxicity or cell death or the regulation of signaling pathways. Characterization of endogenous modification on proteins such as phosphorylation, ubiquitination, sumolation or glycosylation, helps to characterize functional states of proteins and to interpret protein-protein interactions, protein structural changes, degradation and to identify biomarkers for biological states. As noted above, protein sequence coverage is important to accurate mapping of modifications. A challenge in mapping modifications is that modified forms may be present in relatively low abundance. This may be true of modifications by reactive electrophiles, which may modify proteins with low stoichiometry under toxicologically

relevant conditions. There are several approaches that can be taken to address this problem.

First, studies with purified proteins can be valuable for characterizing relationships between electrophile chemistry and protein modification selectivity. If a single protein is to be studied, sufficient amounts of a relatively pure form should be employed (>90% purity). The reason for this can be understood in the context of how the LC_MS-MS instrument operates (see above). The ion trap MS instrument allows a certain number of peptides to be analyzed per unit time. When peptides from proteins with high purity are analyzed by mass spectrometry, more target peptides from that protein are subjected to CID. This increases the probability that low abundance peptide forms (e.g., adducts) from that protein are selected for MS-MS. A single protein can be separated from a protein mixture by 1D or 2D SDS-PAGE. A small piece of gel containing the single protein then can be cut and subjected to in-gel digestion and LC-MS-MS analysis of the tryptic peptides.

Affinity chromatography is commonly used to enrich modified proteins or modified peptides for LC-MS-MS analysis. For example, in a cell, only a small fraction of substrates for protein kinases are phosphorylated at a particular time. To solve this problem, based on the knowledge that phosphate groups have high affinity for polyvalent metal cations, a method call IMAC (immobilized metal affinity chromatography) was developed to isolate phosphopeptides from protein digests (*133*). Antibodies against phosphorylated amino acids can also be used to enrich phosphopepetides by immunoprecipitation.

A good HPLC separation of peptides can also increase coverage, because the number of peptides entering the mass spectrometer per unit time decreases and therefore more peptides in the mixture can be analyzed. ESI-MS is commonly combined with liquid chromatography, so that different peptides can be separated and concentrated before MS-MS analysis and MS sensitivity increases (*134*). A two dimensional chromatography technique, employing a strong cation-exchange (SCX) column connected upstream of a C18 column, is used to separate peptides by charge and hydrophobicity (*115;135*). With this method, peptides digested from complicated protein mixture can be separated in two chromatographic dimensions and more peptides can untimately be selected for MS-MS fragmentation.

Statement of the problem

The Keap1-Nrf2 system is a sensor-trigger for adaptation to electrophilic and oxidant stresses. Electrophilic phase II inducers trigger the release of the transcription factor Nrf2 from its inhibitor Keap1, leading to the nuclear translocation of Nrf2 and expression of ARE-regulated detoxifying and antioxidant enzymes (*44;136*). However, the mechanism how inducers activate Nrf2 is poorly understood. Current evidence suggests that modification of Keap1 sulfhydryls is important for Nrf2 activation. A theory proposed by Talalay's group is that inducer modifications on specific cysteines on Keap1 cause disulfide bond formation intermolecularly and thereafter the conformational change of Keap1 and release of Nrf2 (*69*). Recent evidence shows that Keap1 functions as an adapter for Nrf2 ubiquitination and inducers can stabilize Nrf2 by inhibiting

ubiquitin-dependent degradation of Nrf2 (*66;74*). It is also demonstrated that Keap1 contains a nuclear export signal (NES) that functions as a regulator to shuttle Nrf2 from the nuclei (*137*). Inducers may inactivate Keap1's NES and block export of the Keap1-Nrf2 complex from the nuclei, leading to the nuclear accumulation of Nrf2. So far, only a little work has been done on modification specificity, but only with a model compound Dex-mes that was not evaluated as an inducer in an *in vivo* model (*70*). Moreover, inducers of significance such as chemopreventive agents (e.g., sulforaphane) have not been studied. The question of target selectivity for Keap1 modifications and how specific adducts on Keap1 contribute to Nrf2 activation remains unknown. To address this problem and with the assistance of mass spectrometric analyses, we will use different electrophiles, Nrf2 gene inducers or non-inducers, to explore the mechanism how electrophilic phase II inducers can activate and stabilize Nrf2.

CHAPTER TWO – SPECIFIC PATTERNS OF ELECTROPHILE ADDUCTION TRIGGER KEAP1 UBIQUITINAION AND NRF2 ACTVIATION

Introduction

Cellular defense against electrophiles and oxidants relies on detoxication by phase II biotransformation enzymes, antioxidant enzymes and related stress response proteins (*24;138;139*). Many of these inducible genes, such as the glutathione S-transferases, NAD(P)H oxidoreductase, heme oxygenase 1 (HO-1), γ-glutamyl cysteine ligase phase II genes are regulated at the transcriptional level through cis-acting DNA sequences known as antioxidant/electrophile response elements (ARE/EpREs) (*140;141*). Inducers of ARE/EpRE driven genes share the property of electrophilicity (*142*), which enables them to modify sulfhydryl groups in proteins that regulate signaling pathways involved in toxicity and stress. Oxidative stress induced by tBHQ (*tert*-butylhydroquinone) and diethylmalate also can induce ARE-driven genes to restore redox homeostasis and reduce oxidative damage (*141*).

A critical regulator of ARE-dependent transcription is the transcription factor Nrf2, which is a member of the basic-leucine zipper NF-E2 family (*41;143*). Upon formation of heterodimers with one of the small Maf proteins, Nrf2 activates expression of ARE-driven phase II genes (*47-50*). Recent studies with Nrf2$^{-/-}$ mice indicate that Nrf2 regulates a variety of genes, including chaperones, antioxidant genes, and genes regulating protein degradation (*50;53;62*). Nrf2$^{-/-}$ mice are more susceptible to toxic chemicals and stress (*53*).

Keap1 is cysteine-rich cytoplasmic protein that negatively regulates the activation of Nrf2(*44*). Keap1 has five domains, an N-terminal domain, a BTB domain, a central linker domain, a Kelch repeat domain and a C-terminal domain. The Kelch repeat domain is tethered to cytoskeletal actin and binds to Nrf2 directly (*44;66*). The central linker domain is the most cysteine-rich domain and is required for cytoplasmic sequestration of Nrf2 (*74*). Intact cytoskeleton is also critical for allowing Keap1 to bind Nrf2 and prevent its translocation to nuclei (*74;144*).

Recent studies have shown that Keap1 functions as a substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex and targets Nrf2 for ubiquitination and proteasomal degradation (*75-77*). Site-directed mutation of Keap1 led to the hypothesis that the rate limiting step for Nrf2 stabilization resided at the level of Keap1-directed ubiquitination of Nrf2 (Zhang et al, 2004). Dissociation of Nrf2 from Keap1 may represent a second regulatory step that allows Nrf2 to translocate to the nucleus and activate transcription of ARE-dependent genes.

The mechanism by which electrophiles induce Nrf2-dependent transcription is inadequately understood, but appears to involve modification of specific Keap1 cysteine residues. Only four cysteine residues in murine Keap1 (C257, C273, C288 and C297) have been shown to preferentially react with the prototypical alkylating agent dexamethasone 21-mesylate (Dex-mes) *in vitro* (*69*). Mutation of Keap1 at Cys273 or Cys288 impaired the ability of murine Keap1 to repress Nrf2-dependent transcription (*70*). These studies led to the suggestion that dissociation of Nrf2 from Keap1-Nrf2may be regulated by the redox status of these specific cysteine residues.

We employed two cell permeable, thiol-reactive biotin-tagged electrophiles to probe the relationship between Keap1 alkylation and Nrf2 activation. Electrophile-specific adduction patterns were observed in LC-MS-MS analyses of purified human Keap1. Electrophile-specific adduction was coincident with electrophile-specific Keap1 ubiquitination, Nrf2 stabilization, nuclear Nrf2 translocation and ARE/EpRE-dependent gene activation. These novel data provide compelling support for the hypothesis that electrophile-directed Keap1 ubiquitination represents a rate limiting step for electrophile-directed Nrf2-dependent gene activation.

Experimental Procedures

Cell culture, transfection and treatment - Human embryonic kidney 293 cells were obtained frozen at low passage from Master Cell Bank cultures from GIBCO Life Technologies (Grand Island, NY). Cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS), 2mM glutamine, 100μg/mL penicillin and 100μg/mL streptomycin. Human hepatoma HepG2 cells containing an ARE/TK-GFP reporter (ProCertus Biopharm) were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS), 2mM glutamine, 100μg/mL penicillin and 100μg/mL streptomycin and containing 1 mg/mL geneticin. Confluent cells in 100 mm dishes were transfected with pCMV-FLAG-Keap1 using Lipofectamine 2000 reagent at a ratio of 1:6 DNA/Lipofectamine and then were incubated in DMEM for 5 h at 37^{o}C. Medium with 500 μg/mL geneticin was used to select the single cell colonies stably expressing FLAG-Keap1. Confluent cells in 100 mm plates were washed with phosphate buffered saline (PBS) and treated with 100 μM IAB, BMCC or equal volumes of vehicle (DMSO at 0.3% of total volume) were delivered in 4 mL DMEM with 5% FBS.

Cell fractionation and immunoblot analyses - Confluent cells in 75cm flasks were washed with cold PBS and lysed in cold nuclear lysis buffer (5 mM Tris, pH 7.4, 50 mM NaCl, 10 mM EDTA, 0.5 % NP-40) containing 10 μL/mL protease inhibitor cocktail (Sigma Catalog No. P8340). Lysate was centrifuged at 12,000 rpm for 2 min and

supernatant was collected as the cytoplasmic fraction. The pellet was washed twice with nuclear lysis buffer to remove cytosolic contaminants, and nuclear pellet was lysed by treatment with all protein lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% NP-40, 10 mM EDTA, 0.1% SDS, 0.5% deoxycholic acid, pH 8.0) containing 5μL/mL protease inhibitor cocktail, followed by sonication for 5 sec. The nuclear lysate was centrifuged at 13,000 rpm for 2 min to remove cell debris. Protein concentration was measured with BCA protein assay kit (Pierce). Cell lysate proteins were diluted 1:1 (v:v) with 5 x SDS loading buffer and separated by SDS-polyacrylamide gel electrophoresis in 4-20% Tris HCl Ready Gels (Bio-Rad, Hercules, CA). Resolved proteins then were transferred to PVDF membranes, which were blocked with 5% milk in TBST buffer (20 mM Tris HCl, pH 7.5, 200 mM NaCl, 0.1% Tween 20) and then probed with anti-HO-1 (proved by Dr. Chris Ferris), anti-FLAG (Sigma), anti-Keap1 (Santa Cruz), anti-Biotin (Zymed), and anti-Nrf2 (Santa Cruz) in 5% milk in TBST buffer, respectively. After treatment with appropriate secondary antibody conjugated to horseradish peroxidase (Santa Cruz), immunostained proteins were detected by enhanced chemiluminescence with Western blotting luminol reagent (Santa Cruz).

Conforcal microscopic analysis - FLAG-Keap1 transfected HEK293 cells were cultured in poly-L-lysine coated chambers for 24 h. Confluent cells then were fixed with 2% formaldehyde in Hanks' balanced saline solution (HBSS) for 5 min, washed with methanol/0.1% Triton X-100 for 2 min and then washed with HBSS. The cells then were blocked with 1% milk in HBSS for 30 min and incubated with primary antibody (anti-

FLAG) at 80 μg/mL for 3 h at 37 °C. Cells then were washed with HBSS and incubated with secondary antibody (FITC-mouse anti-IgG) in 1% milk in HBSS for 1 h and then washed with HBSS. The chamber slides were mounted in Cytoseal (Stevens Scientific) and sealed under #1.5 cover slips. Images were acquired with Zeiss LSM510 confocal microscope. FITC fluorescence was excited at 488 nm.

Covalent adduction of model electrophiles to His_6-Keap1 in vitro - Escherichia coli BL21(DE3) was transfected with plasmid pET-15b(+)hKeap1, which encodes a full-length cDNA copy of human Keap1 inserted between the *NdeI* and *XhoI* sites. His_6-Keap1 was expressed by the method developed by Dinkova-Kostova (*69*) except that LB medium was used to grow bacteria. To prevent protein aggregation and precipitation by disulfide bond formation, 10 mM β-mercaptoethanol was added to lysate buffer. His_6-Keap1 protein then was purified by Ni-NTA agrose affinity chromatography (Qiagen). To maintain Keap1 solubility, the purified protein was dialyzed against 25 mM Tris-HCl, pH 8.4, containing 5 mM EDTA and 5 mM mercaptoethanol prior to concentration and storage.

Ultrafree-MC low binding regenerated cellulose centrifugal spin filter devices with a 30,000 molecular weight cutoff (MWCO) were obtained from Millipore. Prior to sample addition, spin filters were sequentially rinsed with 200 μL methanol and 200 μL distilled water by centrifugation at 12,000 x g. Keap1 (60-70 μg) was loaded into the upper chamber of the spin filter and the sample then was centrifuged at 12,000 x g to remove the solution. The proteins on the filter were washed with 200 μL 1 M ammonium

bicarbonate followed by centrifugation. The proteins then were suspended in 50 µL of 1 M ammonium bicarbonate (pH 8.4) containing 100 µM IAB or BMCC (Pierce) in 0.3% DMSO (v/v). After incubation at 37 ^{o}C in the dark, the reaction was terminated by centrifugation to remove the buffer followed by an additional wash with 1 M NH_4HCO_3. The filtrates were discarded. The protein on the filter was suspended in 50 µL 0.1 M ammonium bicarbonate containing 12 µL 40 mM *tris*(carboxyethyl)phosphine (Pierce) and incubated at 50 ^{o}C for 15 min. Then 20 µL of 0.2 M iodoacetamide was added to the sample for 15 min to convert free thiols to carboxamidomethyl derivatives. Modifed porcine sequencing grade trypsin (Promega) then was added in a 1:50 protein:trypsin ratio and the sample was incubated at 37 ^{o}C for 18-24 h. Tryptic peptides were collected by centrifugation through the filter at 5,000 x *g* and the filtrate was acidified with 1 µL of concentrated formic acid for LC-MS-MS analyses.

Affinity capture and digestion of FLAG-Keap1 - Cytoplasmic protein fractions first were incubated with protein G agarose (Sigma) in batch format for 3 h at 4 ^{o}C. The sample was centrifuged at 2,000 x g for 2 min and the supernatant was incubated with anti-FLAG agrose (Sigma) overnight. After washing with nuclear lysis buffer (5 mM Tris, pH 7.4, 50 mM NaCl, 10 mM EDTA, 0.5 % NP-40) and TBS buffer (50 mM Tris-HCl/150 mM NaCl, pH 7.6), the FLAG-Keap1 was eluted with 100 µg/mL FLAG peptide in TBS buffer. Eluted proteins were concentrated with an Amicon ultracentrifugal filter device, 10,000 MWCO (Millipore) at 4,000 rpm and then reduced, alkylated and digested with trypsin as described above.

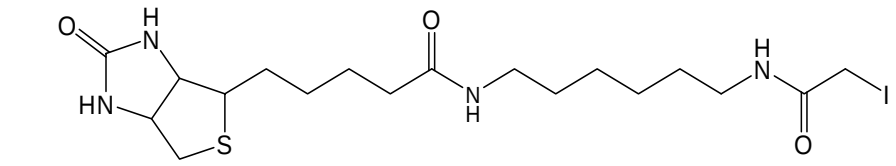
In-gel digestion of purified Keap1 separated through SDS-PAGE - FLAG-Keap1 proteins were captured from cytoplasmic fractions as described above and then separated by electrophoresis on 4-20% SDS-polyacrylamide gel (Bio-Rad ready gel) and stained with colloidal coomassie blue (Invitrogen). Gel bands then were minced into 1mm cubes, which were washed with 0.1M ammonium bicarbonate. Another 50 µL of 0.1 M ammonium bicarbonate was added to cover the cut gel cubes, then 12 µL of 40 mM *tris*(carboxyethyl)phosphine was added and the samples were incubated at 50 ^{o}C for 15 min. Finally, 20 µL 0.2 M iodoacetamide was added to the sample for 15 min in the dark. The liquid was then removed and replaced with 30 µL of 50% acetonitrile, 50 mM ammonium bicarbonate for 15 min. This step was repeated twice to further remove residual stain and reagents. The liquid was removed and the gel cubes were desiccated in a vacuum centrifuge. The dehydrated gel cubes were then reswelled in 20 µL of 0.01 µg/µL trypsin in 25 mM ammonium bicarbonate and incubated at 37 ^{o}C overnight. The digested peptides were then extracted twice from the gel with 40 µL of 60% acetonitrile/0.1% formic acid. The extract then was evaporated by vacuum centrifugation and the peptides were dissolved in 20 µL of 1 % formic acid for liquid chromatography-tandem mass spectrometry (LC-MS-MS) analyses.

LC-MS-MS analyses - Peptide digests were analyzed on a Themo LTQ linear ion trap instrument equipped with Thermo Surveyor LC system and microelectrospray scource (Themo Electron). LC-MS-MS analyses were done by reverse phase chromatography on

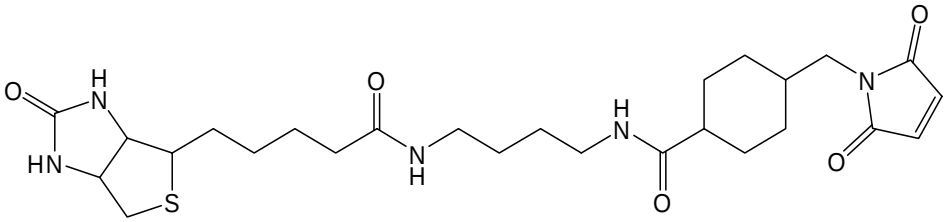
an 11 cm fused silica capillary column (100 µm ID) packed with Monitor C-18 (5 µm) (Column Engineering) and eluted first with water/acetonitrile/formic acid (95:2:0.1, v/v/v) for 5 min. A linear gradient then increased acetonitrile to 30% by 28 min, to 80% by 37 min, and decreased acetonitrile to 7% by 40 min, and then to 1% by 45 min. MS-MS spectra were acquired in data-dependent scanning mode with one full scan followed by one MS-MS scan on the most intense precursor with dynamic exclusion of previously selected precursors for a period of 3 min. MS-MS spectra were analyzed with TurboSequest (Thermo) with specification of carboxamidomethyl- (+57 Da) or IAB adducts (+382.5 Da) as variable modifications.

Results

Identification of Keap1 cysteine residues modified by electrophiles in vitro - In ongoing studies of the cellular effects of reactive electrophiles, we have employed biotin-tagged electrophiles (Fig. 2-1). These compounds display chemistries typical of a number of electrophilic metabolites of drugs and chemical and of endogenous electrophiles. To identify targets for electrophile adduction in Keap1, we used LC-MS-MS to analyze adducts formed with Keap1 *in vitro*. Treatment of His_6-Keap1 with IAB at a molar ratio of 5 to 1 for 2 h at 37°C yielded a total of 6 IAB-modified cysteines, each of which had a mass increase of 382.5, corresponding to IAB adducts (Table 2-1, Appendix 1). All of these cysteines (Cys196, Cys226, Cys241, Cys257, Cys288 and Cys319) are in the central linker domain, which is required for cytoplasmic sequestration of Nrf2 (*74*). For LC-MS-MS analyses, Keap1 protein was first reduced with tris(carboxyethylphosphine) and reduced cysteine thiols were alkylated with iodoacetamide to produce S-carboxamidomethyl derivatives. LC-MS-MS analyses routinely generated MS-MS spectra corresponding to approximately 80% of the protein sequence. All of the cysteine-containing tryptic peptides were detected, except for Cys151, Cys395 and Cys406. In analyses of electrophile-treated Keap1, cysteine-containing peptides were detected as S-carboxamidomethylated derivatives (corresponding to unadducted cysteines) or as electrophile-adducted cysteines (Table 2-1). All adducts were characterized by mass shifts to b- and/or y-ions that confirmed sequence location of the adducts. MS-MS spectra of all adducts are presented in appendix 1.



IAB



BMCC

Figure 2-1. Structures of IAB and BMCC.

Table 2-1. Sites of modification of Keap1 by electrophiles *in vitro*[1].

Residue	Domain[2]	IAB	BMCC
C77	BTB		1,2,3[3]
C196	CL	1,2,3	1,2,3
C226	CL	1	
C241	CL	1,2,3	
C249	CL		3
C257	CL	1,3	
C288	CL	1,2,3	
C319	CL	2	
C368	Kelch		1,2,3
C489	Kelch		1,2,3
C513	Kelch		
C518	Kelch		
C583	Kelch		
C622	CT		
C624	CT		

[1]Human His_6-keap1 (1 nmol) was incubated with 100 μM electrophile (5-fold molar excess) for 2 hrs in 1 M ammonium bicarbonate at 37°C and then reduced, alkylated, digested and the peptides were analyzed by LC-MS-MS as described under Experimental Procedures. Empty cells indicate that the corresponding peptide was detected as the S-carboxamidomethyl derivative following reduction and alkylation with iodoacetamide.

[2]Domains: BTB, BTB domain; CL, central linker domain; Kelch, Kelch domain; CT, C-terminal domain.

[3]Indicates experiment number in which adducted peptide was detected

A two hour treatment of Keap1 with 100 μM IAB, at a 5 to 1 molar ratio should be adequate for the critical reactive cysteines to be modified, since in our *in vivo* studies, Nrf2 was activated and its downstream biological effects were observed within 2 h of electrophile treatment (see below). Longer exposure to IAB (4 h and 6 h) resulted in alkylation of cysteines outside the central linker domain (Table 2-2). These include cysteines in the BTB domain (Cys77), in the kelch repeat domain (Cys368, Cys489, Cys518 and Cys583) and in the C-terminal domain (Cys622).

The biotin-tagged N-alkylmaleimide BMCC displayed a very different pattern of alkylation than IAB. At 37°C and at a molar ratio of 5 to 1, BMCC modified 5 cysteines, including Cys196 and Cys249 in the central linker domain Cys77 in BTB domain and Cys368 and Cys489 in the C-terminal Kelch repeat domain (Table 2-1). The only residue attacked by both BMCC and IAB is Cys196. Thus, the results shown in Table 1 indicate that Keap1 cysteines exhibit different chemical reactivity towards the two electrophiles.

IAB and BMCC produce differential Nrf2 activation in vivo - IAB was modestly toxic to HEK293 cells as measured by the LDH leakage assay and 100 μM IAB induced < 5 % LDH leakage at 4 h. Treatment of ARE/TK-GFP transfected HepG2 cells with 100 μM IAB for 2 hrs at 37°C increased ARE-directed GFP expression 2.78 fold (Fig. 2-2A), which was similar to the induction by the known ARE activator tBHQ (3.72 fold) and which demonstrated that IAB can induce expression of ARE-driven genes. Furthermore, a 2 hr treatment of HEK293 cells with 100 μM IAB induced the accumulation of nuclear Nrf2 and increased the expression of HO-1 protein (Fig 2-2). Purity of cytosolic

Table 2-2. Modification of Keap1 by IAB as a function of time *in vitro*[1]

Residue	Domain[2]	IAB treatment, hr		
		2	4	6
C77	**BTB**		**1,3**[3]	**1,2,3**
C196	**CL**	**1,2,3**	**1,2,3**	**1,2,3**
C226	**CL**	**1**	**1,2,3**	**1,2,3**
C241	**CL**	**1,2,3**	**1,2,3**	**1,2,3**
C249	**CL**			**1**
C257	**CL**	**1,3**	**1,2,3**	**1,2,3**
C288	**CL**	**1,2,3**	**1,2,3**	**1,2,3**
C319	**CL**	**2**	**1,2,3**	**1,2,3**
C368	**Kelch**		**1,2,3**	**1,2,3**
C489	**Kelch**		**1,2,3**	**1,2,3**
C513	**Kelch**			**3**
C518	**Kelch**			**1,2,3**
C583	**Kelch**		**1,2,3**	**1,2,3**
C622	**CT**		**1,2,3**	**1,2,3**
C624	**CT**		**1,2,3**	

[1]Human His_6-keap1 (1 nmol) was incubated with 100 μM of IAB (5-fold molar excess) for 2 hrs in 1 M ammonium bicarbonate at 37°C and then reduced, alkylated, digested and the peptides were analyzed by LC-MS-MS as described under Experimental Procedures. Empty cells indicate that the corresponding peptide was detected as the S-carboxamidomethyl derivative following reduction and alkylation with iodoacetamide.

[2]Domains: BTB, BTB domain; CL, central linker domain; Kelch, Kelch domain; CT, C-terminal domain.

[3]Indicates experiment number in which adducted peptide was detected

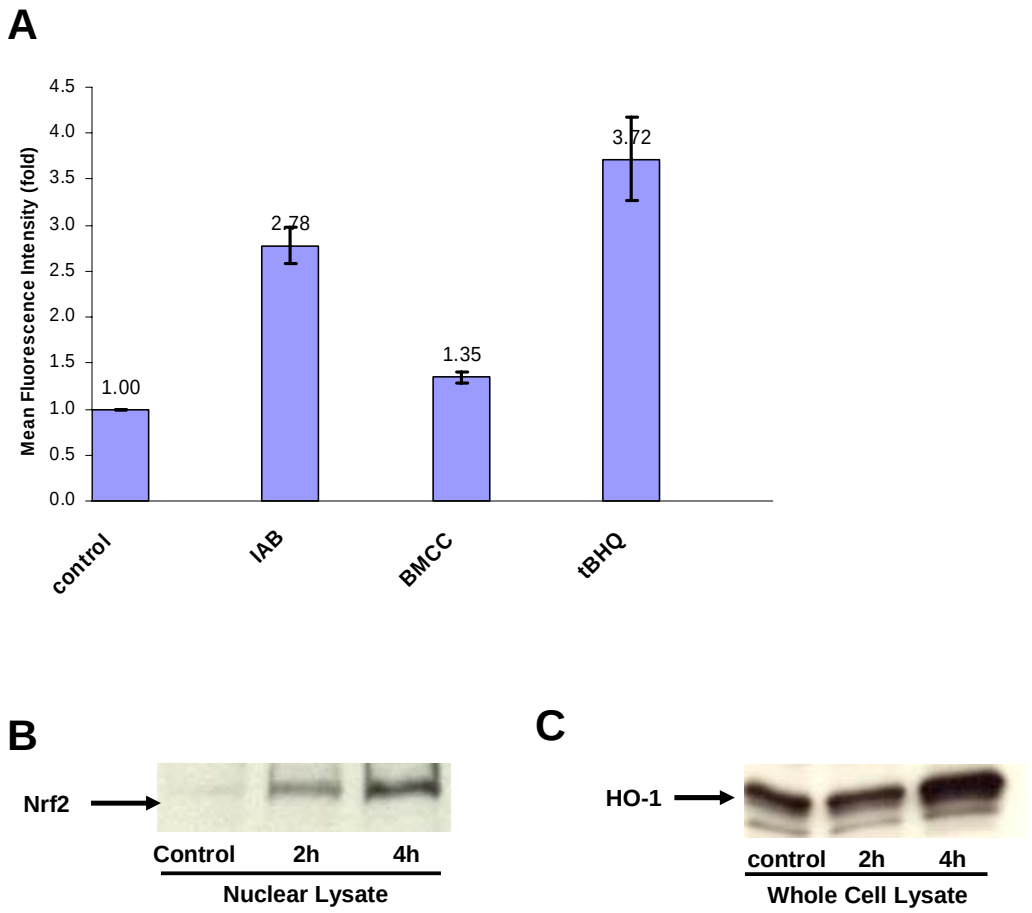


Figure 2-2. Activation of Nrf2 and ARE gene expression induced by IAB in HEK293 cells. (A) ARE-GFP transfected HepG2 cells were treated with 100 µM IAB, BMCC or tBHQ for 24 h. Cells were harvested and GFP expression level was determined using flow cytometer. The mean fluorescence intensity fold was calculated using a value of 1 for the control cells. Data are expressed as the means ± SD of three experiments. (B) Nuclear proteins were obtained from cells treated with 100 µM IAB for 0 h, 2h, 4h and 20 µg from each sample were analyzed by SDS-PAGE and Nrf2 protein was detected by western blot. (C) Cell lysates were analyzed by SDS-PAGE following treatment for 0 h, 2 h, and 4 h with100 µM IAB and HO-1 protein was detected by western blot.

and nuclear extracts was established by western blotting for GSTP1-1 and PARP-1, respectively (data not shown)

To further examine the interplay of electrophiles with the Keap1-Nrf2 system, we generated HEK293 cells stably expressing FLAG-Keap1 (Fig. 2-3). Treatment of FLAG-Keap1 transfected cells with 100 μM IAB for 2 hrs at 37°C resulted in a robust induction of nuclear Nrf2, as well as increased expression of HO-1 (Fig. 2-4). tBHQ also induced nuclear Nrf2 accumulation and induction of HO-1 protein. In contrast, neither nuclear Nrf2 accumulation nor expression of HO-1 protein were affected by a 2 hr BMCC treatment. This result suggests that IAB, but not BMCC can induce Keap1-Nrf2 dissociation, nuclear translocation of Nrf2, and Nrf2-mediated expression of HO-1.

Differential effects of IAB and BMCC on formation of high molecular weight Keap1 forms in vivo - FLAG-Keap1 transfected cells were treated with 100 μM IAB for 2 hrs at 37°C, FLAG-Keap1 proteins were captured with anti-FLAG antibodies, and analyzed on reducing SDS-PAGE gels. The majority of the Keap1 protein from IAB-treated cells migrated in a series of high molecular weight (HMW) bands with a molecular mass of greater than 150 kDa, whereas the Keap1 protein from untreated cells migrated with an observed molecular weight of 70 kDa, which corresponds to the Keap1 monomer (Fig. 2-5A, B). Immunoblotting with anti-ubiquitin indicated intense ubiquitin immunoreactivity co-migrating with the HMW Keap1 protein bands from IAB-treated cells, but not from controls (Fig 2-5C). HMW Keap1 forms were not observed in cells

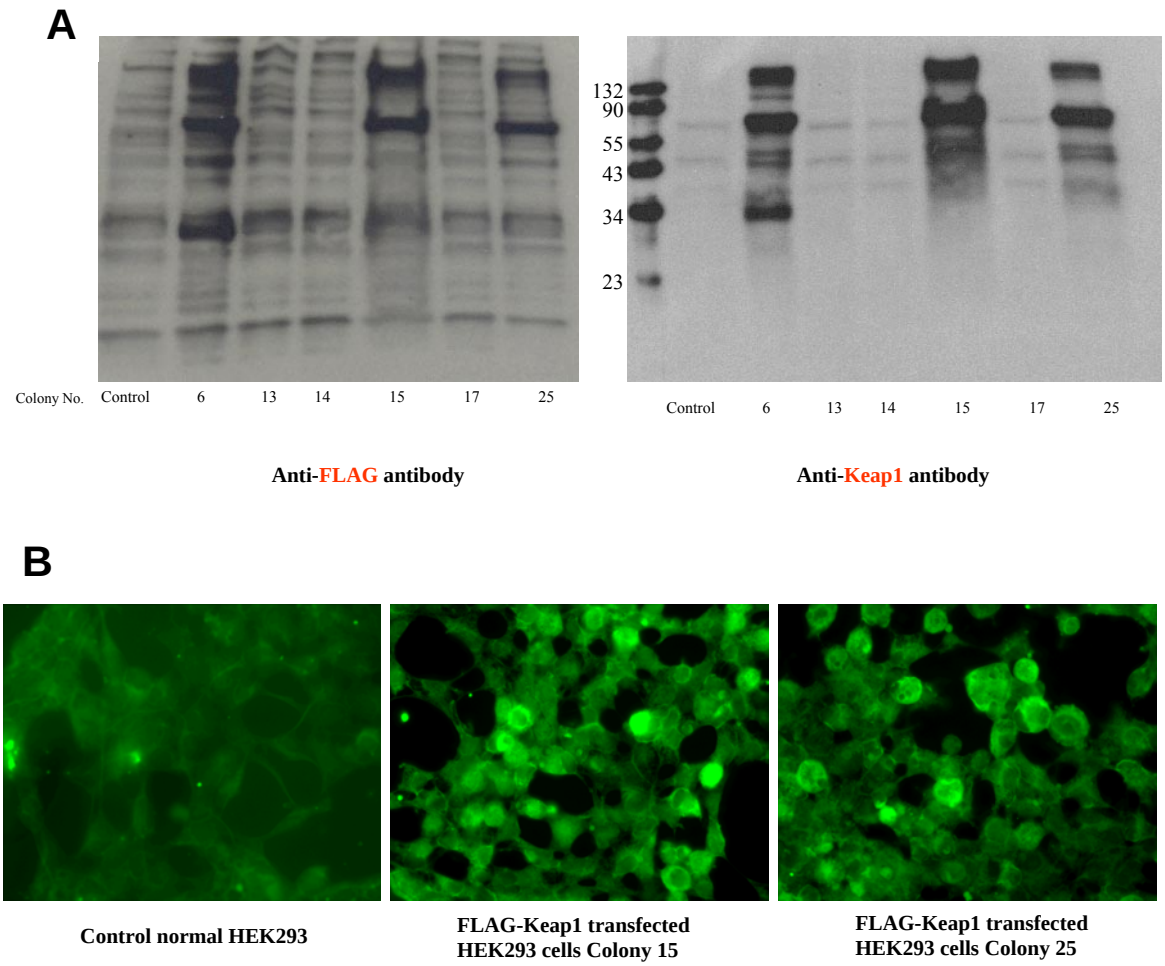


Figure 2-3. Transfected HEK 293 cells stably overexpress Flag-Keap1. (A) HEK293 cells were transfected with plasmid containing FLAG-keap1 sequence. Colonies selected by geneticin were picked for western blotting analyses. Overexpressed FLAG-Keap1 were detected in three colonies with anti-FLAG and anti-keap1 antibodies. (B) Confocal microscopic analyses of transfected cells probed with mouse anti-FLAG antibodies and stained with secondary antibody FITC-mouse anti-IgG confirmed overexpression of FLAG-keap1 in colony 15 and 25.

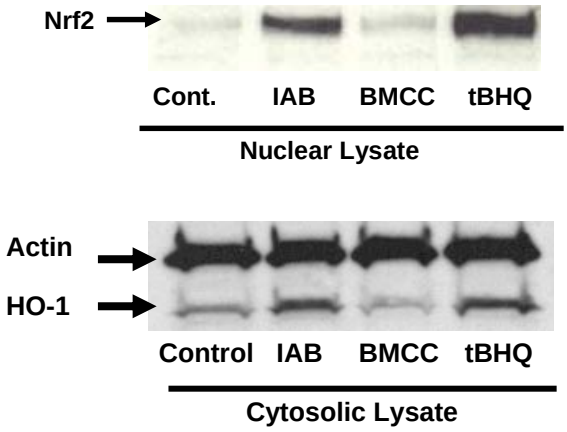


Figure 2-4. Effect of IAB, BMCC and tBHQ on nuclear Nrf2 and cytosolic HO-1 protein levels. Cells stably expressing FLAG-Keap1 were treated with 100 μM IAB, BMCC or tBHQ, and nuclear and cytosolic protein fractions were prepared. Nuclear proteins from each sample (20 μg) were analyzed by SDS-PAGE and immunoblotting with anti-Nrf2 antibody. Cytoplasmic proteins (20 μg) were analyzed by SDS-PAGE and immunoblotting with anti-HO1 antibody.

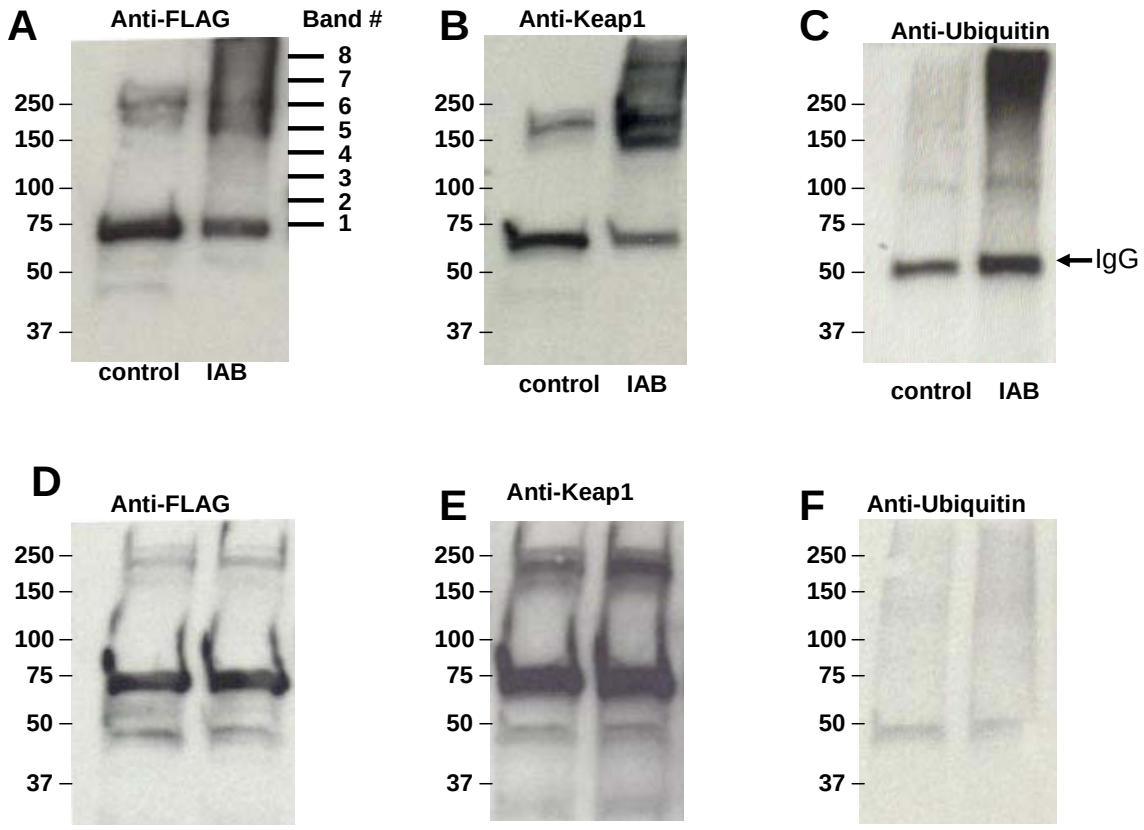


Figure 2-5. Formation of HMW keap1 protein forms in HEK293 cells stably expressing FLAG-Keap1 treated with electrophiles. FLAG-Keap1 proteins isolated from untreated controls and either IAB treated (A-C) or BMCC-treated (D-F) FLAG-Keap1 transfected HEK293 cells were aliquoted to be run on three SDS-PAGE gels and analyzed by western blotting. Keap1 proteins were detected with anti-Keap1 (A,D), anti-FLAG (B,E) and anti-ubiquitin (C,F).

treated with 100 μM BMCC (Fig. 2-5D-F), nor was anti-ubiquitin immunoreactivity detected. The HMW Keap1 protein forms were detected under reducing conditions on SDS-PAGE (15 mM β-mercaptoethanol in the loading buffer). Pretreatment of the samples with 8M urea, reduction with TCEP and alkylation of the reduced protein with iodoacetamide prior to SDS-PAGE failed to alter the migration of the HMW Keap1 products (data not shown). However, these denaturation conditions did result in detection of Cys151 as the S-carboxamidomethyl derivative (data not shown) and is consistent with the work of Wakabayashi et al. (*70*), which indicated that C151 does not undergo adduction. Treatment of His_6-Keap1 with IAB *in vitro* did not generate HMW Keap1 products detectable by immunoblotting (Fig. 2-6).

The formation of HMW Keap1 upon treatment with the prototypical ARE inducer tBHQ was reported previously (*74*). In FLAG-Keap1-expressing 293 cells, both IAB and tBHQ induced a concentration-dependent formation of HMW Keap1 forms (Fig 2-7).

To better understand the nature of the HMW Keap1 forms, we analyzed tryptic digests of the corresponding gel bands by LC-MS-MS. Eight gel sections corresponding to a molecular weight range from 70 kDa and above (indicated in Fig. 2-5A) were analyzed. The majority of peptides detected in these analyses mapped to Keap1 and ubiquitin, both of which were represented by detection of multiple peptides in different bands. Keap1 proteins from both control and IAB-treated cells were found in multiple bands of 70 kDa molecular weight and higher (Fig. 2-8). However, the distribution of HMW Keap1 and ubiquitin are different between these samples. In control cells, Keap1 protein with a molecular weight of 70 kDa is the dominant species (Fig. 2-8A), whereas

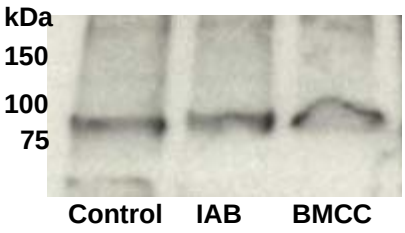


Figure 2-6. IAB itself could not trigger the formation of HMW keap1 *in vitro*. Purified His-keap1 proteins were treated with IAB and BMCC for 2 h, and then were analyzed by SDS-PAGE. Keap1 proteins were detected with anti-Keap1 antibodies by Western blot.

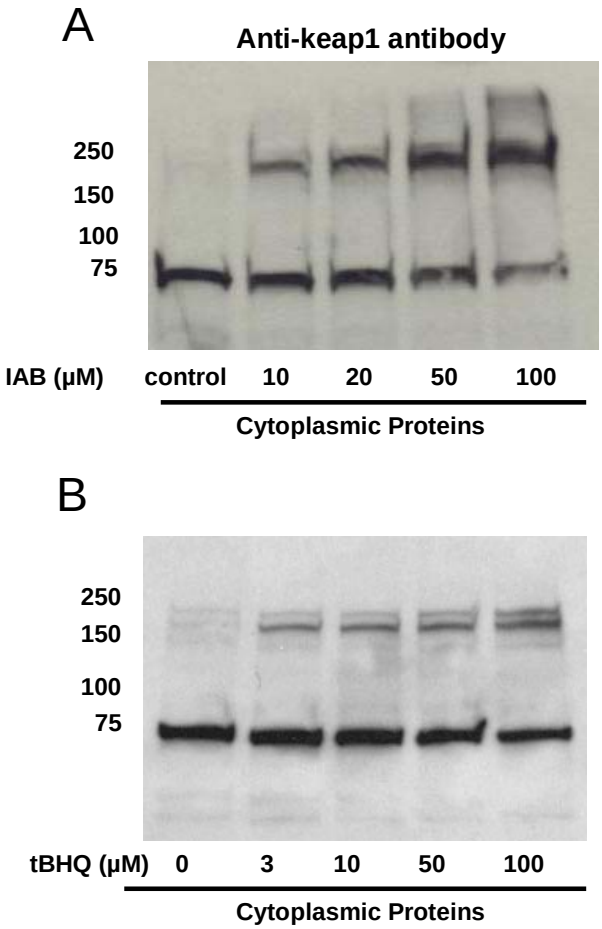


Figure 2-7. Concentration-dependent formation of HMW Keap1 by electrophiles. HEK293 cells stably expressing FLAG-Keap1 were treated 2 hr with (A) IAB or (B) tBHQ and cytosolic proteins were obtained by cellular fractionation. A 20 μg aliquot of cytosolic proteins from each sample was analyzed by SDS-PAGE and immunoblotting with anti-Keap1 antibody.

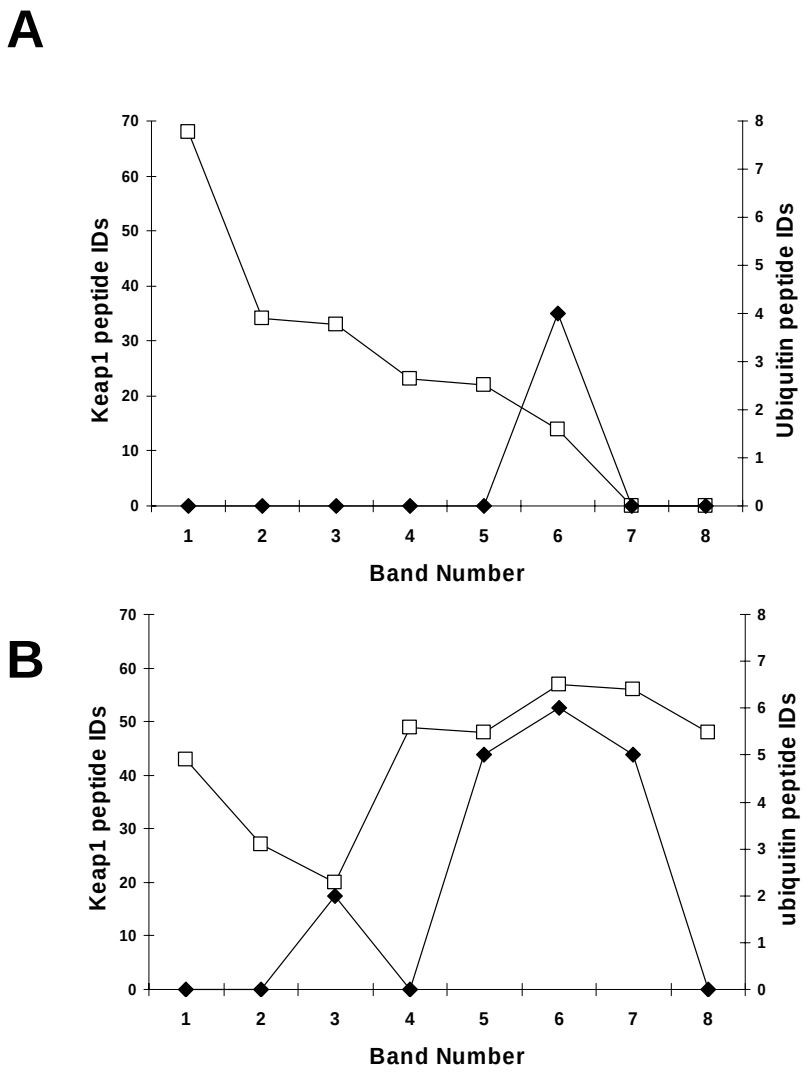


Figure. 2-8. Distribution of Keap1 (□) and ubiquitin (♦) peptide identifications in SDS-PAGE gel fractions from (A) control and (B) IAB-treated transfected cells. FLAG-Keap1 transfected cells were treated with 100 μM IAB for 2 h. FLAG-Keap1 protein then was purified from treated cells and from untreated controls with anti-FLAG antibodies and separated by SDS-PAGE. Eight bands containing proteins with molecular weights higher than 69 kDa were cut and the proteins in these bands were digested and analyzed by LC-MS-MS. Numbers of identifications refer to MS-MS spectra matched to Keap1 or ubiquitin peptides according to criteria described under "Experimental Procedures".

Keap1 protein was detected primarily in bands corresponding to a molecular weight greater than 150 kDa in the IAB-treated cells (Fig. 2-8B). Ubiquitin peptides were only detected in band 6 in control samples, which is consistent with the immunoblot result (Fig. 2-8A). Ubiquitin peptides were found in several bands corresponding to ubiquitin-immunoreactive HMW Keap1 forms from IAB treated cells (Fig. 2-8B), especially in band 6, where six ubiquitin peptides were detected, corresponding to 68.4% of the ubiquitin sequence. The numbers of Keap1 and ubiquitin peptides detected by LC-MS-MS depends in part on levels of background contaminant proteins in each gel band. Thus, numbers of peptide identifications for Keap1 and ubiquitin are at best a semiquantitative measure of protein concentration. However, numbers of detected peptides generally coincide with protein levels detected by immunostaining. Lack of detection of ubiquitin peptides in band 8 (Fig. 2-8B) reflects the presence of contaminating HMW proteins that did not enter the gel.

Trypsin digestion of ubiquitinated proteins leaves a Gly-Gly tag attached to the ε-amino group of the ubiquitin-modified lysine on the target protein, leading to the increased mass shift of 114 on this lysine. One of the ubiquitin peptides was found to be Gly-Gly modified at the site of Lys48 (Fig. 2-9), which indicates that the HMW Keap1 proteins contained Lys48-involved polyubiquitin chains (*145*).

Characterization of ubiquitination and IAB adduction sites on Keap1 in vivo - To map the sites of IAB-induced modifications on Keap1 *in vivo*, FLAG-Keap1 transfected cells were treated with IAB and FLAG-Keap1 protein was captured with an anti-FLAG

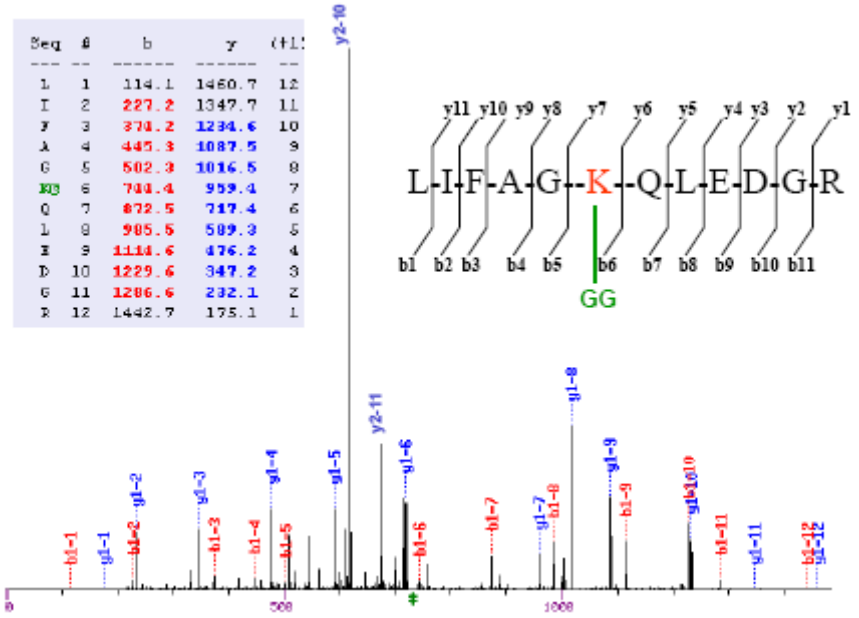


Figure 2-9. MS-MS spectrum of Gly-Gly-modified ubiquitin peptide containing the modification site on Lys48. Ubiquitinated proteins were digested with trypsin, followed by LC-MS-MS analyses. A Gly-Gly tag with a mass of 114 indicates the attachment site of the ubiquitin protein.

affinity column. LC-MS-MS analysis of a tryptic digest of FLAG-Keap1 proteins identified a peptide containing a Gly-Gly tag at Lys298, which is in the central linker domain (Fig. 2-10).

Further analysis of the LC-MS-MS data also revealed three IAB adduct sites, which were identified as Cys241 (detected in 3/3 analyses), Cys257 (2/3) and Cys288 (3/3). These cysteines are also Keap1 adduction sites following IAB treatment *in vitro* (see above). Cys226 was not found to be IAB modified, which is consistent with our *in vitro* studies indicating that this residue was found modified in only one out of three samples when Keap1 protein was incubated with IAB for 2 h (see above).

Nrf2 stabilization coincides with electrophile-specific adduction and ubiquitination of Keap1 in vivo - FLAG-Keap1 transfected cells were exposed to 100 μM IAB for 2hrs. FLAG-Keap1 proteins then were captured with anti-FLAG antibodies, and associated proteins were analyzed by immunoblotting (Fig. 2-11A). In untreated cells, very little Nrf2 was found associated with Keap1, as detected by immunoblotting. These results are consistent with those reported previously (*75*) and were interpreted to indicate rapid destabilization of Nrf2 when it is associated with non-adducted Keap1. Treatment with IAB, which resulted in the adduction and ubiquitination of Keap1 also resulted in Nrf2 stabilization. IAB treatment-induced Nrf2 stabilization was indicated by elevated cytosolic and nuclear Nrf2 (Figure 2-11B).

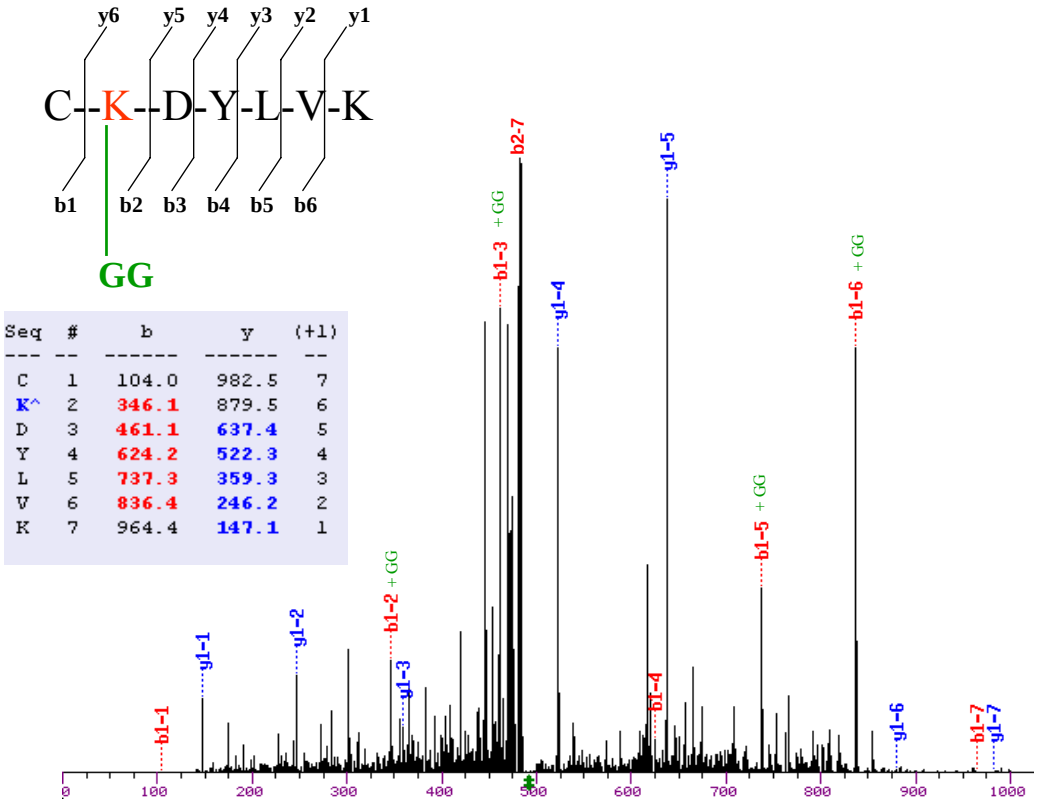


Figure 2-10. MS-MS spectrum of Gly-Gly-modified Keap1 peptide containing the modification site on Lys298. Ubiquitinated proteins were digested with trypsin, followed by LC-MS-MS analyses. A Gly-Gly tag with a mass of 114 indicates the attachment site of the ubiquitin protein.

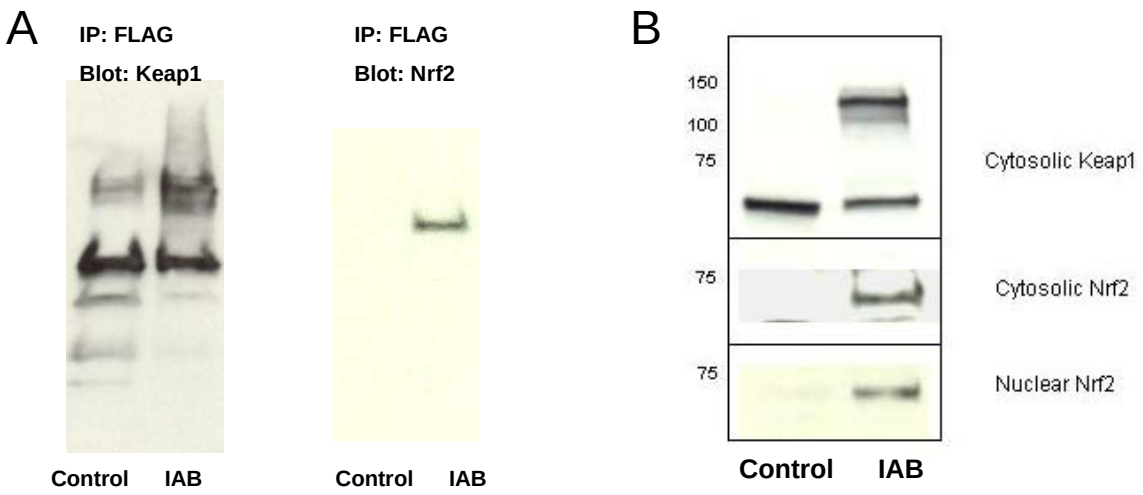


Figure 2-11. Nrf2 stabilization coincides with electrophile-specific adduction and ubiquitination of Keap1 *in vivo*. (A) HEK293 cells stably expressing FLAG-Keap1 were treated with 100 μM IAB for 2 hrs at 37°C. FLAG-Keap1 proteins were immunoprecipitated from whole cell lysates from untreated control and IAB treated cells with anti-FLAG antibodies. The captured proteins then were analyzed by SDS-PAGE and subjected to immunoblot analysis with anti-Keap1 (right) or anti-Nrf2 (left) antibodies. (B) HEK293 cells stably expressing FLAG-Keap1 were treated with 100 μM IAB for 2 hrs at 37°C. Nuclear and cytosolic fractions were isolated and analyzed by immunoblotting with anti-Keap1 and anti-Nrf2 antibodies.

Discussion

Keap1 serves as a sensor-trigger for the activation of Nrf2-regulated genes by electrophiles. Two models have been proposed to account for the role of Keap1. In the first, Keap1 sequesters Nrf2 in association with cytoskeletal actin filaments (*144*). Electrophile modification of Keap1 thiols is proposed to dissociate the Keap1:Nrf2 complex, thus enabling Nrf2 nuclear translocation and activation of Nrf2-regulated genes (*69*). However, recent work indicates that Keap1 serves as an adaptor for Cul3-dependent Nrf2 ubiquitination (*75;78*). This observation suggests that Keap1 does not passively sequester Nrf2, but instead actively directs Nrf2 degradation by facilitating Cul3-dependent Nrf2 ubiquitination. In this model, electrophiles have been shown to block Nrf2 ubiquitination, resulting in Nrf2 stabilization and nuclear translocation (*75*). However, a major unresolved issue is the mechanism by which electrophiles trigger these events. .

Here we have mapped Keap1 modifications by different thiol-reactive electrophiles *in vitro* and *in vivo* and we have demonstrated that site-specific adduction leads to Keap1 ubiquitination and Nrf2 activation. We used two biotin-tagged electrophiles with different thiol-reactive chemistries to probe the relationship between Keap1 modification specificity and Nrf2 activation. The iodoacetamide-containing probe IAB and the N-alkylmaleimide probe BMCC displayed strikingly different specificities for Keap1 alkylation *in vitro* and only IAB activated Nrf2 *in vivo*. Whereas IAB alkylated human Keap1 primarily in the central linker domain, BMCC modified Keap1 primarily in other domains. In our experiments, different electrophiles yielded different

adduct maps. IAB alkylated Keap1 preferentially in the central linker domain. Neither IAB nor BMCC modified Cys273. Our failure to detect adduction at C273 was not due to poor sequence coverage in LC-MS-MS analyses. Indeed, we routinely detected all of the cysteines in question either as adducts or as carboxamidomethylated derivatives following reduction and alkylation of cysteines with iodoacetamide during sample workup. All identifications were based on MS-MS spectra, rather than on mass measurements of intact peptides or peptide adducts. The only cysteines not routinely detected in our analyses were Cys151, Cys395 and Cys406. Of these, Cys151 is of potential interest in Keap1 function (*146*). In the case of IAB, selective targeting of several central linker domain cysteines correlated with Nrf2 stabilization, nuclear translocation and Nrf2-directed gene activation.

Previous MS analyses characterized Cys 273 and Cys288 of murine Keap1 as targets of the electrophile dexamethasone mesylate (*69*) and led the authors to denote these as the "most reactive residues of Keap1". In our experiments, human Keap1 is modified at Cys288, but not Cys273. Moreover, we found that dexamethasone mesylate modified human Keap1 at several central linker domain Cys residues, but not at Cys273 or Cys288 (data not shown). Differences between adduction patterns on murine *versus* human Keap1 may be due in part to sequence differences in the central linker domains, which differ by 12 of 153 residues. Our results suggest that site selectivity is a property not just of the amino acid target, but of the electrophile structure and reactivity. Whether alkylation at specific residues in the central linker domain of Keap1 or elsewhere is a

general trigger for Nrf2 activation cannot be satisfactorily resolved without further studies with other electrophiles.

Wakabayashi et al. proposed that electrophiles induced dimerization of Keap1 monomers via Cys273-Cys288 disulfide linkages of the monomers (*70*). This "physical release" mechanism cannot be an obligatory means of Nrf2 activation, as this linkage is not possible in IAB adduction of Keap1, which nevertheless results in Nrf2 activation. Our data and those of Zhang et al. suggest that activation of Nrf2 reflects not merely changes in Keap1 structure, but also Keap1 ubiquitination (*74;75*). However, the question of whether Keap1 ubiquitination occurs before or coincident with Nrf2 release will require additional study.

The major physical manifestation of change in Keap1 upon treatment with Nrf2 activators is the formation of HMW Keap1 forms. This was first reported by Zhang et al. (*74*). In accord with the work of Zhang et al., treatment of Keap1 with either IAB or the prototypical ARE activato tBHQ formed a series of HMW Keap1 immunoreactive bands (Figure 2-5). LC-MS-MS analyses of tryptic peptides from in-gel digestion of these HMW Keap1 bands indicated the presence of both Keap1 protein and ubiquitin, including K48-(Gly-Gly)-modified ubiquitin peptides. These HMW Keap1 bands thus correspond to a Keap1 dimer and to its polyubiquitinated Keap1 forms. Although the presence of low amounts of a HMW Keap1 form also can be seen in untreated cells, most of the Keap1 migrates as the monomeric species (Fig. 2-5). Thus, the conversion of Keap1 to HMW forms appears to be triggered by specific thiol adduction. We postulate that these adducted Keap1 species are the substrates for ubiquitination.

Recent reports indicated that under basal conditions Keap1 functions as an adaptor protein for Cul3-dependent ubiquitin ligase complex Keap1-Cul3-Rbx1 to target Nrf2 for ubiquitination and thereafter proteosomal degradation (*75-77*). Inducers of Nrf2-dependent transcription inhibit Keap1-dependent ubiquitination of Nrf2 and increase its stability (*75*). Our results suggest that Nrf2-directed transcription may be regulated through site specific modification and ubiquitination of Keap1 by electrophiles. This prosposition is supported by evidence that mutations within the Keap1BTB domain increase association of Keap1 with Cul3, inhibit the ubiquitination of Nrf2 and trigger the ubiquitination of Keap1 (*75*).

Some SCF (Skp1p-cullin-F-Box protein) complexes, which function as E3-ubiquitin ligases, may disassemble once a substrate has been degraded and re-assemble to form new complexes with a different substrate (*75;147;148*). F-box proteins are substrate adaptors that provide substrate specificity and ubiquitin-dependent degradation by SCF complexes of certain G-cyclins involved in the control of cell cycle progression (*149*). Galan et al. demonstrated that F-box proteins are degraded by autoubiquitination within their own SCF complex, which suggests that SCF complex can switch rapidly between substrate (G-cyclin) and substrate adaptor (F-box protein) to balance their levels (*150;151*). Cul3-associated BTB proteins are substrate adaptors for Cullin-3 ubiquitin ligases and these proteins are themselves regulated through ubiquitination and proteasome-mediated degradation (*152*). Keap1 is a BTB protein and functions as an adaptor to target Nrf2 for ubiquitination by Cul3-based E3 ligase (*75-77*). Our results suggest that site-specific modification of Keap1 by electrophiles switches ubiquitin

targeting from Nrf2 to Keap1 and that this target switching mechanism governs Nrf2 activation by electrophiles.

Our studies provide a new perspective on the mechanisms by which electrophile modifications trigger Nrf2 activation. The differences in Keap1 adduction patterns between the electrophiles indicate that adduction site specificity is highly dependent on the chemistry of the electrophile modifier and not simply a function of thiol reactivity in the protein *per se*. Not all electrophiles that modify Keap1 are Nrf2 inducers. On the other hand, electrophiles that are Nrf2 activators may display similar patterns of Keap1 modification, perhaps in targeting the central linker domain of the protein. The identification of consensus electrophile target sites in Keap1 awaits a more complete study of adduction and Nrf2 activation by different electrophile structures.

Our results also indicate that site-specific adduction triggers Keap1 ubiquitination and further imply that adduction triggers a switching of Cul3-dependent ubiquitination from Nrf2 to Keap1, which results in Nrf2 activation (Fig. 2-12). Although this hypothesis certainly requires further testing, it is interesting to note that the ubiquitination target site on Keap1 is Lys 298, which lies adjacent to Cys residues in the central linker domain. Zhang et al. showed that substitution of serine at Cys273 and Cys288 in murine Keap1 blocked Nrf2 ubiquitination and they speculated that these cysteines may participate directly in ubiquitin transfer to Nrf2 (*74;75*). The same line of reasoning suggests the possibility that adduction of cysteines in the central linker domain blocks ubiquitin transfer to Nrf2 and causes ubiquitin deposition instead on Lys298 of the Keap1 protein.

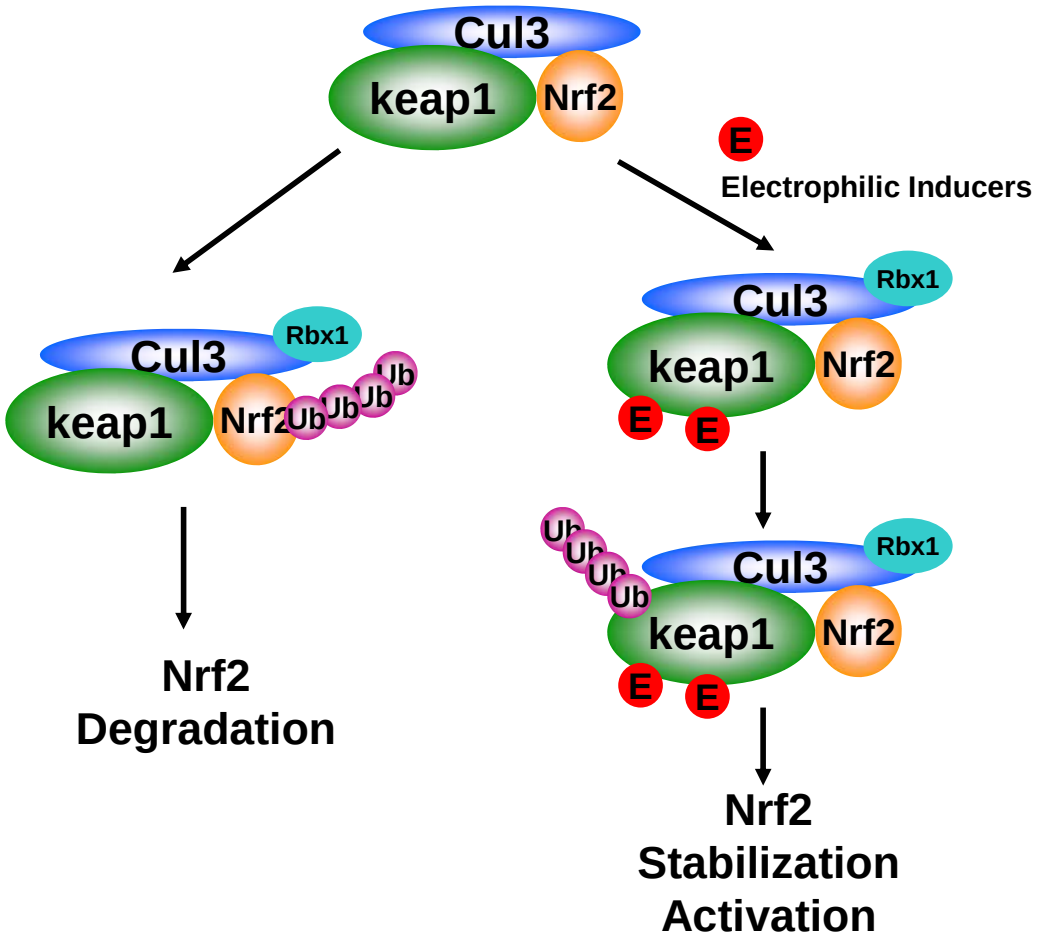


Figure 2-12. Proposed mechanism of Nrf2 activation by electrophilic inducers of Nrf2-dependent gene expression. Under basal conditions, keap1 is an adaptor for Cul3-dependent Nrf2 ubiquitination. When Keap1 is modified by electrophilic inducers, Cul3-dependent ubiquitination switches from Nrf2 to Keap1, leading to the degradation of Keap1 and to the stabilization and activation of Nrf2 protein.

CHAPTER THREE – IDENTIFICATION OF SENSOR CYSTEINES IN HUMAN KEAP1 MODIFIED BY THE CHEMOPREVENTIVE AGENT SULFORAPHANE

Introduction

Epidemiologic studies have demonstrated that individuals who consume high amounts of fruits and vegetables appear to have a low risk of cancer (*153-155*). Among anticarcinogenic vegetables, cruciferous vegetables have among the strongest chemopreventive effects (*155*) and studies have reported an inverse association between consumption of crucifiers and the incidence of a variety of cancers (*156*). Sulforaphane (R-1-isothiocyanato-4-methylsulfinylbutane), an isothiocyanate isolated from cruciferous vegetables, is among the most potent inducers of phase II enzymes (*157-159*) by activating Nrf2 (*51;140*). Sulforaphane induces GST activities and reduces the incidence of preneoplastic lesions in carcinogen-treated mouse mammary glands (*160*). Sulforaphane also effectively reduces colonic aberrant crypt foci formation in carcinogen-treated rats (*161*) and mediates a dose-dependent apoptosis and growth arrest in the prostate cancer cells (*21*).

Chemopreventive agents in cruciferous vegetables protect cells against carcinogens by inducing phase II dextoxifying and antioxidant enzymes to enhance cellular detoxication and antioxidant capacity (*23*). Induction of phase II enzymes by inducers involves the antioxidant/electrophile response element (ARE/EpRE), which is found in the 5'-flanking region of the phase II and antioxidant genes(*162;163*). Nuclear factor-E2-related factor 2 (Nrf2) was identified as the ARE-binding transcription factor

(*41*) and shown to play a critical role in the ARE-mediated gene expression. Small Maf protein was later found to form heterodimer with Nrf2 and this dimer was shown to bind the AREs of phase II genes (*47*). Overexpression of Nrf2 mediated up-regulation of ARE-regulated reporter genes, which were induced by a variety of chemopreventive agents (*38;164*). Numerous studies showed that the Nrf2 knockout mice exhibited decreased levels and induction of phase II detoxification and antioxidant enzymes and were highly sensitive to cytotoxic electrophiles compared to wild-type mice (*55-57*). Several Nrf2-dependent genes were found to be sulforaphane-inducible by comparing transcriptional profiles of small intestine of wild-type and Nrf2 knockout mice treated with sulforaphane (*54*).

Recently, a cytoskeletal actin-binding protein Keap1 (Kelch-like ECH-associated protein 1) was isolated and identified to be an inhibitory regulator of Nrf2 that binds to the Neh2 domain of Nrf2 (*44;136*). The Cys-rich Keap1 protein has five domains: an N-terminal domain, a BTB domain, a central linker domain, a Kelch repeat domain and a C-terminal domain. The Kelch repeat domain binds to Nrf2 and actin directly (*44;66*), the BTB domain is required for the dimerization of Keap1 (*67*) and the central linker domain is essential for cytoplasmic sequestration of Nrf2 (*74*). Under basal conditions, Keap1 functions as an adaptor for Cul3-dependent ubiquitination and degradation of Nrf2 (*75-77*).

A key event in the activation of Nrf2 is the reaction of electrophiles with Keap1 to form covalent adducts, which stabilizes Nrf2 and results in enhanced nuclear Nrf2 levels. The mechanism by which Keap1 modifications trigger Nrf2 activation presents an

interesting problem. Although its structure has not been determined, Keap1 has 27 Cys residues, most or all of which appear to be available to react with electrophiles. Previous studies have mapped Keap1-electrophile adducts by tandem MS analysis (*70;165*). Although studies of modification of murine Keap1 by the model alkylating agent dexamethasone mesylate suggested that central linker domain cysteines, particularly Cys273 and Cys288 were the most reactive towards electrophiles, our more recent studies of Keap1 reactions with other electrophiles indicate that targeting selectivity varies with electrophile structure (*165*). We reported that two different electrophiles IAB and BMCC both modified Keap1, but displayed strikingly different adduct maps, as analyzed by tandem mass spectrometry (MS-MS). These differences were also manifested in differences in Nrf2 activation *in vivo*. IAB activated ARE reporter genes, stabilized Nrf2 and alkylated Keap1 primarily in the central linker domain. In contrast, BMCC, which neither activated ARE reporter genes nor stabilized Nrf2, was alkylated primarily outside the central linker domain. Most interestingly, IAB and the prototypical inducer tBHQ both induced the formation of polyubiquitinated Keap1 *in vivo*. Keap1 ubiquitination was mapped to Lys298. These results suggested that alkylation of Keap1 central linker domain cysteines led to a switch in ubiquination from Nrf2 to Keap1, thus enabling Nrf2 stabilization. Moreover, the results suggested that alkylation site specificity is an important determinant of the ability of electrophiles to activate ARE/ERE-regulated genes.

Here we describe mechanisms of Nrf2 activation resulting from Keap1 modifications by sulforaphane. Sulforaphane is an isothiocyanate, which reacts with

thiols to form thionoacyl adducts. Unlike the Keap1 alkylation adducts studied previously, thionoacyl adducts are labile to hydrolysis and transacylation reactions. We have developed a LC-MS-MS method to map sulforaphane modification sites formed on Keap1 in vitro. Our studies indicate that sulforaphane displays a pattern of Keap1 modification distincly different from ARE/ERE inducers that modify Keap1 by alkylation. Moreover, the modification of Keap1 *in vivo* by sulforaphane is not accompanied by Keap1 ubiquitination. This suggests a novel mechanism for Nrf2 stabilization by sulforaphane thionoacyl adduct formation.

Experimental procedures

Cell culture, transfection and treatment

FLAG-Human Keap1 transfected human embryonic kidney 293 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 2mM glutamine, 100μg mL^{-1} penicillin,100μg/mL streptomycin and 500 μg mL^{-1} geneticin. Human hepatoma HepG2 cells containing an ARE/TK-GFP reporter (ProCertus Biopharm) were maintained in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 2mM glutamine, 100μg mL^{-1} penicillin and 100μg mL^{-1} streptomycin and containing 1 mg mL^{-1} geneticin. Confluent cells in 100 mm plates were washed with PBS and treated with sulforaphane (ICN Biomedicals Inc.) or equal volumes of vehicle (DMSO) delivered in 4 mL DMEM with 5% FBS.

Cell fractionation and immunoblot analyses

Confluent cells in 75 cm flasks were washed with cold PBS and lysed in cold nuclear lysis buffer (5 mM Tris, pH 7.4, 50 mM NaCl, 10 mM EDTA, 0.5 % NP-40) containing 10 μL mL^{-1} protease inhibitor cocktail (Sigma Catalog No. P8340). Lysate was centrifuged at 12,000 rpm for 2 min and supernatant was collected as the cytoplasmic fraction. The pellet was washed twice with nuclear lysis buffer to remove cytosolic contaminants, and nuclear pellet was lysed by treatment with all protein lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% NP-40, 10 mM EDTA, 0.1% SDS, 0.5% deoxycholic acid, pH 8.0) containing 5μL/mL protease inhibitor cocktail, followed

by sonication for 5 sec. The nuclear lysate was centrifuged at 13,000 rpm for 2 min to remove cell debris. Protein concentration was measured with BCA protein assay kit (Pierce). For electrophoresis, 40 μg cytoplasmic proteins or 20 μg nuclear proteins were diluted 1:1 (v:v) with 4 x NuPAGE LDS sample buffer respective and separated on NuPAGE 4-12% Bis-Tris Gels using MOPS SDS running buffer (Invitrogen). Resolved proteins then were transferred to PVDF membranes, which were blocked with 5% milk in TBST buffer (20 mM Tris HCl, pH 7.5, 200 mM NaCl, 0.1% Tween 20) and then probed with anti-HO-1 (proved by Dr. Chris Ferris) and anti-Nrf2 (Santa Cruz) in 5% milk in TBST buffer, respectively. After treatment with appropriate secondary antibody conjugated to horseradish peroxidase (Santa Cruz), immunostained proteins were detected by enhanced chemiluminescence with Western blotting luminol reagent (Santa Cruz).

Flow cytometry analysis of GFP expression induced by sulforaphane

ARE/TK-GFP transfected HepG2 cells were treated with 0, 2, 6, 20 or 60 μM sulforaphane for 6, 12 or 24 h, respectively. After washed with PBS, treated cells were harvested and resuspended in PBS buffer. Fluorescent cells were then analyzed with BD-LSR II flow cytometer (Becton Dickinson) with excitation at 488 nm and fluorescence emission was detected using 530/15 bandpass filter. 30,000 events were collected from each experiment. Mean fluorescence intensity was calculated with software FACS DiVa 4.1.1 using a value of 1 for the control cells. This experiment was repeated twice and data were expressed as the means ± SD of three experiments.

Modification of TpepC by sulforaphane

TpepC (AVAGCAGAR; SynPep Co) and sulforaphane, each at a concentration of 1 mM, were mixed in 100 µL of 0.1 M ammonium bicarbonate (pH 7.5) and the mixture was incubated at 37°C for 2 h or 16 h. TpepC modified by sulforaphane for 2 h at 37°C was incubated at 50°C for 15 min, followed by alkylation with 20mM iodoacetamide at room temperature for 15 min in the dark. Concentrated formic acid (2-3 µL) then was added to acidify each sample, and samples then were diluted 1:1000 with H_2O for LC-MS-MS analysis.

Covalent adduction of sulforaphane to His_6-Keap1 in vitro

His_6-Keap1 was expressed in *Escherichia coli* and purified as described in chapter 2. Ultrafree-MC low binding regenerated cellulose centrifugal spin filter devices with a 30,000 molecular weight cutoff (MWCO) were obtained from Millipore. Prior to sample addition, spin filters were sequentially rinsed with 200 µL methanol and 200 µL distilled water by centrifugation at 12,000 x g. Keap1 (60-70 µg) was loaded into the upper chamber of the spin filter and the sample then was centrifuged at 12,000 x g to remove the solution. The protein on the filter was washed with 200 µL 1 M ammonium bicarbonate followed by centrifugation. The protein then was suspended in 50 µL of 1 M ammonium bicarbonate (pH 7.5) containing 5mM TCEP (*tris*(carboxyethyl)phosphine) (Pierce) and sulforaphane at various concentrations. After incubation at 37°C for 15 min in the dark, the reaction was terminated by centrifugation to remove the buffer followed by an additional wash with 1 M NH_4HCO_3. The filtrates were discarded. The protein on

the filter was suspended in 6 M Urea in 40 µL 0.1 M ammonium bicarbonate containing 5 µL 40 mM TCEP and incubated at 37°C for 30 min. Then 5 µL of 0.2 M iodoacetamide was added to the sample for 15 min to convert free thiols to carboxamidomethyl derivatives. Modifed porcine sequencing grade trypsin (Promega) then was added in a 1:50 protein:trypsin ratio and the sample was incubated at 37°C for 2 h. Tryptic peptides were collected by centrifugation through the filter at 5,000 x *g* and the filtrate was acidified with 1 µL of concentrated formic acid for LC-MS-MS analyses.

LC-MS-MS analyses

TpepC and its modified form were analyzed on a Thermo TSQ Quantum triple qualdrople mass spectrometer equipped with Thermo Surveyor LC system and electrospray scource (Thermo Electron). A 150 x 2 mm LUNA 5µ C18 column (Phenomenex) was eluted first with water/acetonitrile/formic acid (95:5:0.1, v/v/v) for 5 min. A linear gradient then increased acetonitrile to 28% by 33 min, to 80% by 40 min, and decreased acetonitrile to 1% by 47 min, followed by a 13 equilibration period. MS-MS spectra were acquired with one full scan followed by three MS-MS scans on selected precursors at *m/z* 388.2, 476.7 and 416.7. MS-MS spectra were analyzed manually.

Keap1 peptide digests were analyzed on a Themo LTQ linear ion trap instrument equipped with Thermo Surveyor LC system and microelectrospray scource (Thermo Electron). LC-MS-MS analyses were done by reverse phase chromatography on an 11 cm fused silica capillary column (100 µm ID) packed with Monitor C-18 (5 µm) (Column Engineering) and eluted first with water/acetonitrile/formic acid (95:5:0.1,

v/v/v) for 5 min. A linear gradient then increased acetonitrile to 50% by 50 min, to 80% by 52 and to 90% by 55min, and decreased acetonitrile to 5% by 61 min, and then to 1% by 66 min for 10 min. MS-MS spectra were acquired in data-dependent scanning mode with one full scan followed by one MS-MS scan on the most intense precursor with dynamic exclusion of previously selected precursors for a period of 3 min. MS-MS spectra were analyzed with TurboSequest (Thermo) and manually with specification of carboxamidomethyl- (+57 Da) or sulforaphane adducts (+177 Da) as modifications.

Results

Induction of ARE-regulated gene expression by sulforaphane

Previous studies have shown that sulphoraphane activates a number of ARE-regulated genes (*54;166-168*). To define the concentration-response relationship for ARE gene activation in cultured cells, we utilized HepG2 cells containing ARE-regulated GFP construct. ARE-GFP transfected cells were treated with sulforaphane (2 μM, 6 μM, 20 μM, 60 μM) for 0, 8, 12, or 24 h. Cells then were harvested for flow cytometry analysis to measure fluorescence intensity of GFP induced by sulforaphane. At concentrations ranging from 2 to 20 μM, sulforaphane induced ARE-regulated GFP expression in a time-dependent manner ($p<0.001$, ANOVA) (Figure 3-1). At a concentration of 60 μM, ARE-regulated GFP expression was not significantly different from vehicle control ($p>0.05$, ANOVA). Exposure to these higher sulforaphane concentrations (>20 μM) caused significant cytotoxicity (data not shown). The relationship between ARE activation and sulforaphane concentration was analyzed. Maximum activation was observed following exposure to 2 μM ($p<0.02$, Student's test). Higher concentrations did not produce a statistically significant increase in GFP fluorescence ($p>0.05$, ANOVA).

Sulphoraphane-induced stabilization and nuclear accumulation of Nrf2

In our previous studies of Keap1 alkylation by electrophiles and Nrf2 activation, we employed HEK293 cells stably transfected with FLAG-Keap1 (*165*). Cells

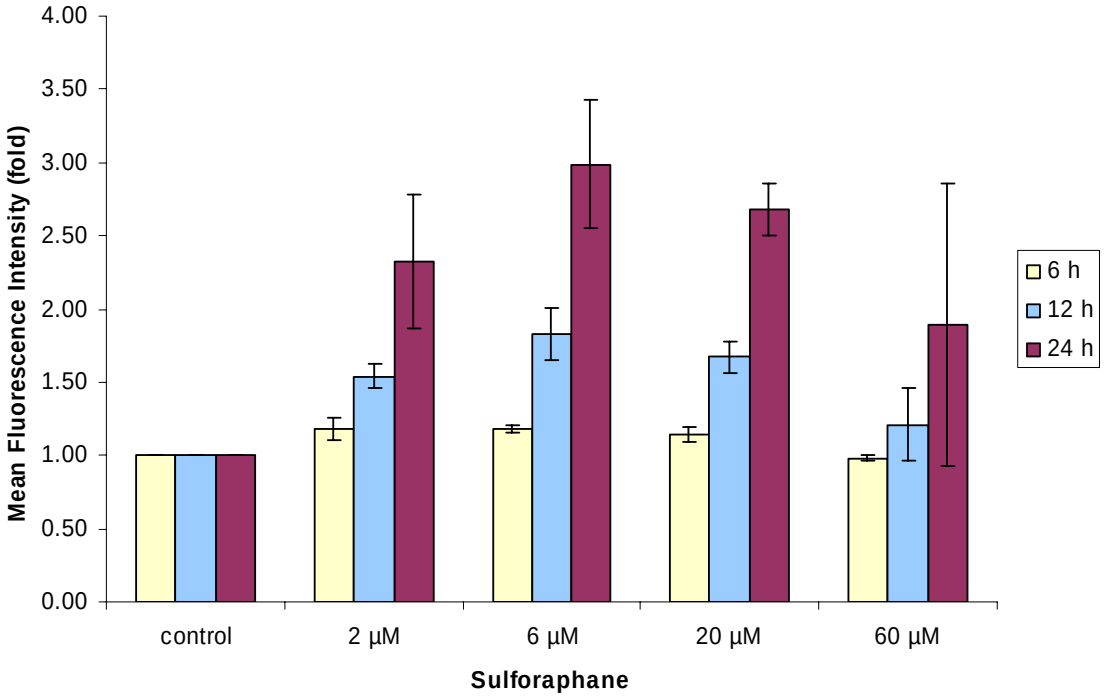


Figure 3-1. Activation of ARE gene expression induced by sulforaphane in HepG2 cells. ARE-GFP transfected HepG2 cells were treated with with sulforaphane (2 µM, 6 µM, 20 µM, 60 µM) for 0, 8, 12, or 24 h. Cells then were harvested for flow cytometry analysis. The mean fluorescence intensity fold was calculated using a value of 1 for the control cells. Data are expressed as the means ± SD of three experiments.

expressing the epitope-tagged Keap1 protein display Nrf2 activation response comparable to untransfected cells. Treatment of these cells with 6 μM sulforaphane caused a time and concentration dependent stabilization of of Nrf2 that was maximal at approximately 4 h (Figure 3-2A, B). Treated cells were fractionated and cytoplasmic and nuclear proteins were analyzed by reducing SDS-PAGE and western blotting using anti-Nrf2 as primary antibody. Treatment with 20 mM sulforaphane for 6 h produced a significantly greater elevation of Nrf2. Stabilization of Nrf2 in cytoplasm generated two Nrf2 species of slightly different MW (Figure 3-2B, arrow). However, only the higher MW species was observed in the nuclei. It is not clear whether this modified Nrf2 form results from modified by sulforaphane or from some other posttranslational modification, although phosphorylation of Nrf2 has been reported (*88;89*).

Previous work had shown that sulforaphane inhibits Keap1-dependent Nrf2 degradation by the ubiquitin-proteasome system (*74;75*). To examine this relationship in our experimental system, cells were treated with sulforaphane at both low (7 μM) and high (15 μM) concentrations for 6, 7, or 8 h. Some cells were treated for 8 h with both sulphoraphane and the proteasome inhibitor MG-132 (5 μM). At the 7 μM concentration, sulforaphane inhibited degradation of Nrf2 and this effect was enhanced by MG-132. At 15 μM, sulforaphane appeared to maximally stabilize Nrf2, as the MG-132 had no additional effect. Sulforaphane also induced expression of the ARE-regulated protein HO-1 together with cytosolic and nuclear Nrf2 stabilization (Figure 3-3). However, in cells treated with both 15 μM sulforaphane and 5 μM MG-132, cytosolic

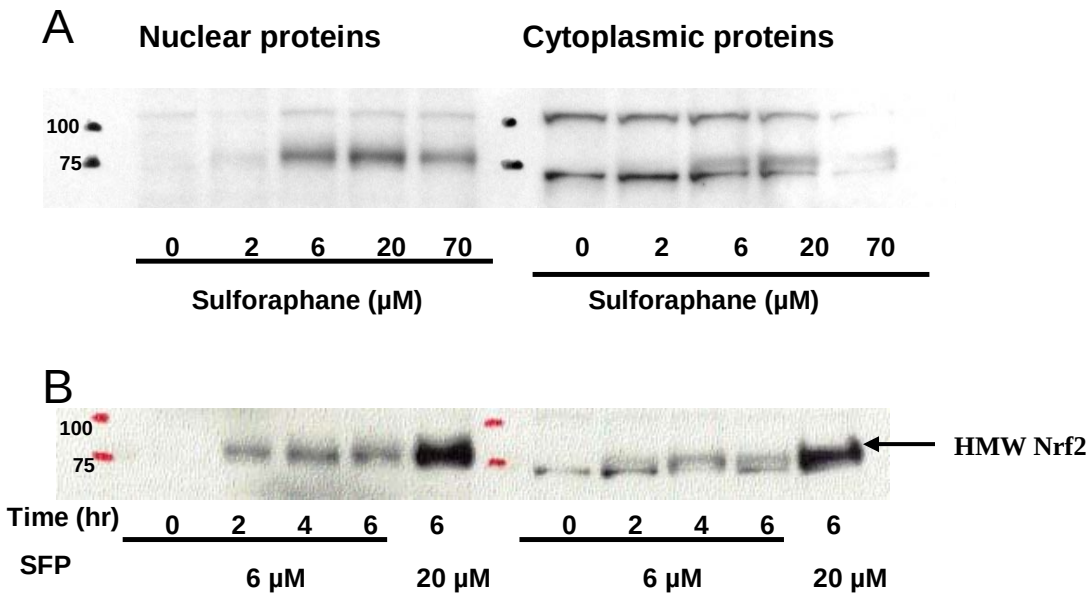


Figure 3-2. Sulforaphane caused a concentration and time dependent stabilization of Nrf2. (A) Transfected cells were treated with sulforaphane (0, 2, 6, 20, 70 μM) for 1h or (B) were treated with 6μM sulforaphane for 0, 2, 4, 6 hrs, respectively. Cell then were fractionated and cytoplasmic and nuclear proteins were analyzed by reducing SDS-PAGE and western blotting using anti-Nrf2 as primary antibody.

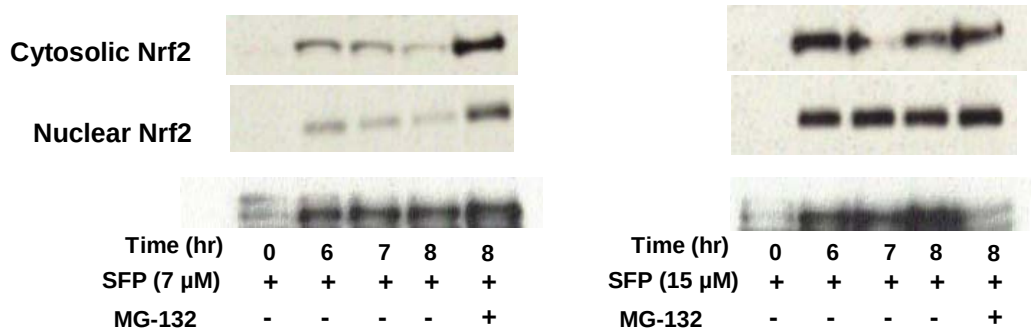


Figure 3-3. Sulforaphane induced expression of HO-1 and sulforaphane inhibited degradation of Nrf2 was enhanced by proteasome inhibitor MG-132. Transfeced HEK293 cells were treated with sulforaphane at both low (7 μM) and high (15 μM) concentrations for 6, 7, or 8 h. Some cells were treated for 8 h with both sulphoraphane and the proteasome inhibitor MG-132 (5 μM). Cell then were fractionated and cytoplasmic and nuclear proteins were analyzed by reducing SDS-PAGE and western blotting using anti-Nrf2 and anti-HO1 as primary antibodies.

HO-1 was not induced. This probably reflects diminished protein synthesis due to the combined cytotoxic effects of both sulforaphane and MG-132 at the concentrations used. These experiments suggest that significant effects of sulforaphane on ARE-regulated gene expression and Nrf2 stabilization occur at concentrations ranging from 5-20 μM. Thus, this range of sulphoraphane concentrations was used for subsequent studies of sulforaphane modification of the Keap1 protein *in vitro* (see below).

Sulphoraphane does not trigger the ubiquitination of Keap1 as tBHQ does

Phase II inducers such as IAB and tBHQ were demonstrated to stabilize and activate Nrf2 by triggering the ubiquitination of Keap1 (*165*). To examine whether sulforaphane utilizes the same mechanism to activate Keap1-Nrf2 system, FLAG-Keap1 transfected cells were treated with 100 μM tBHQ and 20 μM sulforaphane for 2 hrs. It was demonstrated that only tBHQ, but not sulforaphane, can induce the ubiquitination on Keap1 (Fig. 3-4). We note that the apparent levels of high molecular weight, ubiquitinated Keap1 form (approximately 150 KDa) in the controls appears to differ between the controls for the tBHQ treated cells and the sulforaphane-treated cells. We have observed some variability in the level of the ubiquitinated Keap1 from between experiments (*165*) and the apparent difference may also reflect slight differences in exposure times during imaging of the Western blots. However, the tBHQ treatment clearly enhances formation of the ubiquitinated Keap1, whereas the sulforaphane does not.

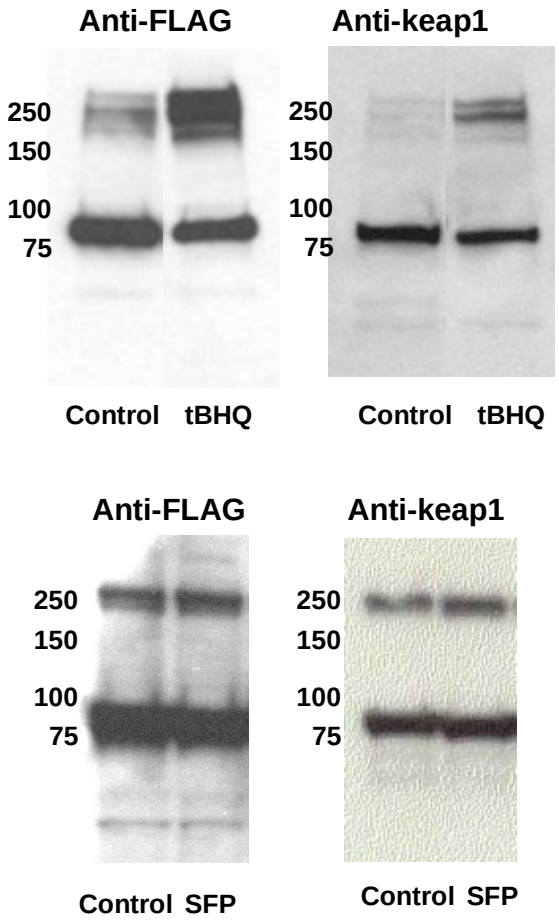


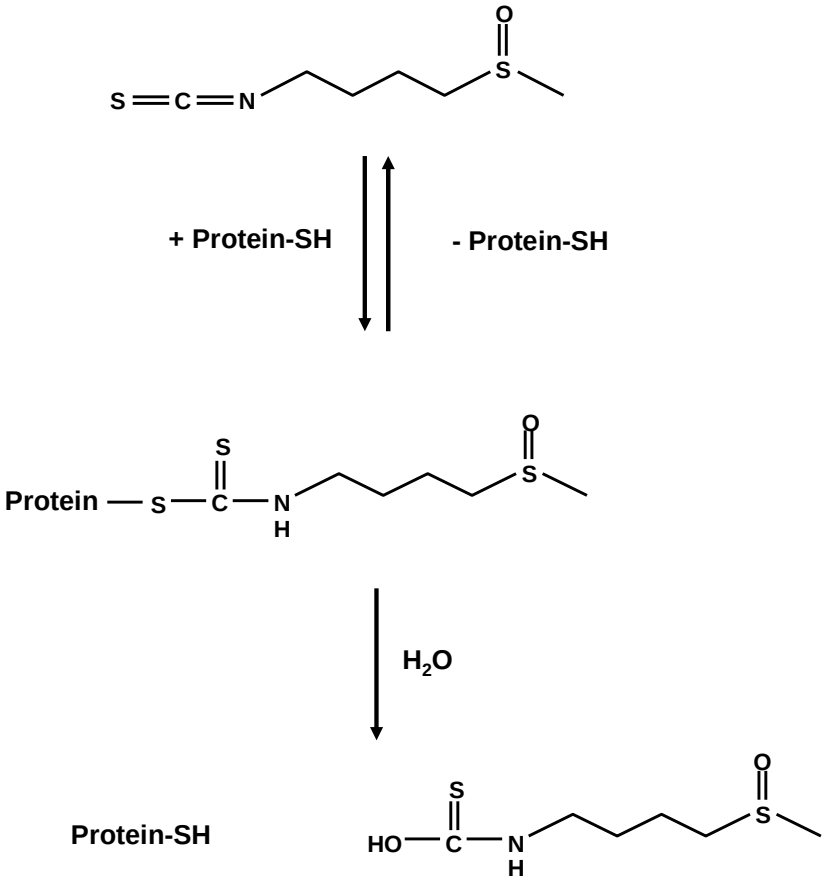
Figure 3-4. tBHQ, but not sulforaphane triggers the ubiquitination of Keap1. FLAG-Keap1 transfected HEK293 cells were treated with 100 μM tBHQ or 20 μM sulforaphane for 2 hrs respectively. Cells then were lysed and 20 μg proteins were loaded on SDS-PAGE gels and were subjected to immunoblotting analyses. Keap1 proteins were probed with anti-Keap1 and anti-FLAG antibodies.

Stability of sulforaphane cysteine adducts

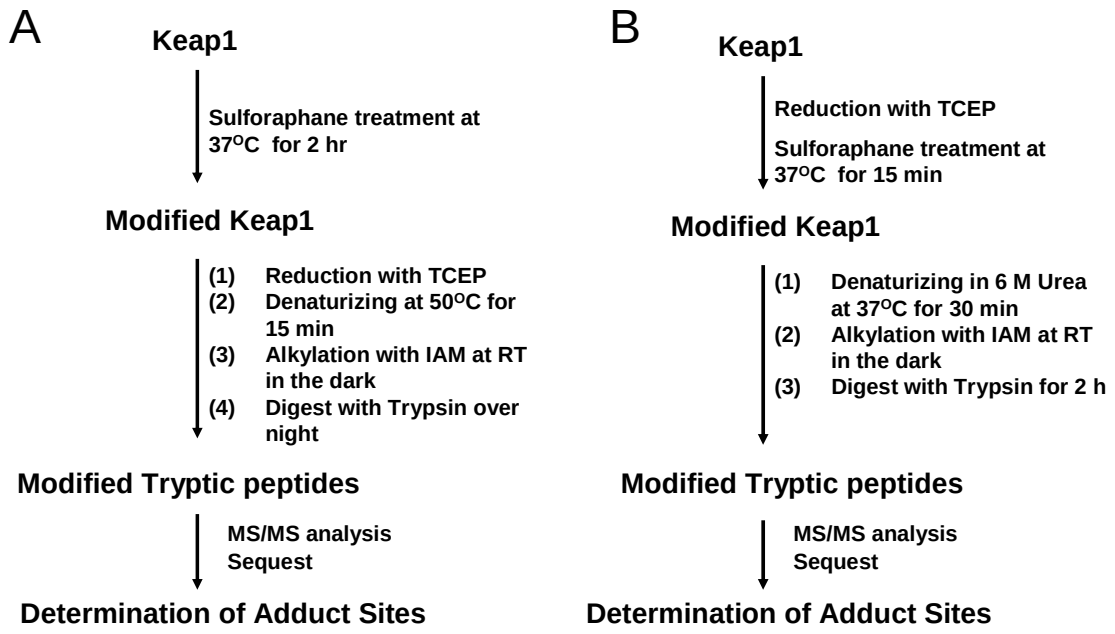
The analysis of sulforaphane cysteine adducts presents a challenge due to the propensity of the adducts to revert to sulforaphane and free cysteine or to hydrolyze (Scheme 1). In our initial attempts to identify targets of sulforaphane adduction on His_6-Keap1 *in vitro*, we employed a standard protocol we used previously (Scheme 2A) to prepare sample for LC-MS-MS analysis (*165*). However, we were unable to detect any sulphoraphane adducts. We suspected that sulforaphane cysteine adducts may have been lost either during the 50°C reduction and denaturation step or during the 16-20 hr tryptic digestion.

To more carefully evaluate sulforaphane cysteine adduct instability during sample workup, we utilized the nine amino acid synthetic tryptic peptide TPepC (AVAGCAGAR), which we have employed previously to model peptide cysteine adduct chemistry (*121;169*). TPepC (1 mM) was treated with 1 mM sulforaphane in 0.1 M ammonium bicarbonate (pH 7.5) at 37 °C for 2 h, which converted most of the TPepC to the sulforaphane cysteine adduct (Fig 3-5a). The MS-MS spectrum (obtained on a triple quadrupole instrument) confirmed adduction on the cysteine residue (Fig 3-6). The spectrum indicates not only the expected b- and y-ions due to peptide fragmentation, but also characteristic neutral losses of sulforaphane from the intact peptide adduct ion and several of the adduct-containing fragment ions. In addition, sulforaphane-derived fragments were observed.

To mimic denaturing step of our standard protocol, sulforaphane-adducted TpepC was heated to 50°C for 15 min in the presence of 15 mM TCEP in 0.1 M ammonium



Scheme 3-1. Sulforaphane cysteine adduct formation and hydrolysis of the adduct to sulforaphane and free cysteine.



Scheme 3-2. Experimental procedures utilized to prepare samples for LC-MS-MS analyses. (A) Prototypical methods used in our lab. (B) Optimized experimental methods by which sulforaphane adducts could be identified with mass spectrometry.

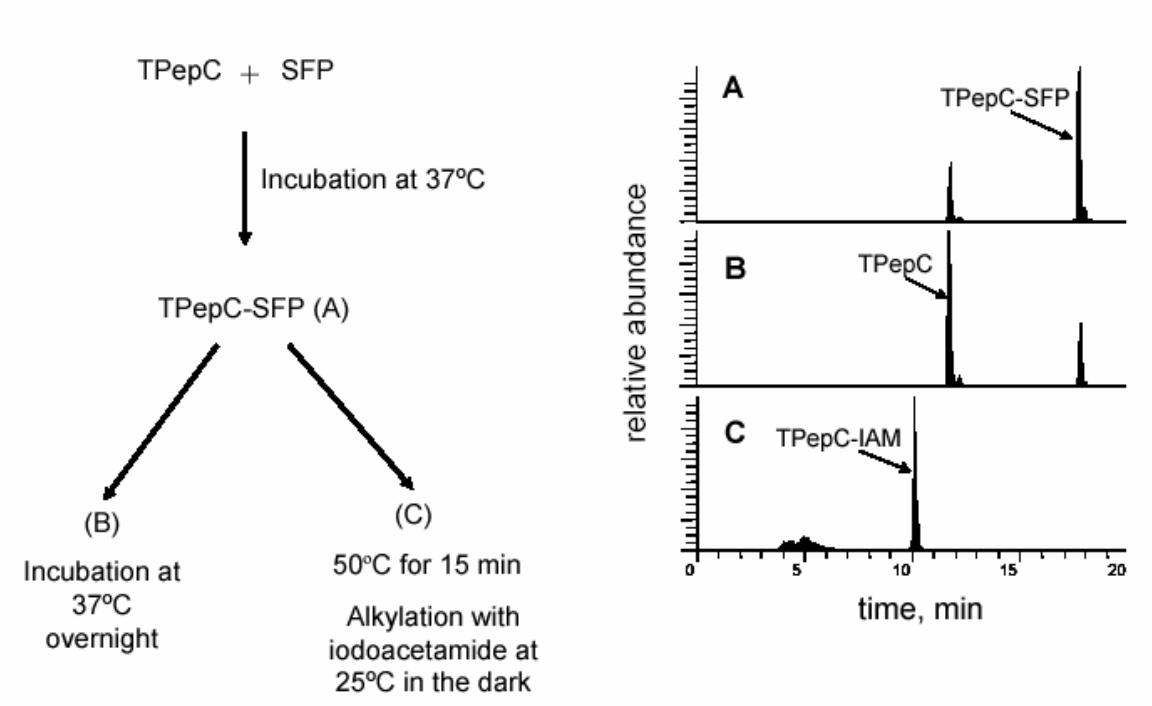


Figure 3-5. Longer incubation and higher temperature facilitate the reversion of sulphoraphane cysteine adduct to sulforaphane and free cysteine. (a) Incubation of sulforaphane and TpepC (molar ratio 1:1) for 2 h at 37^{O}C. (b) Incubation was lengthened to overnight. (c) Sulforaphane adducted TpepC were heated up to 50^{O}C for 15 min.

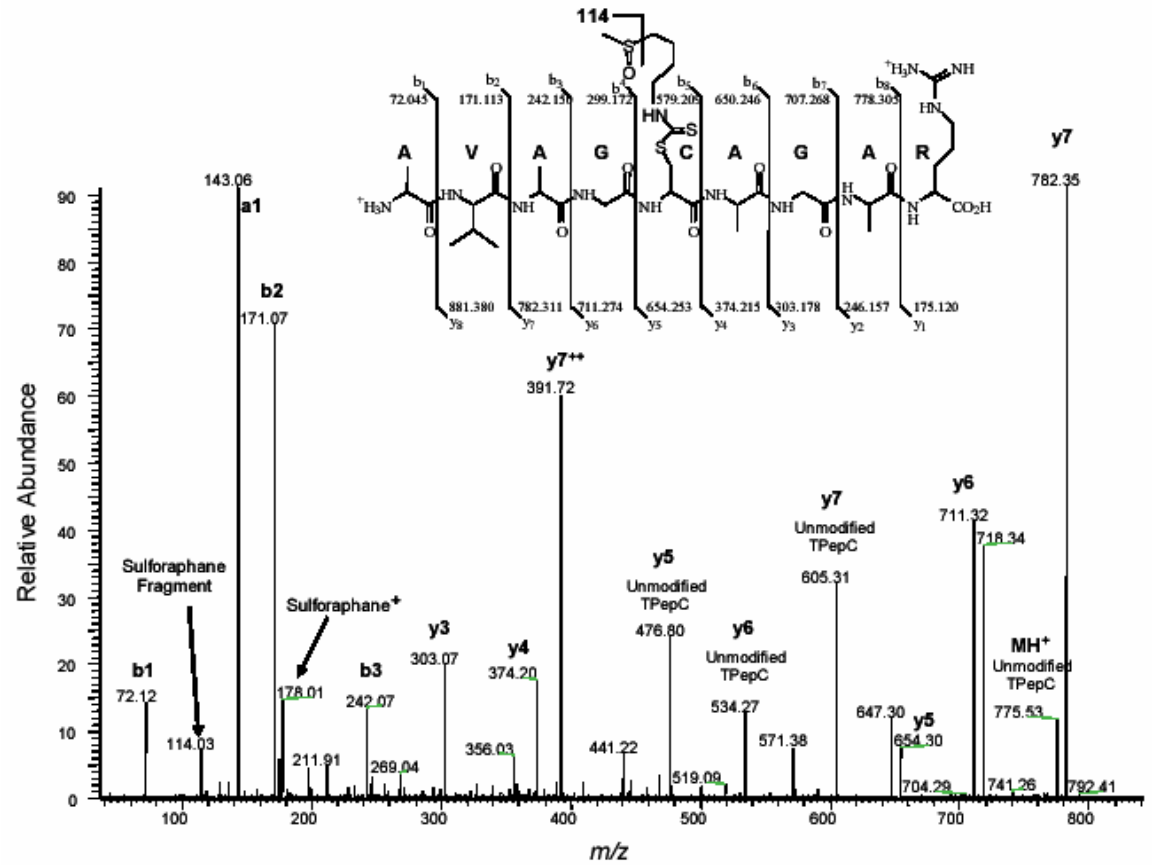


Figure 3-6. MS-MS spectrum of sulforaphane modified TpepC.

bicarbonate, followed by treatment with iodoacetamide at 25°C to alkylate free cysteines. This treatment destroyed all all sulforaphane cysteine adducts and only S-carboxamidomethylated TpepC was observed (Figure 3-5c). To mimic the tryptic digestion step, sulforaphane-adducted TpepC was incubated in 0.1 M ammonium bicarbonate pH 7.5 at 37°C for 16 h, during which most of the adduct was converted back to TpepC (Figure 3-5b). This experiment indicated that sulforaphane cysteine adducts are lost to both the reduction/alkylation step and the tryptic digestion step during sample preparation for LC-MS-MS.

Through a series of experiments summarized in Table 3-1, we optimized conditions for the detection of sulforaphane adduct sites by LC-MS-MS. These experiments generated two key alterations in the workup. The first was a new denaturation/reduction step in which the adducted protein was denatured for 15 min in 0.1 M amonium bicarbonate, pH 7.5 containing 6M urea and 15 mM TCEP at 37°C. The second was a shortening of the tryptic digestion from 16 hr to 2 hr. These changes produced a dramatic improvement in adduct recovery from Keap1 treated with sulforaphane in vitro. Finally, we found that treatment of the Keap1 protein with 5 mM TCEP immediately prior to sulforaphane treatment improved the formation of adducts. When this step was omitted, we only found 3 sulforaphane modified sites on the Keap1 protein and all were located at the very ends of Keap1 (Cys13, Cys14 and Cys624). This suggests that disulfide-mediated aggregation of the cysteine-rich Keap1 protein prevents sulforaphane attacking these cysteines. Coincubation with sulforaphane and TCEP did not interfere with adduct formation, as TCEP is a non-nucleophilic phosphine reductant.

Table 3-1. Procedures to optimize workup conditions to identify SFP adduct sites by LC-MS-MS

Steps	Method				
	A	**B**	**C**	**D**	**E**
Keap1 in 0.1 M Ambic (pH7.5)	+	+	+	+	+
Add 5mM TCEP					+
Add 2mM SFP, incubate at 37 ^{O}C for 15 min	+			+	+
Wash with and resuspend in 0.1 M Ambic (pH7.5)	+			+	+
5mM TCEP, 50 ^{O}C, 15 min	+				
Resuspended in 0.1 M Ambic (pH7.5) containing 6M Urea and 5mM TCEP, 37 ^{O}C, 30 min		+	+	+	+
Wash with and resuspend in 0.1 M Ambic (pH7.5)		+	+	+	+
Add 2mM SFP, incubate at 37 ^{O}C for 15 min		+	+		
Wash with and resuspend in 0.1 M Ambic (pH7.5)		+	+		
Add 20 mM Iodoacetamide, incubate at RT for 15 min	+	+	+	+	+
Wash with and resuspend in 0.1 M Ambic (pH7.5)		+	+		+
Add trypsin (1:50), incubate at 37 ^{O}C for 16 h	+	+			
Add trypsin (1:50), incubate at 37 ^{O}C for 2 h			+	+	+
Detected SPF modified Cysteine residues	**0**	**0**	**16**	**3***	**24**

* Cys 13, Cys 14 and Cys 624

Concentration-dependent formation of sulforaphane cysteine adducts on Keap1

To compare the relative reactivity of cysteines in Keap1 with sulforaphane, His_6-Keap1 was treated with 2 μM, 20 μM or 200 μM sulforaphane for 15 min. Five separate incubations were performed at each sulforaphane concentration and the adducts were mapped by LC-MS-MS analysis. The results of these replicate experiments were found to be quite reproducible (Table 3-2). It should be noted that all Keap1 Cys-containing peptides were detected either as sulforaphane adducts or as S-carboxamidomethylated derivatives. MS-MS spectra of all sulforaphane-adducted peptides are presented as supplementary material (Appendix 2). A notable feature of these spectra is the frequent presence of b- and y-ions for the unmodified peptides and fragmentations involving the sulforaphane moiety. These characteristic fragmentations were noted in the MS-MS spectrum of the TpepC adduct shown in Figure 3-6.

The most reactive cysteines in Keap1 towards sulforphane were judged to be those most reproducibly adducted at the 2 μM concentration. These were located in the Kelch repeat domain, especially Cys489, which was found to be sulforaphane-modified in all experiments at the 2 μM concentration. Cysteines 513, 518 and 583, also in the Kelch domain were modified in at least three experiments, as were cysteines in the central linker domain (Cys226 and Cys249), the BTB domain (Cys77) and the C-terminal domain (Cys624). At higher sulforaphane concentrations, adduction selectivity was increasing less selective, yet reproducible between experiments.

Table 3-2. Modification of Keap1 by sulforaphane as a function of concentration *in vitro*

Keap1 Tryptic Peptides	Cysteine	Domain	Sulforaphane		
			2 µM	20 µM	200 µM
PSGAGACCR	Cys 12	N-terminal	2,4	1,3,5	1,2,3,5
	Cys 13	N-terminal	2,4	3,4	3,4,5
FLPLQSQCPEGAGDAVMYASTECKAEVTPSQHGNR	Cys 23	N-terminal			4,5
	Cys 38	N-terminal	3	1,2,3,4,5	1,2,3,4,5
LSQQLCDVTLQVK	Cys 77	BTB	1,2,4	1,2,3,4,5	1,2,3,4,5
CVLHVMNGAVMYQIDSVVR	Cys 151	BTB		1,5	1,2,3,4,5
ACSDFLVQQLDPSNAIGIANFAEQIGCVELHQR	Cys 171	Central linker	2	1,2,3,4	1,2,3,4,5
	Cys 196	Central linker	3	4	1,2,5
QEEFFNLSHCQLVTLISR	Cys 226	Central linker	1,2,4	1,2,3,5	1,2,3,4,5
CESEVFHACINWVK	Cys 241	Central linker			1,2,3,4,5
	Cys 249	Central linker	1,2,4	1,2,3,4,5	1,2,3,4,5
YDCEQR	Cys 257	Central linker	1,4	1,2,3,4,5	1,2,3,4,5
CHSLTPNFLQMQLQK	Cys 273	Central linker		1,2	1,2,3,4,5
CEILQSDSR	Cys 288	Central linker		1,2,3,4,5	1,2,3,4,5
CK(DYLVK)	Cys 297	Central linker			
IFEELTLHKPTQVMPCR	Cys 319	Central linker			2,3,5
SGLAGCVVGGLLYAVGGR	Cys 368	Kelch	3	1,2,3,4,5	1,2,3,4,5
NNSPDGNTDSSALDCYNPMTNQWSPCAPMSVPR	Cys 395	Kelch		1,5	1,2,3,4,5
	Cys 406	Kelch			2,3,5
IGVGVIDGHIYAVGGSHGCIHHNSVER	Cys 434	Kelch		5	5
LNSAECYYPER	Cys 489	Kelch	1,2,3,4,5	2,3,4,5	1,2,3,4,5
SGAGVCVLHNCIYAAGGYDGQDQLNSVER	Cys 513	Kelch	2,4,5	1,2,3	1,2,3,4,5
	Cys 518	Kelch	1,2,4	1,2,4,5	1,2,3,4,5
IYVLGGYDGHTFLDSVECYDPDTDTWSEVTR	Cys 583	Kelch	1,2,3,4	1,2,3,4,5	1,2,3,4,5
SGVGVAVTMEPCR	Cys 613	C-terminal		5	5
QIDQQNCTC	Cys 622	C-terminal			
	Cys 624	C-terminal	1,2,3,4	1,2,3,4,5	1,2,3,4,5

Discussion

Sulforaphane has become the prototype for a class of isothiocyanate compounds that induce the transcription of ARE-regulated genes encoding carcinogen detoxication and antioxidant enzymes (*157;170*). Sulforaphane is thought to activate ARE regulated genes by modifying Keap1, thus stabilizing Nrf2 and promoting its nuclear translocation (*70;74;75*). However, sulforaphane modifications of Keap1 had not been characterized and thus the mechanisms by which sulforaphane triggers Nrf2 stabilization have remained speculative. Here we applied LC-MS-MS to map Keap1 adducts formed with sulforaphane *in vitro* and we found that sulforaphane adduction selectivity differs significantly from that of other ARE-inducing electrophiles characterized to date. Moreover, sulforaphane adduction of Keap1 does not lead to Keap1 polyubiquitination, which we recently observed with other electrophilic inducers. The results suggest that the Keap1 protein is a versatile electrophile sensor capable of stabilizing Nrf2 by at least two mechanisms.

The function of Keap1 as an electrophile sensor apparently depends on the site-selectivity of electrophile modifications. We have recently shown that adduction site specificity is largely dependent on the chemistry of the electrophile modifier (*165*). Although previous reports have suggested that sulforaphane modifies critical Cys residues in the central linker domain (*69;70;171*), Keap1-sulforaphane adducts have never been directly detected. A major barrier to analysis of these adducts is the lability of the sulforaphane cysteine adducts, which are susceptible to decomposition during standard sample preparation protocols for LC-MS-MS analysis. Here we described a

modified sample workup that minimized adduct decomposition and enabled mapping of sulforaphane cysteine adducts by LC-MS-MS.

How Keap1 modification site selectivity governs Nrf2 stabilization and ARE activation presents an interesting problem. Sulforaphane displayed a different Keap1 modification pattern than did other electrophiles studied previously. Even at very low concentrations (2 μM), sulforaphane modified all five Keap1 domains. Preferred sulforaphane adduction sites included Cys77 in BTB domain, Cys 226, Cys249 and Cys257 in the central linker domain, Cys489, Cys513, Cys518 and Cys583 in the Kelch repeat domain, and Cys624 in the C-terminal domain. This is in marked contrast to modification by the model electrophiles and ARE activators dexamethasone mesylate (Dex-Mes) (*69*) and N-iodoacetyl-N-biotinylhexylenediamine (IAB) (*165*), which modify Keap1 preferentially in the central linker domain. Although we were only able to study Keap1 sulforaphane adducts generated *in vitro*, our previous studies with IAB indicated that adducts formed *in vitro* also were detected on FLAG-Keap1 *in vivo* (*165*).

Our identification of sulforaphane modification sites on Keap1 further underscores the diversity of electrophile adduction patterns on this sensor protein. The ARE activating electrophiles Dex-Mes and IAB display selectivity for Cys residues in the central linker domain, although specificity for targeting the 10 central linker domain Cys residues differs (*69;70;165*). Sulforaphane produces central linker domain adducts, but also adducts in the BTB, Kelch and C-terminal domains. Interestingly, the electrophile 1-biotinamido-4-(4'-[maleimidoethyl-cyclohexane]-carboxamido)butane (BMCC), which is not an ARE activator, modified human Keap1 primarily on cysteines in the BTB and

Kelch repeat domains (*165*). Thus, of the four electrophiles whose Keap1 adducts have been mapped by MS-MS, all display a unique pattern of target modifications. These data indicate that Keap1 modifications are not driven just by high reactivity of specific central linker domain Cys residues, but instead reflect both individual Cys residue reactivities and chemical properties of the electrophiles.

Inferences about critical targets in Keap1 have been drawn from functional analyses of Keap1 mutants. Cys273Ser and Cys288Ser mutations in the central linker domain blocked Keap1-dependent ubiquitination of Nrf2 (*74*) and Cys273Ala and Cys288Ala mutations blocked Keap1-dependent repression of an ARE reporter (*70*). Zhang et al. also reported that Cys151 in the Keap1 BTB domain is required for inhibition of Keap1-dependent degradation of Nrf2 by sulforaphane and oxidative stress, although this mutant does function as a constitutive repressor of Nrf2 (*74;75*). Although these data might suggest that Cys 273, Cys288 and Cys151 would be key targets for electrophiles, MS studies of adducts do not support this view. Although Cys273 and Cys288 are preferred targets for Dex-Mes, only Cys288 is modified by IAB. Neither is modified by sulforaphane, except at high concentrations. Despite being critical to Keap1 function, Cys273 is not a preferred target of either sulforaphane or IAB. Cys151 is equally critical for Nrf2 activation by Keap1 modifiers (*74*), but is itself not a target for electrophiles. Thus, the MS adduct maps clearly show that residues essential for function are not necessarily targets for electrophile modification.

An interesting consequence of Keap1 alkylation by electrophiles *in vivo* is the formation of HMW Keap1 protein forms (*146*), which we have recently shown to be

polyubiquitinated Keap1 (*165*). Studies by several groups have demonstrated that ubiquitination of Nrf2 is mediated through formation of a Keap1-CUL3-Roc1 complex (*75;77;78*). Our identification of ubiquitinated Keap1 following IAB and tBHQ treatment led us to hypothesize that electrophilic modification of sensor cysteines triggers a switching of ubiquitination from Nrf2 to Keap1, thus stabilizing Nrf2 (*165*). Zhang et al. indicated that sulforaphane induced formation of a HMW Keap1 form, but did not present the data (*74*). In our experiments, we were able to detect formation of ubiquitinated Keap1 upon treatment with tBHQ, but not with sulforaphane (Fig. 3-4). This somewhat surprising result suggests that sulforaphane may affect the Keap1-Nrf2 interaction in a different way than other electrophiles studied previously. Our MS data shows that sulforaphane modifies Keap1 in the central linker domain as well as in the Kelch and C-terminal domains. This modification pattern is much less specific than what we observed for IAB (*165*). However sulforaphane adduction of Keap1 may prevent ubiquitination of Nrf2 *without* shifting ubiquitination to Keap1 and this action may be sufficient to stabilize Nrf2.

Another interesting observation from this study is that sulforaphane treatment of cells induced the formation of a slightly higher mass form of Nrf2 in cytoplasm and that only this higher mass Nrf2 was translocated into nuclei (Fig. 3-2). Although we have no data to indicate the nature of the modification in our system, previous studies have shown that PKC directly phosphorylates Nrf2 and that PKC inhibitors significantly suppressed ARE-regulated gene expression (*87*). In Nrf2, Serine 40 in the Neh2 domain plays an important role in Keap1-Nrf2 association (*88*). Studies by several groups confirmed that

phosphorylation of Nrf2 at Ser40 by protein kinase C in response to antioxidants and electrohiles leads to the release of Nrf2 from Keap1 (*89;172*). The phosphatidylinositol-3-kinase (PI3P) pathway was reported to be involved in Nrf2 activation in cells treated with tBHQ (*82;85*). Since atypical protein kinase C (aPKC) is a downstream enzyme of PI3K, Numazawa proposed that stimuli utilize this pathway to initiate ARE regulated gene expression (*173*). These considerations suggest that the higher mass Nrf2 is a phosphorylated form produced by a sulforaphane-triggered PKC phosporylation signaling pathway. Direct analysis of Nrf2 modifications will be needed to test this hypothesis.

CHAPTER FOUR – DISCUSSION

Several studies have shown that activation of Nrf2 by electrophilic phase II inducers is through their direct interaction with Keap1 to form covalent adducts. However, there is no solid evidence to explain how modifications on Keap1 by electrophiles interrupt Keap1-Nrf2 interaction. It was already established by Yamamoto group (*78*) and Hannink group (*75*) that, under basal conditions, Keap1 serves as an adaptor for the ubiquitination and degradation of Nrf2. In our studies, we first reported that ubiquitination of Keap1, which is triggered by modification on the central linker domain of Keap1, is associated with activation and stabilization of Nrf2. We propose that, under electrophile and oxidative stress, adduct formation on the central linker domain of Keap1 interferes with the ubiquitination of Nrf2 and that the ubiquitination is switched from Nrf2 to Keap1, which stabilizes and activates Nrf2.

It was proposed by Talalay's group that the most reactive cysteines in Keap1 may serve as a sensor responsible for the release of Nrf2 from Keap1 triggered by eletrophilic inducers. Four cysteines (Cys257, Cys273, Cys288, and Cys297) of 25 cysteines in murine Keap1 were identified to be the most reactive residues toward Dex-mes (*69*). However, in our studies using phase II inducer IAB and non-inducer BMCC to modify human Keap1, we found different eletrophiles may have different modification patterns on Keap1, which indicates adduct formation is not solely dependent on reactivity of cysteine residue, but also a function of the chemistry of electrohiles. IAB adduction on Keap1 is exclusively on the central linker domain and BMCC adducts on other domains,

which indicates that the central linker domain of Keap1 is the sensor for electrophile-induced Nrf2 activation.

A large number of studies have reported that sulforphane is a chemopreventive agent by inducing expression of phase II enzymes through activation of Nrf2. However, there is no study showing that sulforaphane can modify Keap1. In our studies, we found that sulforaphane adducts on keap1 are unstable, which prevented identification of sulforaphane adduct sites with the typical procedure to prepare analyte for LC-MS-MS analysis used in our lab. With a modified method, we finally identified sulforaphane modification pattern on Keap1. Sulforaphane is a very powerful electrophile and can modify most of the cysteine residues on Keap1 at a concentration of 20 μM within 15 min. Unexpectedly, the most reactive cysteines are located on the Kelch repeat domain, even though cysteines on the central linker domain are still easily attacked by sulforaphane. Moreover, in this study, it was the first time Cys151 was identified to be modified by electrophile although Cys151, compared to other cysteine residues, is relatively unreactive toward sulforaphane modification. Cys151 was demonstrated by Zhang et al. to be critical for Nrf2 activation by Keap1 modifiers (*75;146*). Our MS data show that residues essential for function are not necessarily targets for electrophile modification. In addition, with immunoblotting analysis of cytosolic and nuclear proteins in sulforaphane treated and untreated controls, we found that sulforaphane triggered formation of a slightly higher mass Nrf2 and that only the high mass Nrf2 form was transferred to the nucleus for ARE-regulated gene activation. We suspect that the high mass Nrf2 is its phosporylated form, although this certainly will require experimental

verification. However, this suggestion appears reasonable because many kinase pathways seem to be redox-sensitive. For example, protein tyrosine phosphatases contain cysteine residues that are prone to oxidation and alkylation (*174*), which may lead to sustained tyrosine phosphorylation and extended activation of singaling pathways such as MAPK and PI3K pathways.

Recently, there are more and more studies have focused on exploring the mechanism of the Keap1-Nrf2 signaling pathway. New developments in this field include a report that Keap1 is a zinc metalloprotein (*171*) and evidence of Keap1 as a shuttle for Nrf2 into and out of the nucleus (*137*) .

The theory of disulfide bond formation induced by eletrophiles for Nrf2 activation was proposed by Talalay's group (*70*). Further examination of Keap1 protein by the same group has shown that Keap1 is a zinc-thiol protein that controlled by both metal-binding and thiol reactivity (*171*). Studies by Dinkova-Kostova have identified four cysteines of 25 cysteines in murine Keap1 the most reactive residues to Dex-mes (*69*). Further examination of these cysteines by the same group showed that Cys273 and Cys288 are related to the release of Nrf2 from Keap1 sequestration and it was proposed disulfide bond formation between these two cysteines intermolecularly leads to a Keap1 conformational change and thereafter the release and activation of Nrf2 (*70*). To solve the problem that what distinguishes these cysteines in terms of reactivity, Dinkova-Kostova (*171*) reported recent studies that indicate Keap1 is a zinc metalloprotein and that reactivity of cysteine residues is modulated by zinc binding. They proposed that "involvement in metal coordination could keep the cysteine residues from being hyper-

reactive under basal conditions yet be poised to react with an inducer because they are already in their thiolate state and do not require deprotonation". However, the experimental results may not be adequate and accurate for drawing this conclusion. First of all, the author calls all the electrophiles (Dex-mes, H_2O_2 and 4,4′-dipyridyl disulfide) used in this study inducers. But "inducer" means that the chemical can induce the expression of ARE regulated genes. None of these chemicals have been proved to be an inducer *in vivo* studies. Second, in the paper, it is claimed that different inducers (sulforaphane, H_2O_2 and 4,4′-dipyridyl disulfide) can trigger the release of zinc from Keap1. However, no data shows suforaphane has such an effect. Third, in a experiment to examine whether Keap1 competes with the chelator 4-(2-pyridylazo)resorcinol (PAR) for zinc binding, 100 μM sulforaphane was added twice to 5 μM Keap1 to show addition of sulforaphane release zinc from Keap1. However, their data have not shown a statistically significant change in the amount of released zinc within 20 min after the first addition of sulforaphane. Whereas in our studies, sulforaphane has been shown to be a very powerful electrophile and can modify most of cysteine residues in Keap1 at the concentration of 20 μM within 15 min. Fourth, in their study, 4,4′-dipyridyl disulfide was used to induce disulfide bond formation on Keap1 inter- or intramolecularly and demonstrate that the formation of disulfide bonds can change Keap1 conformation. However, since 4,4′-dipyridyl disulfide is not a phase II inducer, it is not clear whether these results would hold with true inducers such as sulforaphane. These considerations suggest that there is no adequate evidence to support the conclusion that inducer activates Nrf2 by binding zinc bonded cysteines.

Current models established by Yamamoto (*144*) group propose that Keap1 sequester Nrf2 to the actin cytoskeleton and serve as an adaptor to target Nrf2 for ubiquitination and proteasomal degradation. However, studies by Velichkova and Hasson (*137*) have shown that, instead of anchoring Nrf2 in the cytoskeleton, Keap1 is located in cytoplasm through an active nuclear export pathway. They demonstrated that the central linker domain of Keap1 contains a nuclear export signal (NES) that is dependent on Crm/exportin1 pathway and that deletion of the NES region results in nuclear accumulation of both Keap1 and Nrf2. They proposed a new model to explain how Keap1 regulates Nrf2. A complex of Keap1 and Nrf2 can move between the nucleus and the cytoplasm. Under basal conditions, the Keap1-Nrf2 complex is dominantly located in the cytoplasm because of the NES sequence and Nrf2 is targeted for ubiquitination and degradation in the cytoplasm. A very small fraction of Keap1-Nrf2 complex is translocated to the nucleus for the expression of ARE genes. Under oxidative stress, nuclear export pathway for Keap1 is blocked, probably because oxidation of Keap1 changes Keap1 conformation, which masks NES sequence on Keap1. In that case, both Nrf2 and Keap1 accumulate in the nucleus. However, nuclear accumulation of Nrf2 alone is not sufficient to induce the expression of ARE-regulated GST gene. It is highly possible that phosphorylation of Nrf2 by PKC account for the dissociation of Nrf2 from Keap1. Velichkova et al. proposed that the conserved NES is between amino acids 272 and 312 in the central linker domain. Our data have shown electrophilic inducer IAB can modify cysteine residue 288 *in vivo* within this proposed NES, which suggests that

adduction of Keap1 disrupts recognition of the NES sequence in Keap1 and thereafter the accumulation of Keap1-Nrf2 complex in nuclei.

In chapter one, we have introduced some typical phase II inducers. However, all of them are exogenous chemicals. Endogenous substances were also found to activate Nrf2. 15-Deoxy-$\Delta^{12,14}$-prostaglanin J_2 (15d-PGJ_2) is synthesized from arachidomic acid by the action of cyclooxygenase (COX). 15d-PGJ_2 was found to be a potent phase II inducer by its directly modifying Keap1 and thereafter triggering the release of Nrf2 from Keap1 (*175*). Another endogenous inducer of Nrf2 is nitric oxide (NO). When endothelial (*176*) and neuroblastoma cells (*177*) were exposed to NO donors, Nrf2 accumulated in the nucleus and induced its regulated genes. Our collaborators in Dr. Michael Freeman's group at Vanderbilt found that oxidized fish oil possesses the ability to induce the expression of the Nrf2- and ARE dependent gene heme oxygenase 1. When FLAG-Keap1 transfected HEK293 cells were exposed to 50 μM oxidized fish oil mixture for 2 h, I found that Nrf2 was translocated to the nucleus, which was coincident with the formation of ubiquitinaion of Keap1 (data not shown).

Cellular stress responses are important mechanisms for cells to defend themselves against damage to macromolecules. Cellular stress response assesses and counteracts stress-induced damage. When the damage is not tolerated, damaged cells will be removed by programmed cell death, apoptosis. An important form of damage is macromolecule adduction by electrophiles. However, the specific mechanisms by which electrophile adducts trigger specific responses were largely unknown. In our studies, we found that specific patterns of Keap1 adduction result in its ubiquitination. Through this

ubiquitination mechanism, Keap1 could be degraded by proteasomal system and its associated transcription factor Nrf2 can be activated to initiate the transcription of ARE-regulated genes. Thus through the studies of Keap1-Nrf2 system, the new mechanism of stress response was discovered.

APPENDIX 1

Index of detected electrophile modified keap1 peptides

IAB modified peptides

a. C_{288}EILQSDSR
b. C_{241}ESEVFHACINWVK
c. YDC_{257}EQR
d. LSQQLC_{77}DVTLQVK
e. ACSDFLVQQLDPSNAIGIANFAEQIGC_{196}VELHQR (1: singly charged; 2: doubly charged)
f. QEEFFNLSHC_{226}QLVTLISR (1: singly charged; 2: doubly charged)
g. IFEELTLHKPTQVMPC_{319}R
h. SGLAGC_{368}VVGGLLYAVGGR
i. LNSAEC_{489}YYPER
j. SGAGVCVLHNC_{518}IYAAGGYDGQDQLNSVER (1: singly charged; 2: doubly charged)
k. IYVLGGYDGHTFLDSVEC_{583}YDPDTDTWSEVTR (1: singly charged; 2: doubly charged)
l. KQIDQQNC_{622}TC

BMCC modified peptides

m. LSQQLC$_{77}$DVTLQVK

n. ACSDFLVQQLDPSNAIGIANFAEQIGC$_{196}$VELHQR (1: singly charged; 2: doubly charged)

o. CESEVFHAC$_{249}$INWVK

p. SGLAGC$_{368}$VVGGLLYAVGGR

q. LNSAEC$_{489}$YYPER

Seq	#	b	y	(+1)
C#	1	486.0	1432.5	9
E	2	615.1	947.5	8
I	3	728.1	818.4	7
L	4	841.2	705.4	6
Q	5	969.3	592.3	5
S	6	1056.3	464.2	4
D	7	1171.3	377.2	3
S	8	1258.4	262.2	2
R	9	1414.5	175.1	1

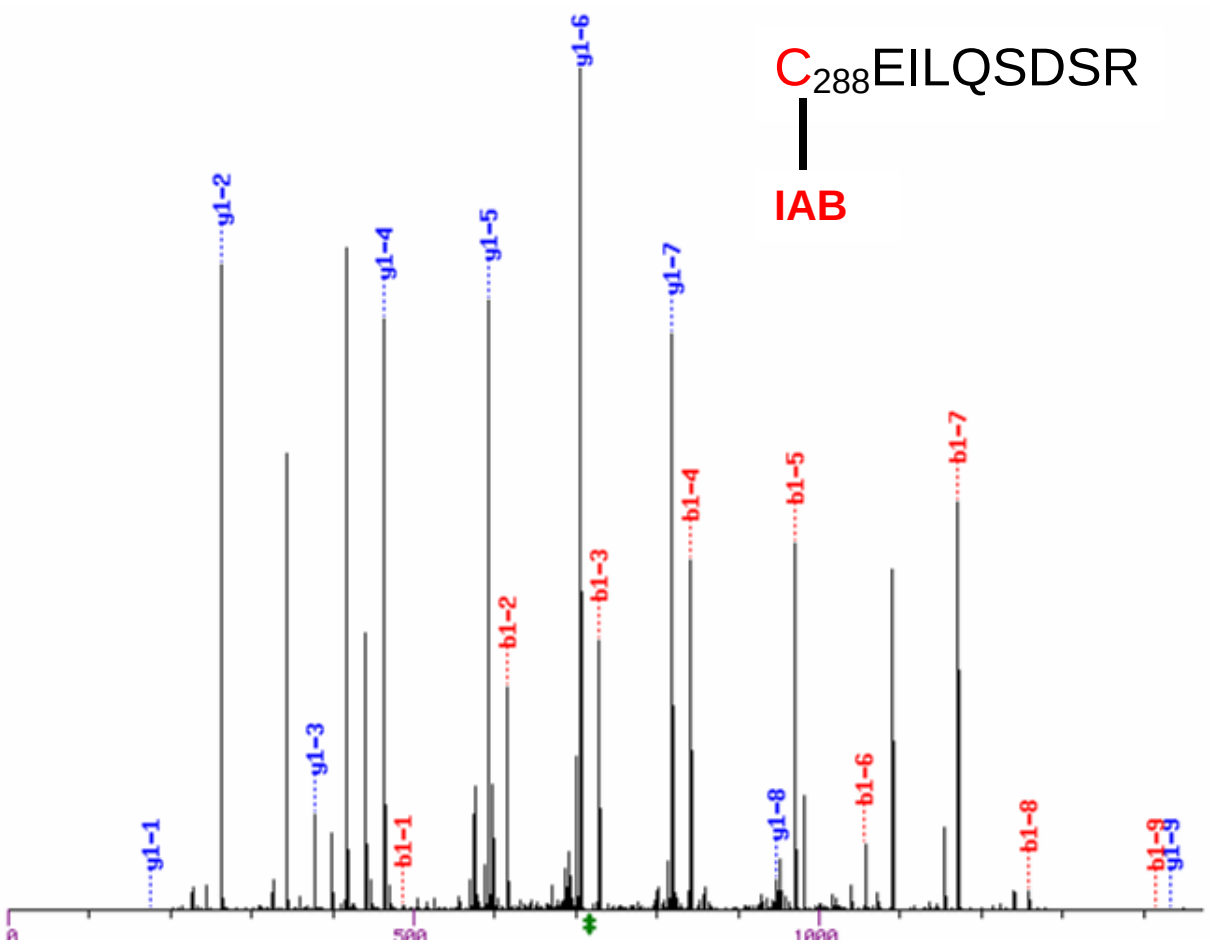


Figure a

Seq	#	b	y	(+1)
C@	1	487.0	2104.8	14
E	2	616.1	1618.7	13
S	3	703.1	1489.7	12
E	4	832.1	1402.7	11
V	5	931.2	1273.6	10
F	6	1078.3	1174.6	9
H	7	1215.3	1027.5	8
A	8	1286.4	890.4	7
C~	9	1446.4	819.4	6
I	10	1559.5	659.4	5
N	11	1673.5	546.3	4
W	12	1859.6	432.3	3
V	13	1958.7	246.2	2
K	14	2086.7	147.1	1

Seq	#	b	y	(+2)
C@	1	244.0	1052.9	14
E	2	308.5	809.9	13
S	3	352.0	745.4	12
E	4	416.6	701.8	11
V	5	466.1	637.3	10
F	6	539.6	587.8	9
H	7	608.2	514.3	8
A	8	643.7	445.7	7
C~	9	723.7	410.2	6
I	10	780.2	330.2	5
N	11	837.3	273.7	4
W	12	930.3	216.6	3
V	13	979.8	123.6	2
K	14	1043.9	74.1	1

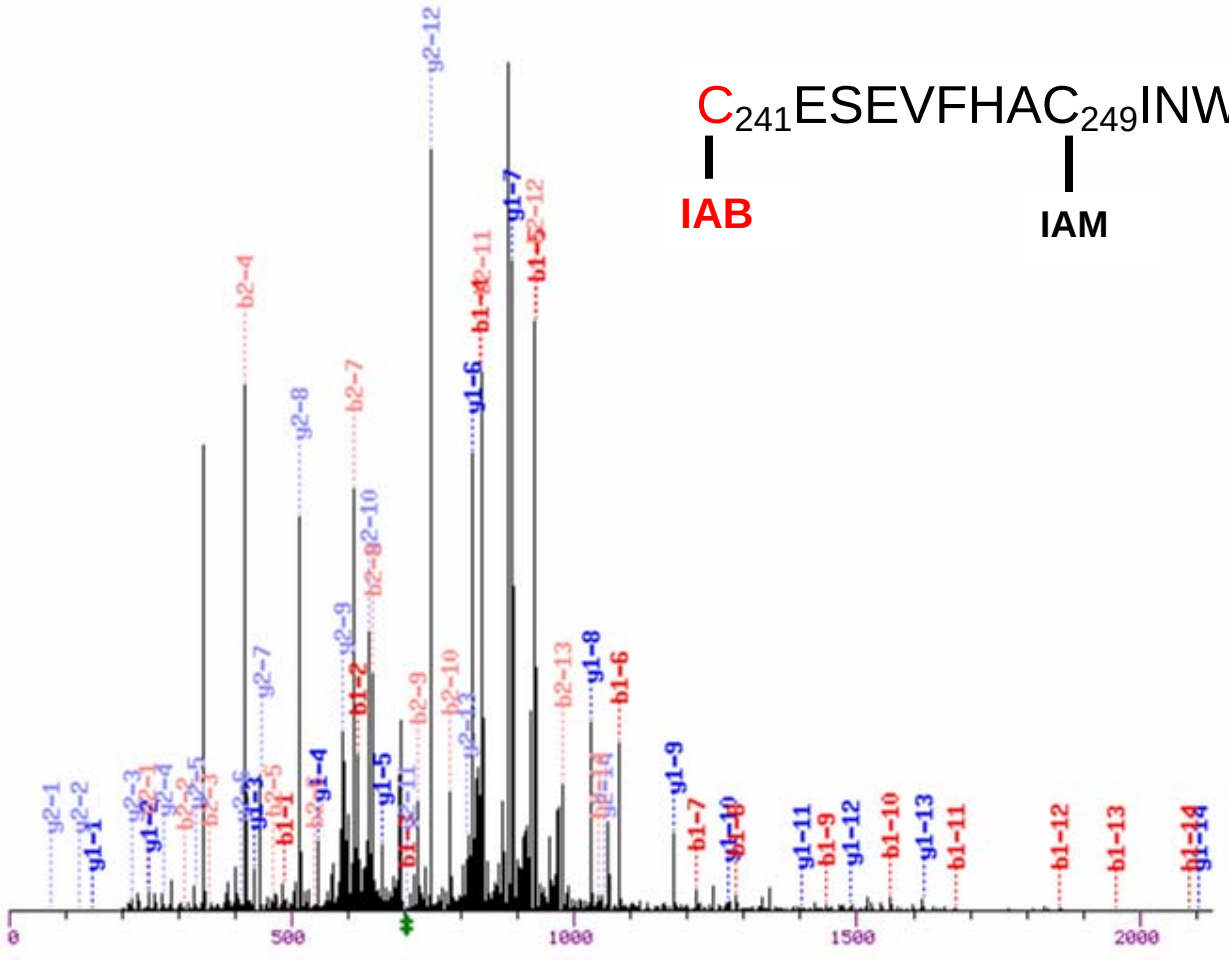


Figure b

Seq	#	b	y	(+1)
Y	1	164.1	1195.3	6
D	2	279.1	1032.3	5
C#	3	764.1	917.2	4
E	4	893.1	432.2	3
Q	5	1021.2	303.2	2
R	6	1177.3	175.1	1

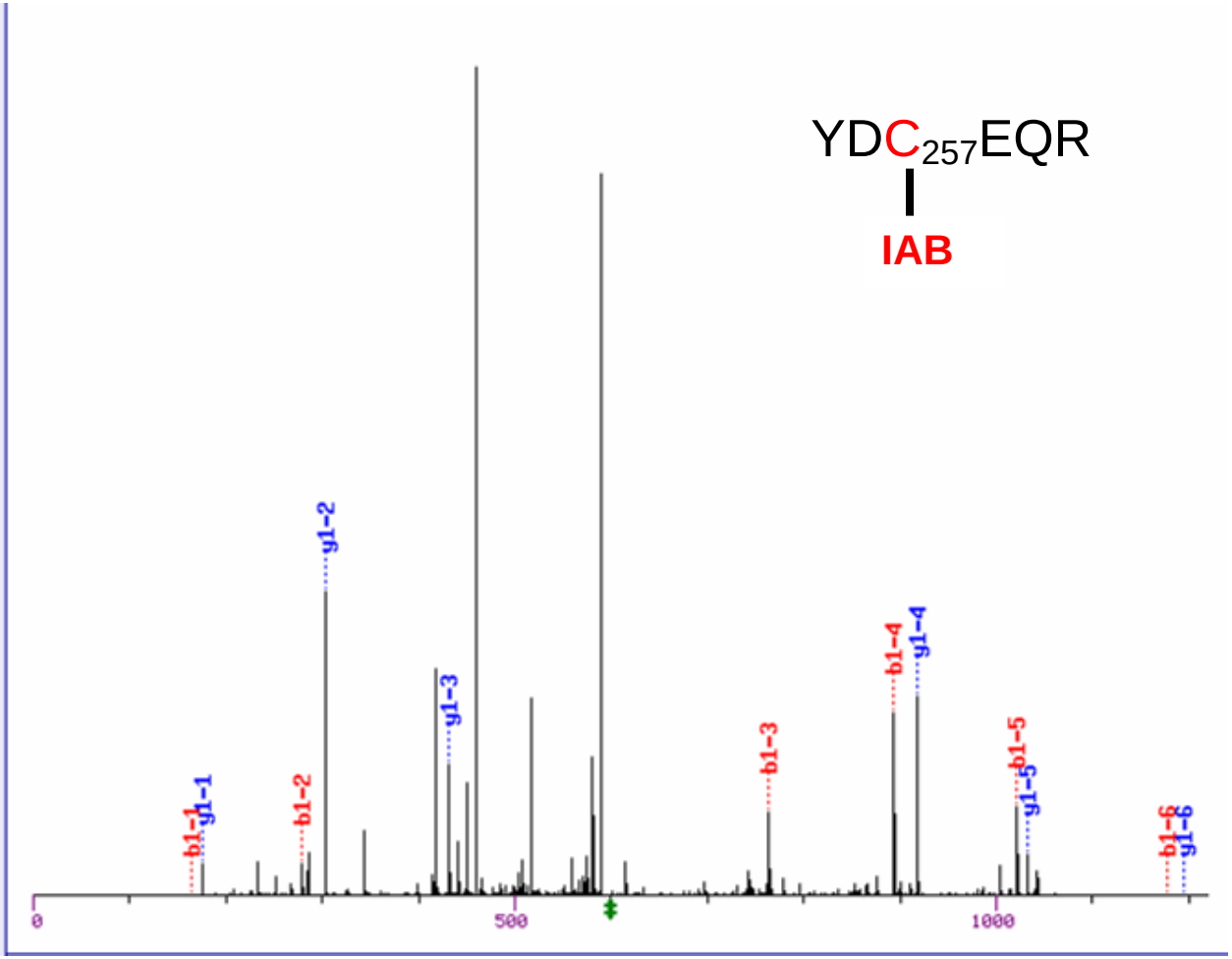


Figure c

Seq	#	b	y	(+1)
L	1	114.1	1856.8	13
S	2	201.1	1743.7	12
Q	3	329.2	1656.7	11
Q	4	457.2	1528.6	10
L	5	570.3	1400.6	9
C#	6	1055.3	1287.5	8
D	7	1170.4	802.5	7
V	8	1269.4	687.4	6
T	9	1370.5	588.4	5
L	10	1483.6	487.3	4
Q	11	1611.6	374.2	3
V	12	1710.7	246.2	2
K	13	1838.8	147.1	1

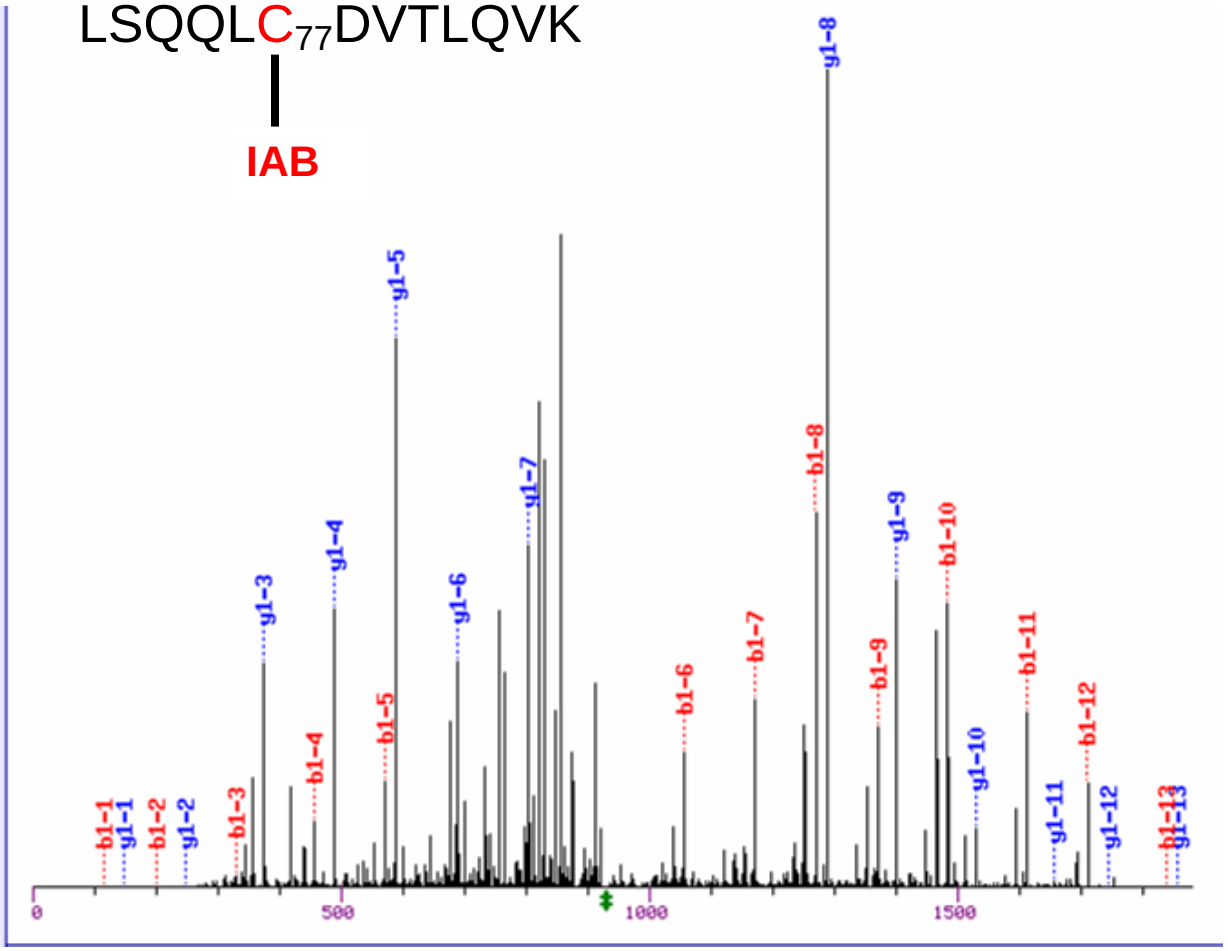


Figure d

Seq	#	b	y	(+1)
A	1	72.0	4026.7	33
C~	2	232.1	3955.7	32
S	3	319.1	3795.7	31
D	4	434.1	3708.7	30
F	5	581.2	3593.6	29
L	6	**694.3**	3446.6	28
V	7	**793.3**	3333.5	27
Q	8	**921.4**	3234.4	26
Q	9	**1049.5**	3106.4	25
L	10	**1162.5**	2978.3	24
D	11	**1277.6**	2865.2	23
P	12	**1374.6**	2750.2	22
S	13	1461.6	2653.1	21
N	14	1575.7	2566.1	20
A	15	**1646.7**	2452.1	19
I	16	**1759.8**	2381.0	18
G	17	1816.8	2267.9	17
I	18	**1929.9**	2210.9	16
A	19	2001.0	2097.8	15
N	20	2115.0	2026.8	14
F	21	2262.1	**1912.8**	13
A	22	2333.1	**1765.7**	12
E	23	2462.1	**1694.6**	11
Q	24	2590.2	1565.6	10
I	25	2703.3	**1437.5**	9
G	26	2760.3	**1324.5**	8
C@	27	3246.3	1267.4	7
V	28	3345.4	**781.4**	6
E	29	3474.4	**682.4**	5
L	30	3587.5	**553.3**	4
H	31	3724.6	440.2	3
Q	32	3852.6	303.2	2
R	33	4008.7	175.1	1

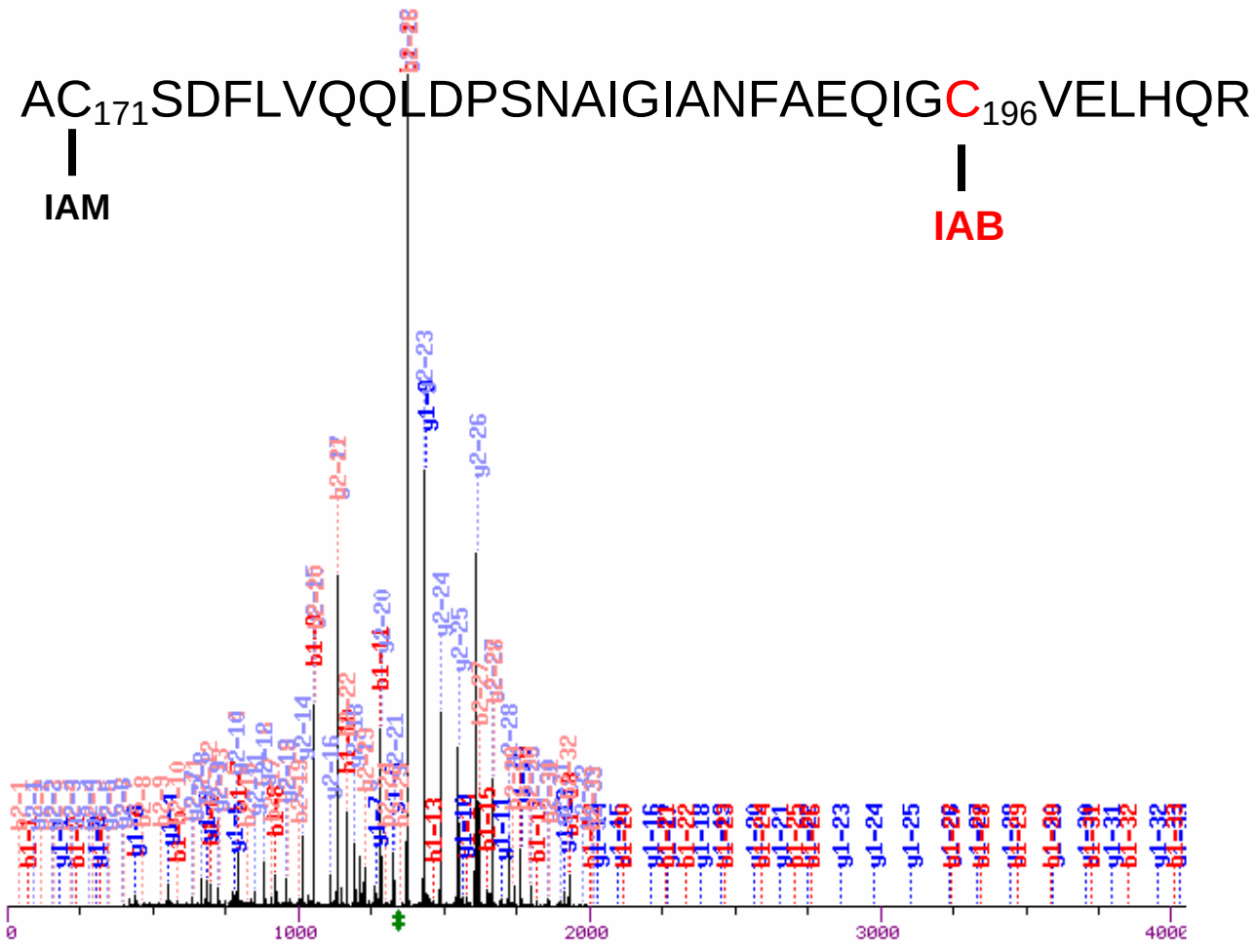


Figure e1

Seq	#	b	y	(+2)
A	1	36.5	2013.9	33
C~	2	116.5	1978.4	32
S	3	160.0	1898.4	31
D	4	217.6	1854.8	30
F	5	291.1	**1797.3**	29
L	6	347.6	**1723.8**	28
V	7	397.2	**1667.2**	27
Q	8	461.2	**1617.7**	26
Q	9	525.2	**1553.7**	25
L	10	581.8	**1489.7**	24
D	11	639.3	**1433.1**	23
P	12	687.8	**1375.6**	22
S	13	731.3	**1327.1**	21
N	14	788.3	**1283.6**	20
A	15	823.9	**1226.5**	19
I	16	880.4	**1191.0**	18
G	17	908.9	**1134.5**	17
I	18	965.5	**1106.0**	16
A	19	1001.0	**1049.4**	15
N	20	1058.0	**1013.9**	14
F	21	1131.5	**956.9**	13
A	22	1167.1	**883.3**	12
E	23	1231.6	**847.8**	11
Q	24	1295.6	**783.3**	10
I	25	1352.1	719.3	9
G	26	1380.7	**662.7**	8
C@	27	1623.7	634.2	7
V	28	1673.2	391.2	6
E	29	1737.7	341.7	5
L	30	1794.3	277.2	4
H	31	1862.8	220.6	3
Q	32	**1926.8**	152.1	2
R	33	2004.9	88.1	1

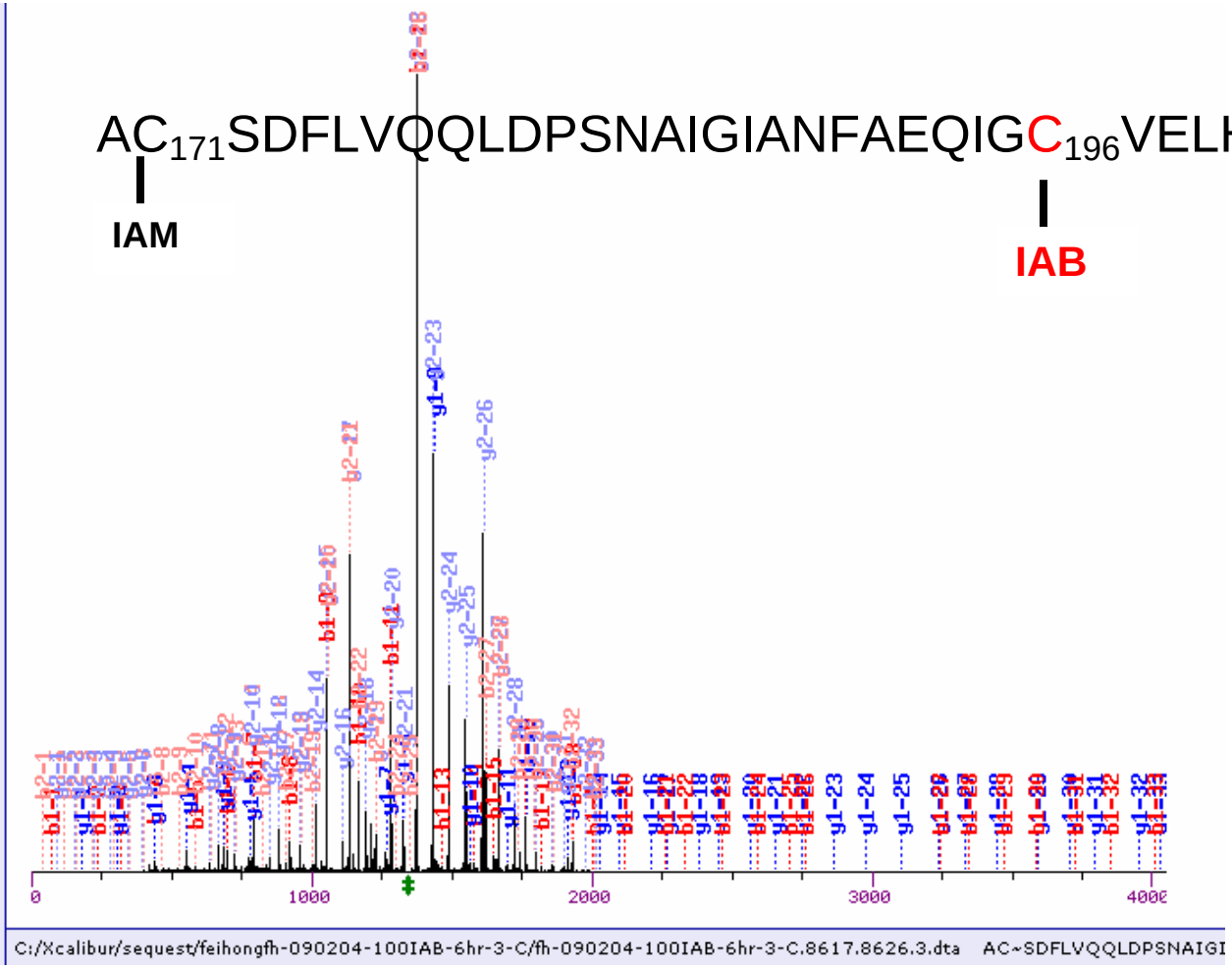


Figure e2

Seq	#	b	y	(+1)
Q	1	129.1	2547.1	18
E	2	258.1	2419.0	17
E	3	387.2	2290.0	16
F	4	534.2	2160.9	15
F	5	681.3	2013.9	14
N	6	795.3	1866.8	13
L	7	908.4	1752.8	12
S	8	995.4	1639.7	11
H	9	1132.5	1552.6	10
C@	10	1618.5	1415.6	9
Q	11	1746.6	929.6	8
L	12	1859.7	801.5	7
V	13	1958.7	688.4	6
T	14	2059.8	589.4	5
L	15	2172.9	488.3	4
I	16	2285.9	375.2	3
S	17	2373.0	262.2	2
R	18	2529.1	175.1	1

Seq	#	b	y	(+2)
Q	1	65.0	1274.0	18
E	2	129.6	1210.0	17
E	3	194.1	1145.5	16
F	4	267.6	1081.0	15
F	5	341.1	1007.4	14
N	6	398.2	933.9	13
L	7	454.7	876.9	12
S	8	498.2	820.3	11
H	9	566.8	776.8	10
C@	10	809.8	708.3	9
Q	11	873.8	465.3	8
L	12	930.3	401.3	7
V	13	979.9	344.7	6

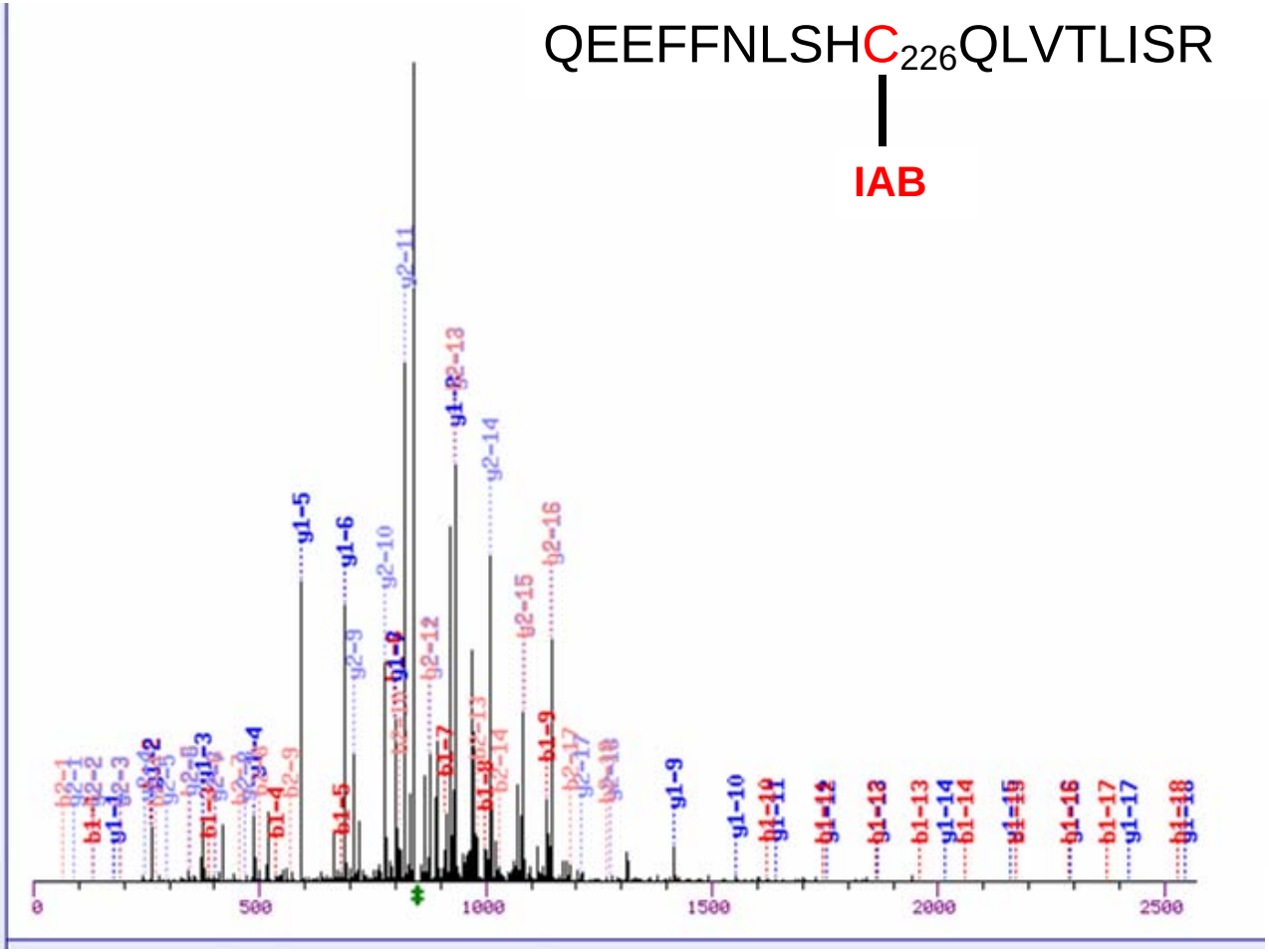


Figure f1

F	4	534.2	2160.9	15
F	5	681.3	2013.9	14
N	6	795.3	1866.8	13
L	7	908.4	1752.8	12
S	8	995.4	1639.7	11
H	9	1132.5	1552.6	10
C@	10	1618.5	1415.6	9
Q	11	1746.6	929.6	8
L	12	1859.7	801.5	7
V	13	1958.7	688.4	6
T	14	2059.8	589.4	5
L	15	2172.9	488.3	4
I	16	2285.9	375.2	3
S	17	2373.0	262.2	2
R	18	2529.1	175.1	1

Seq	#	b	y	(+2)
Q	1	65.0	1274.0	18
E	2	129.6	1210.0	17
E	3	194.1	1145.5	16
F	4	267.6	1081.0	15
F	5	341.1	1007.4	14
N	6	398.2	933.9	13
L	7	454.7	876.9	12
S	8	498.2	820.3	11
H	9	566.8	776.8	10
C@	10	809.8	708.3	9
Q	11	873.8	465.3	8
L	12	930.3	401.3	7
V	13	979.9	344.7	6
T	14	1030.4	295.2	5
L	15	1086.9	244.7	4
I	16	1143.5	188.1	3
S	17	1187.0	131.6	2
R	18	1265.0	88.1	1

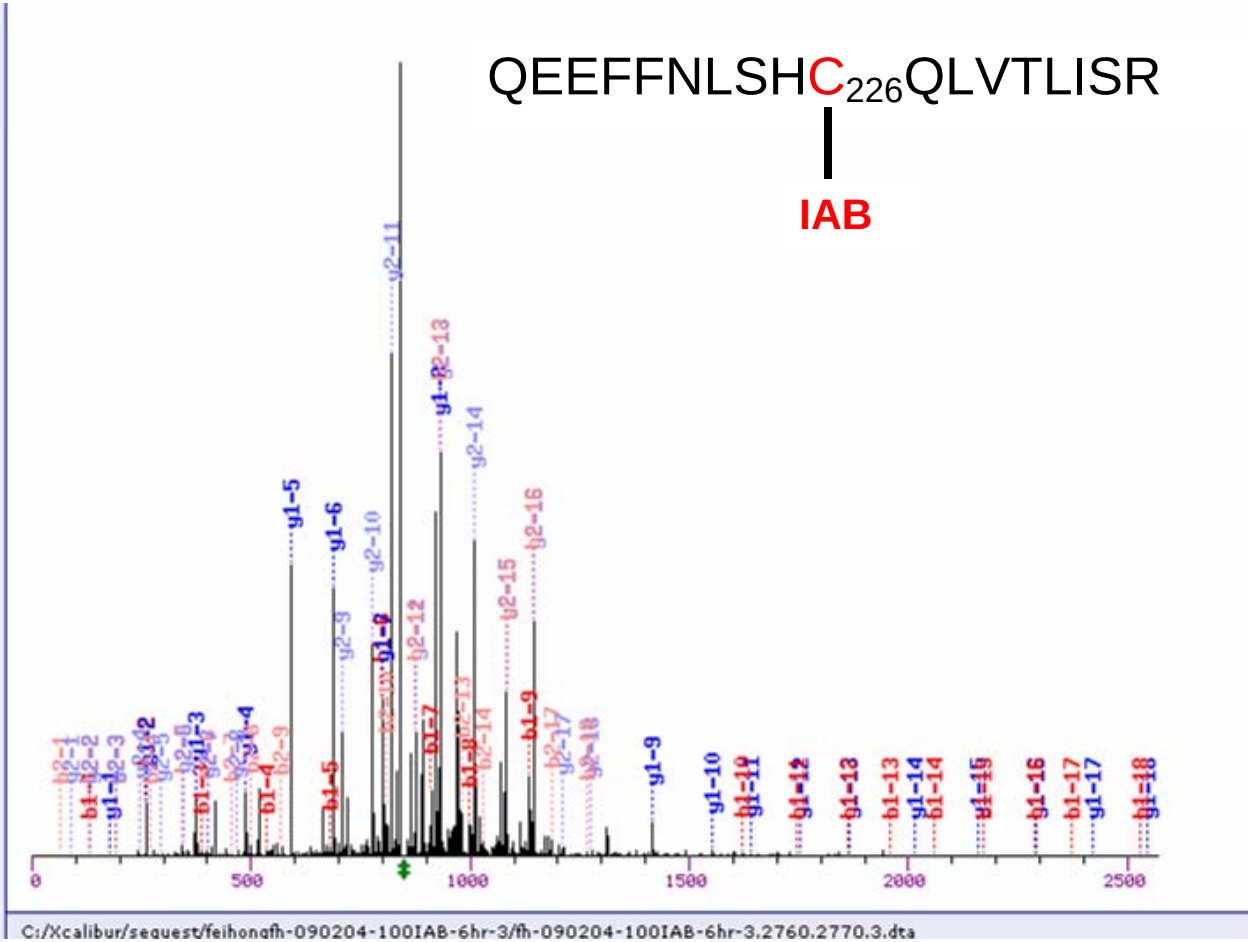


Figure f2

Seq	#	b	y	(+1)
I	1	114.1	2424.1	17
F	2	261.2	2311.0	16
E	3	390.2	2163.9	15
E	4	519.2	2034.9	14
L	5	632.3	1905.8	13
T	6	733.4	1792.7	12
L	7	846.5	1691.7	11
H	8	983.5	1578.6	10
K	9	1111.6	1441.5	9
P	10	1208.7	1313.4	8
T	11	1309.7	1216.4	7
Q	12	1437.8	1115.3	6
V	13	1536.8	987.3	5
M	14	1667.9	888.2	4
P	15	1764.9	757.2	3
C#	16	2249.9	660.1	2
R	17	2406.0	175.1	1

Seq	#	b	y	(+2)
I	1	57.5	1212.5	17
F	2	131.1	1156.0	16
E	3	195.6	1082.5	15
E	4	260.1	1017.9	14
L	5	316.7	953.4	13
T	6	367.2	896.9	12
L	7	423.7	846.3	11
H	8	492.3	789.8	10
K	9	556.3	721.3	9
P	10	604.8	657.2	8
T	11	655.4	608.7	7
Q	12	719.4	558.2	6
V	13	768.9	494.1	5
M	14	834.4	444.6	4
P	15	883.0	379.1	3
C#	16	1125.5	330.6	2
R	17	1203.5	88.1	1

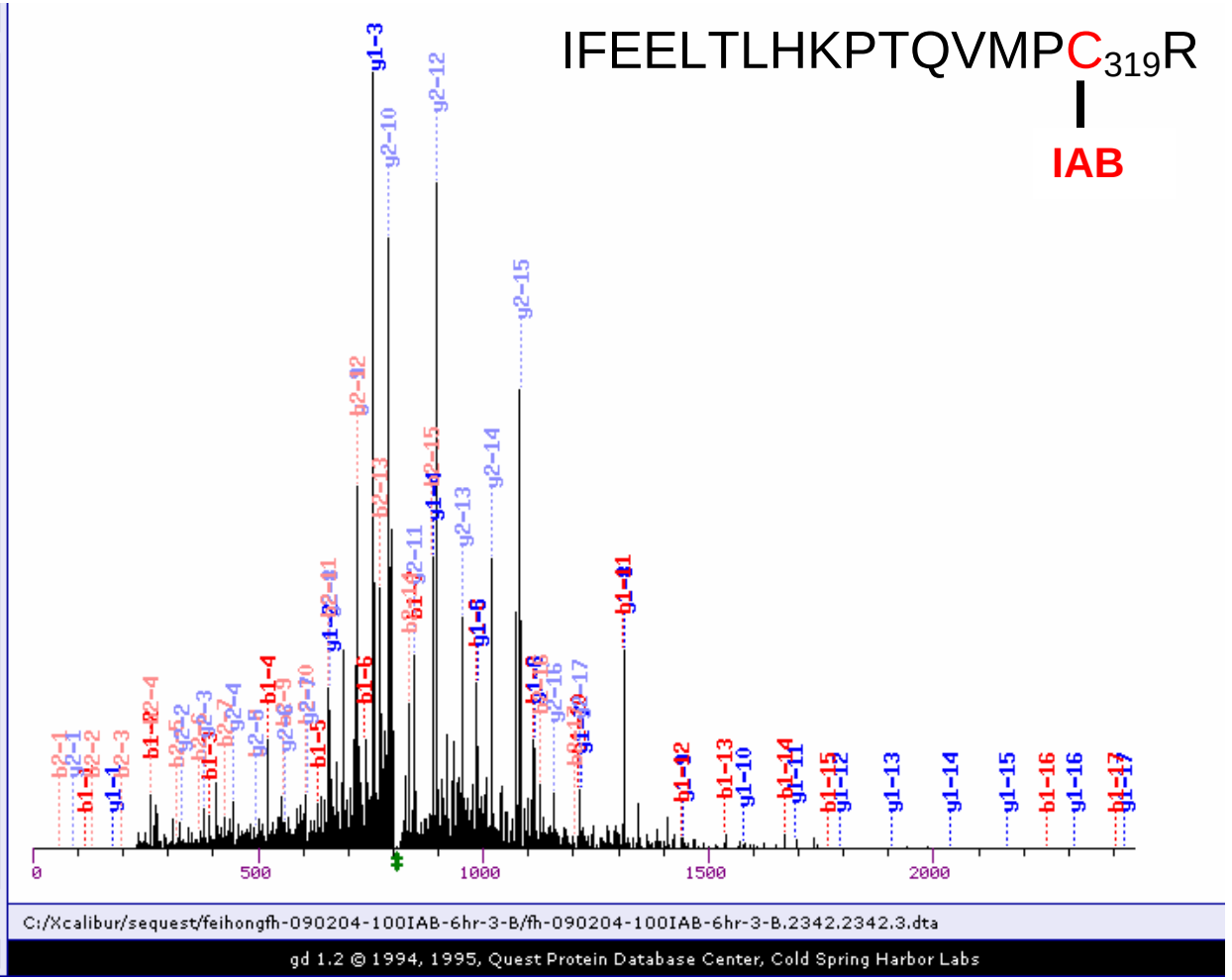


Figure g

Seq	#	b	y	(+1)
S	1	88.0	2030.9	18
G	2	145.1	1943.9	17
L	3	258.1	1886.8	16
A	4	329.2	1773.7	15
G	5	386.2	1702.7	14
C#	6	871.2	1645.7	13
V	7	970.3	1160.7	12
V	8	1069.3	1061.6	11
G	9	1126.4	962.5	10
G	10	1183.4	905.5	9
L	11	1296.5	848.5	8
L	12	1409.6	735.4	7
Y	13	1572.6	622.3	6
A	14	1643.7	459.3	5
V	15	1742.7	388.2	4
G	16	1799.8	289.2	3
G	17	1856.8	232.1	2
R	18	2012.9	175.1	1

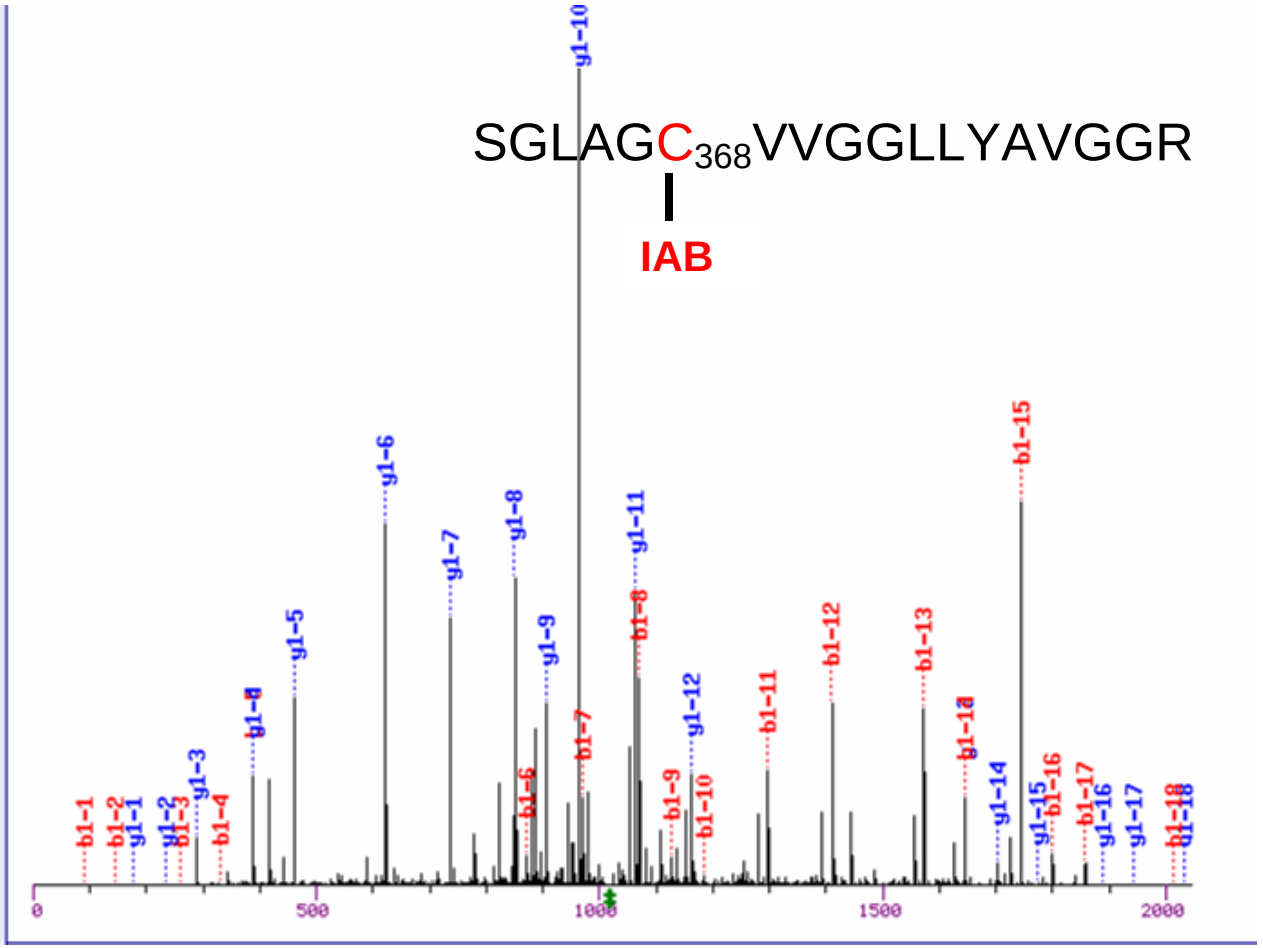


Figure h

Seq	#	b	y	(+1)
L	1	114.1	1727.6	11
N	2	228.1	1614.5	10
S	3	315.2	1500.5	9
A	4	386.2	1413.4	8
E	5	515.2	1342.4	7
C@	6	1001.3	1213.4	6
Y	7	1164.3	727.3	5
Y	8	1327.4	564.3	4
P	9	1424.4	401.2	3
E	10	1553.5	304.2	2
R	11	1709.6	175.1	1

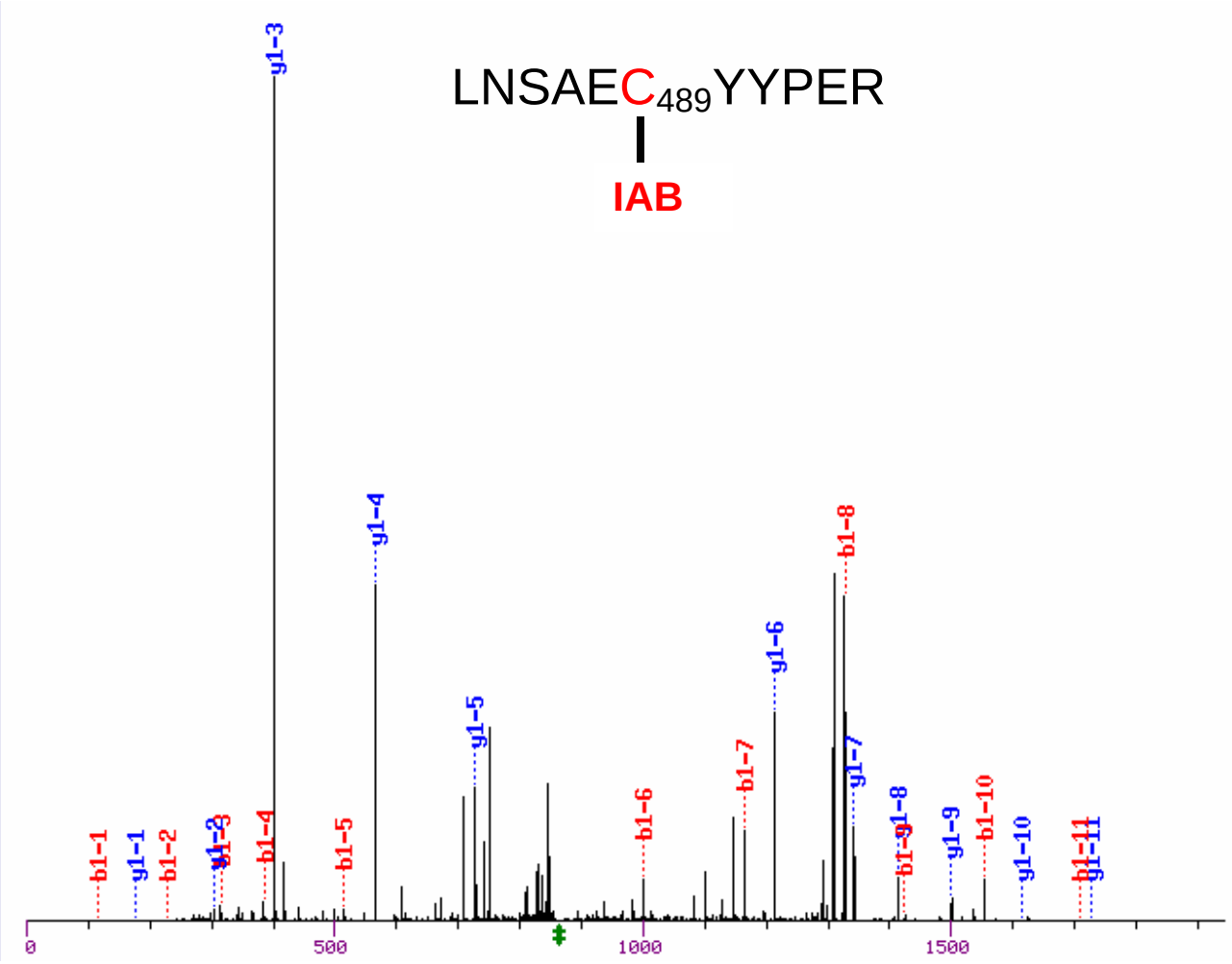
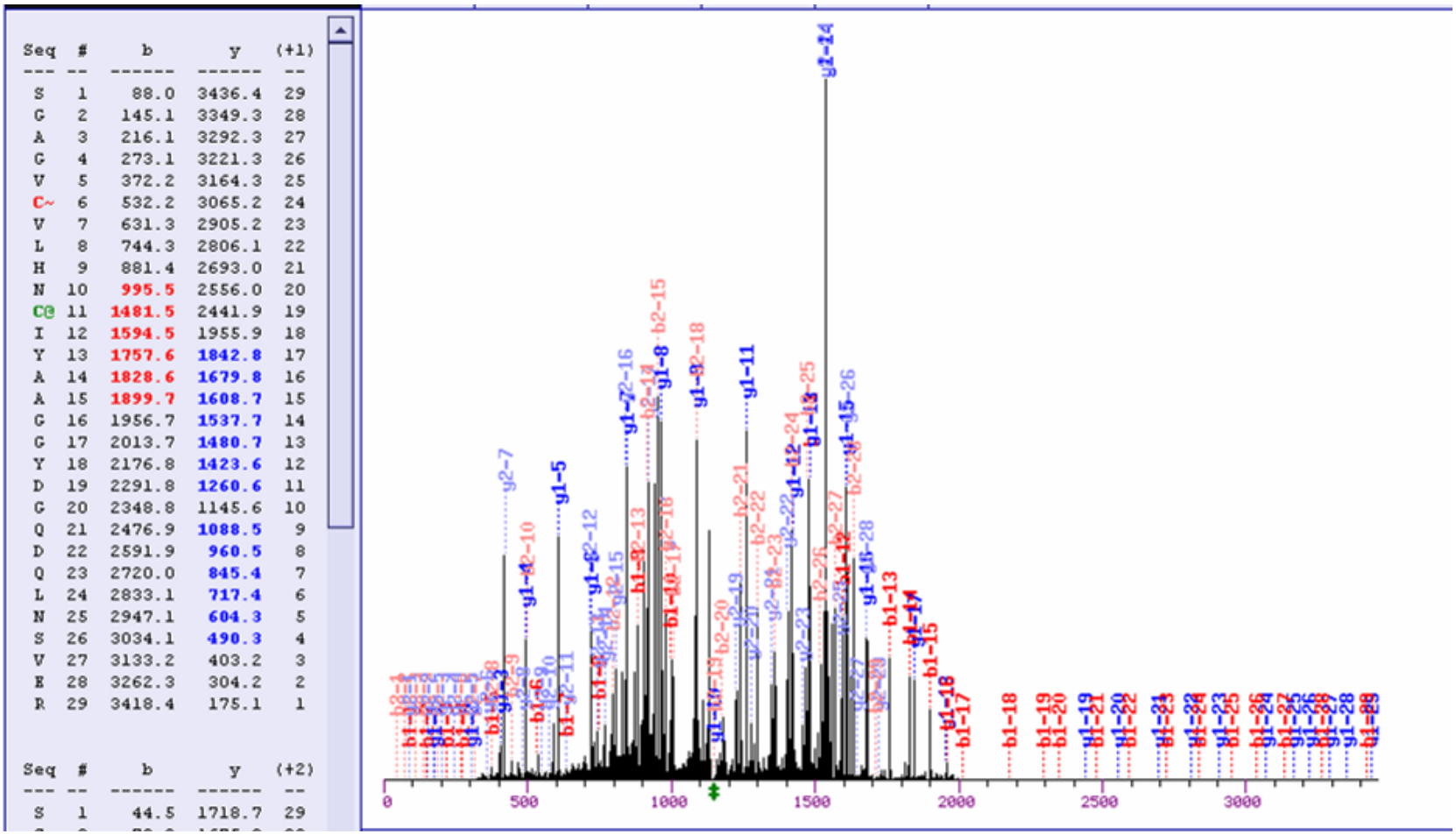


Figure i



Seq	#	b	y	(+1)
S	1	88.0	3436.4	29
G	2	145.1	3349.3	28
A	3	216.1	3292.3	27
G	4	273.1	3221.3	26
V	5	372.2	3164.3	25
C~	6	532.2	3065.2	24
V	7	631.3	2905.2	23
L	8	744.3	2806.1	22
H	9	881.4	2693.0	21
N	10	995.5	2556.0	20
C@	11	1481.5	2441.9	19
I	12	1594.5	1955.9	18
Y	13	1757.6	1842.8	17
A	14	1828.6	1679.8	16
A	15	1899.7	1608.7	15
G	16	1956.7	1537.7	14
G	17	2013.7	1480.7	13
Y	18	2176.8	1423.6	12
D	19	2291.8	1260.6	11
G	20	2348.8	1145.6	10
Q	21	2476.9	1088.5	9
D	22	2591.9	960.5	8
Q	23	2720.0	845.4	7
L	24	2833.1	717.4	6
N	25	2947.1	604.3	5
S	26	3034.1	490.3	4
V	27	3133.2	403.2	3
E	28	3262.3	304.2	2
R	29	3418.4	175.1	1

Seq	#	b	y	(+2)
S	1	44.5	1718.7	29

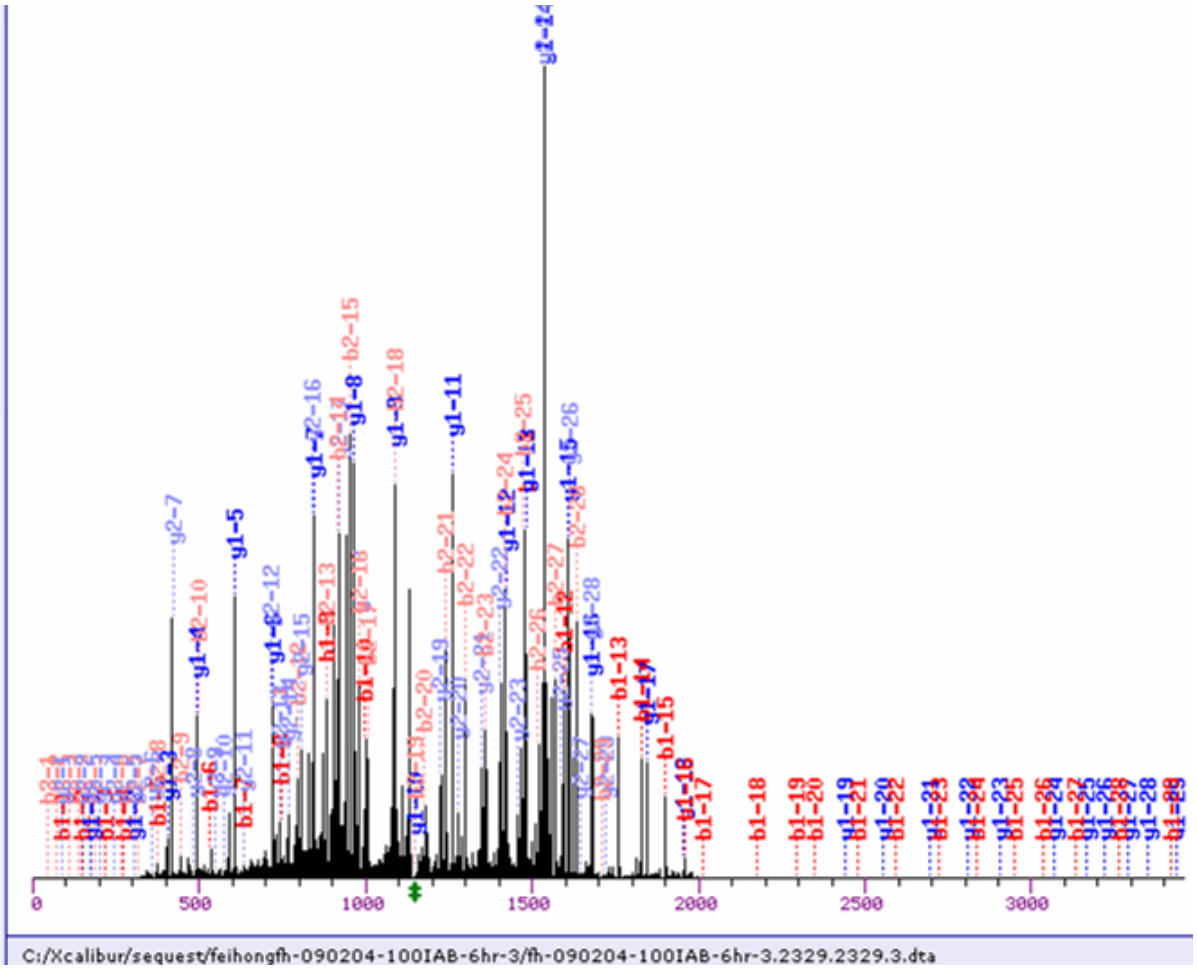
SGAGVC$_{513}$VLHNC$_{518}$IYAAGGYDGQDQLNSVER

IAM **IAB**

Figure j1

S	26	3034.1	490.3	4
V	27	3133.2	403.2	3
E	28	3262.3	304.2	2
R	29	3418.4	175.1	1

Seq	#	b	y	(+2)
S	1	44.5	1718.7	29
G	2	73.0	1675.2	28
A	3	108.6	1646.7	27
G	4	137.1	1611.1	26
V	5	186.6	1582.6	25
C~	6	266.6	1533.1	24
V	7	316.1	1453.1	23
L	8	372.7	1403.6	22
H	9	441.2	1347.0	21
N	10	498.2	1278.5	20
C@	11	741.2	1221.5	19
I	12	797.8	978.5	18
Y	13	879.3	921.9	17
A	14	914.8	840.4	16
A	15	950.3	804.9	15
G	16	978.9	769.3	14
G	17	1007.4	740.8	13
Y	18	1088.9	712.3	12
D	19	1146.4	630.8	11
G	20	1174.9	573.3	10
Q	21	1239.0	544.8	9
D	22	1296.5	480.7	8
Q	23	1360.5	423.2	7
L	24	1417.0	359.2	6
N	25	1474.1	302.7	5
S	26	1517.6	245.6	4
V	27	1567.1	202.1	3
E	28	1631.6	152.6	2
R	29	1709.7	88.1	1



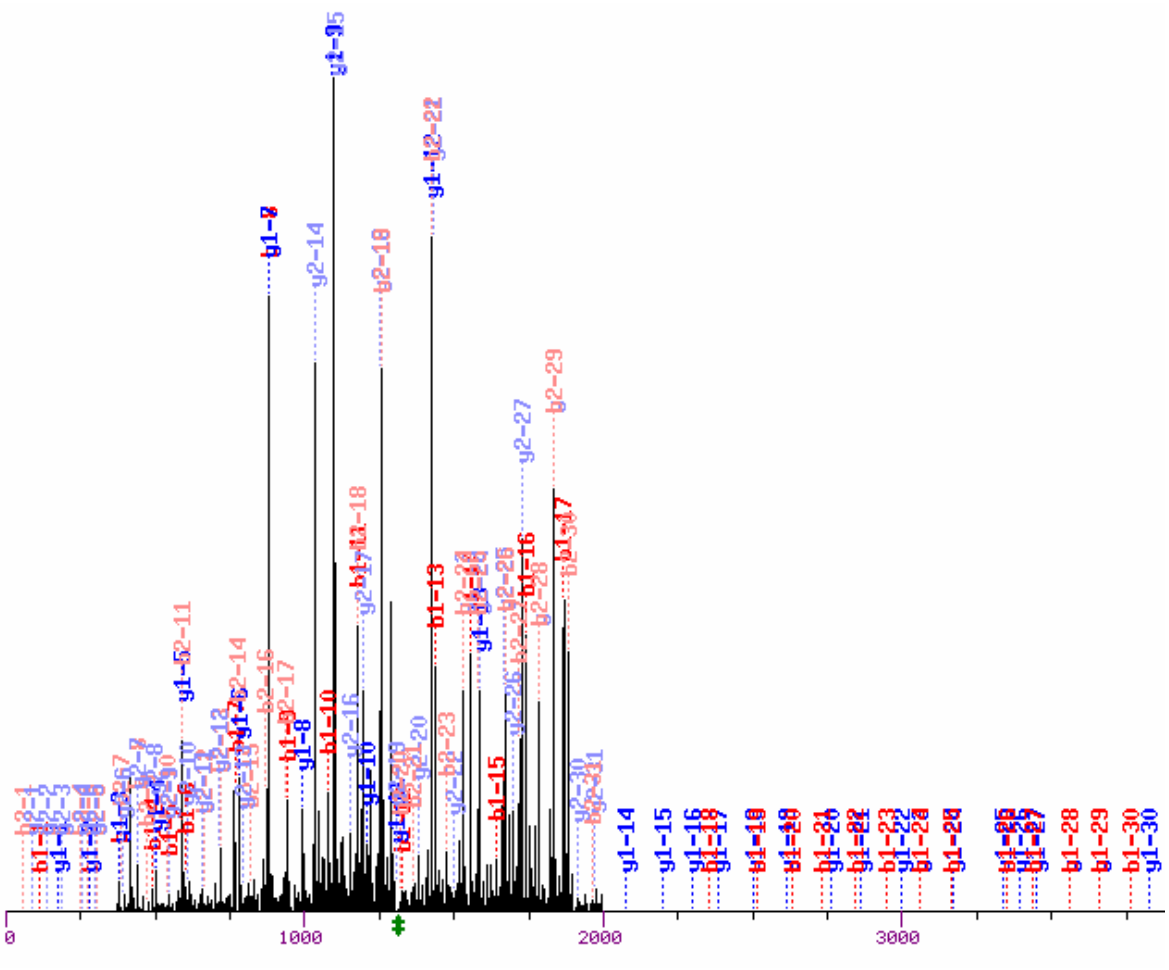
SGAGVC$_{513}$VLHNC$_{518}$IYAAGGYDGQDQLNSVER

IAM IAB

Figure j2

Seq	#	b	y	(+1)
I	1	114.1	3937.6	31
Y	2	277.2	3824.5	30
V	3	376.2	3661.4	29
L	4	489.3	3562.4	28
G	5	546.3	3449.3	27
G	6	603.4	3392.3	26
Y	7	**766.4**	3335.2	25
D	8	**881.4**	3172.2	24
G	9	938.5	3057.1	23
H	10	1075.5	3000.1	22
T	11	**1176.6**	2863.1	21
F	12	1323.6	2762.0	20
L	13	**1436.7**	2614.9	19
D	14	**1551.7**	2501.9	18
S	15	1638.8	2386.8	17
V	16	**1737.8**	2299.8	16
E	17	**1866.9**	2200.7	15
C^	18	2353.9	2071.7	14
Y	19	2517.0	**1584.7**	13
D	20	2632.0	**1421.6**	12
P	21	2729.0	1306.6	11
D	22	2844.1	1209.5	10
T	23	2945.1	**1094.5**	9
D	24	3060.1	**993.5**	8
T	25	3161.2	**878.4**	7
W	26	3347.3	**777.4**	6
S	27	3434.3	**591.3**	5
E	28	3563.3	504.3	4
V	29	3662.4	375.2	3
T	30	3763.5	276.2	2
R	31	3919.6	175.1	1

Seq # b y (+2)



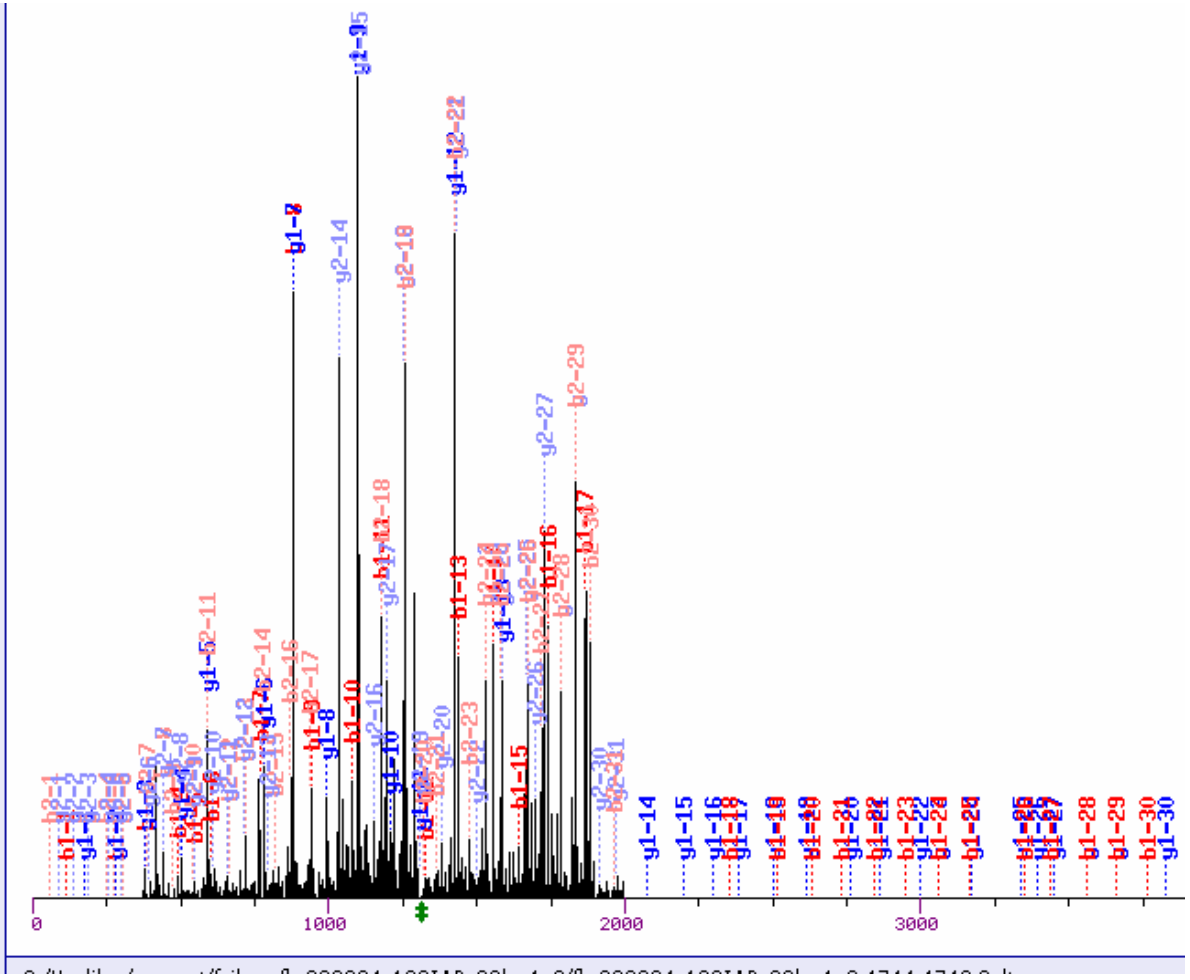
IYVLGGYDGHTFLDSVEC$_{583}$YDPDTDTWSEVTR

IAB

Figure k1

V	29	3662.4	375.2	3
T	30	3763.5	276.2	2
R	31	3919.6	175.1	1

Seq	#	b	y	(+2)
I	1	57.5	1969.3	31
Y	2	139.1	1912.7	30
V	3	188.6	**1831.2**	29
L	4	245.2	**1781.7**	28
G	5	273.7	**1725.1**	27
G	6	302.2	**1696.6**	26
Y	7	383.7	**1668.1**	25
D	8	441.2	**1586.6**	24
G	9	469.7	**1529.1**	23
H	10	538.3	1500.6	22
T	11	588.8	1432.0	21
F	12	662.3	**1381.5**	20
L	13	718.9	1308.0	19
D	14	776.4	**1251.4**	18
S	15	819.9	**1193.9**	17
V	16	869.4	**1150.4**	16
E	17	933.9	**1100.9**	15
C^	18	**1177.5**	**1036.3**	14
Y	19	**1259.0**	792.8	13
D	20	1316.5	711.3	12
P	21	1365.0	653.8	11
D	22	**1422.5**	605.3	10
T	23	**1473.1**	547.8	9
D	24	**1530.6**	497.2	8
T	25	**1581.1**	439.7	7
W	26	**1674.1**	389.2	6
S	27	**1717.7**	296.2	5
E	28	**1782.2**	252.6	4
V	29	**1831.7**	188.1	3
T	30	**1882.2**	138.6	2
R	31	1960.3	88.1	1



IYVLGGYDGHTFLDSVEC$_{583}$YDPDTDTWSEVTR

IAB

Figure k2

Seq	#	b	y	(+1)
K	1	129.1	1619.5	10
Q	2	257.2	1491.4	9
I	3	370.2	1363.4	8
D	4	485.3	1250.3	7
Q	5	613.3	1135.2	6
Q	6	741.4	1007.2	5
N	7	855.4	879.1	4
C#	8	1340.4	765.1	3
T	9	1441.5	280.1	2
C~	10	1601.5	179.0	1

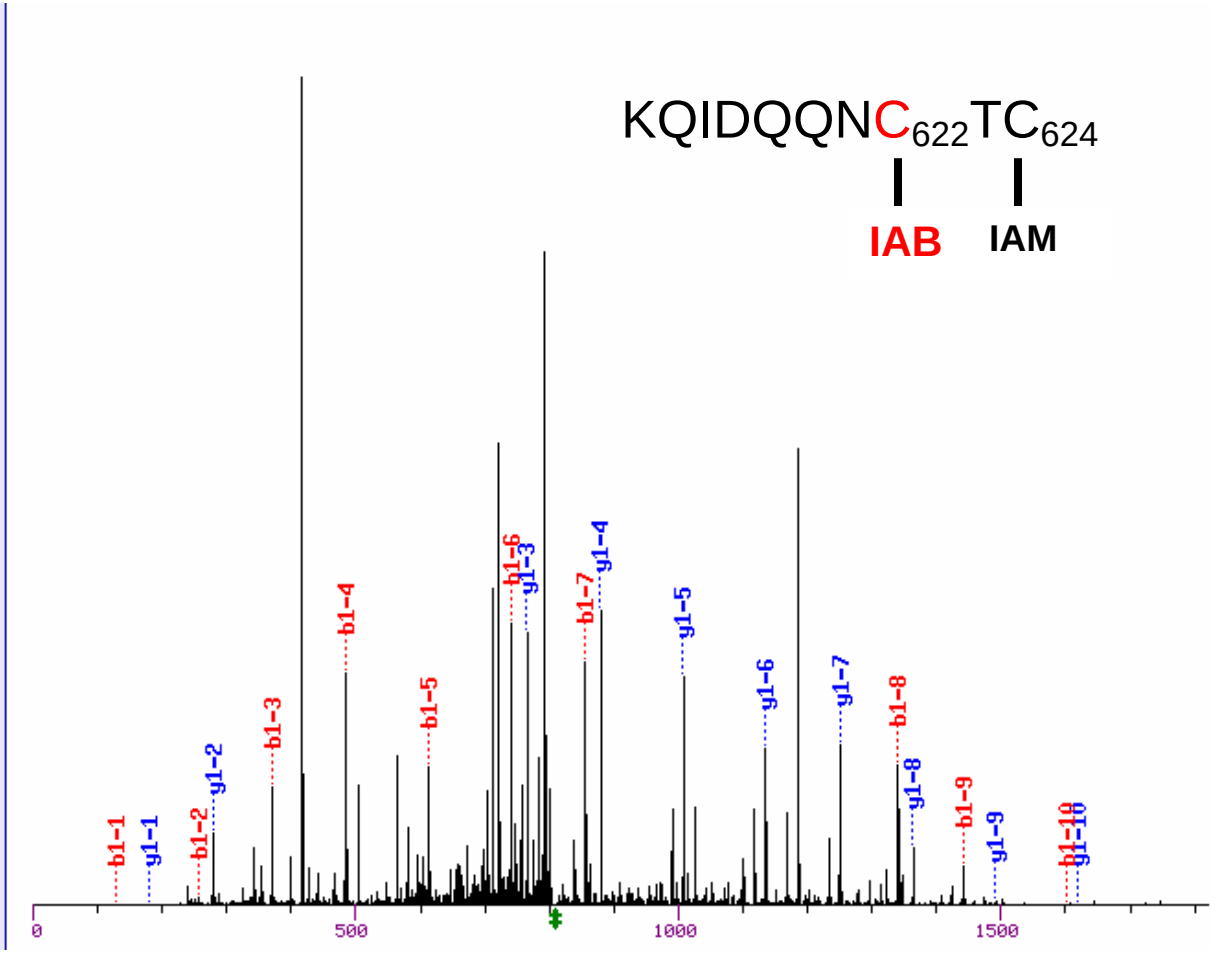


Figure I

Seq	#	b	y	(+1)
L	1	114.1	2009.3	13
S	2	201.1	1896.2	12
Q	3	329.2	1809.2	11
Q	4	457.2	**1681.1**	10
L	5	**570.3**	**1553.1**	9
C#	6	**1207.8**	**1440.0**	8
D	7	**1322.9**	**802.5**	7
V	8	**1421.9**	**687.4**	6
T	9	**1523.0**	**588.4**	5
L	10	**1636.1**	**487.3**	4
Q	11	**1764.1**	**374.2**	3
V	12	**1863.2**	246.2	2
K	13	1991.3	147.1	1

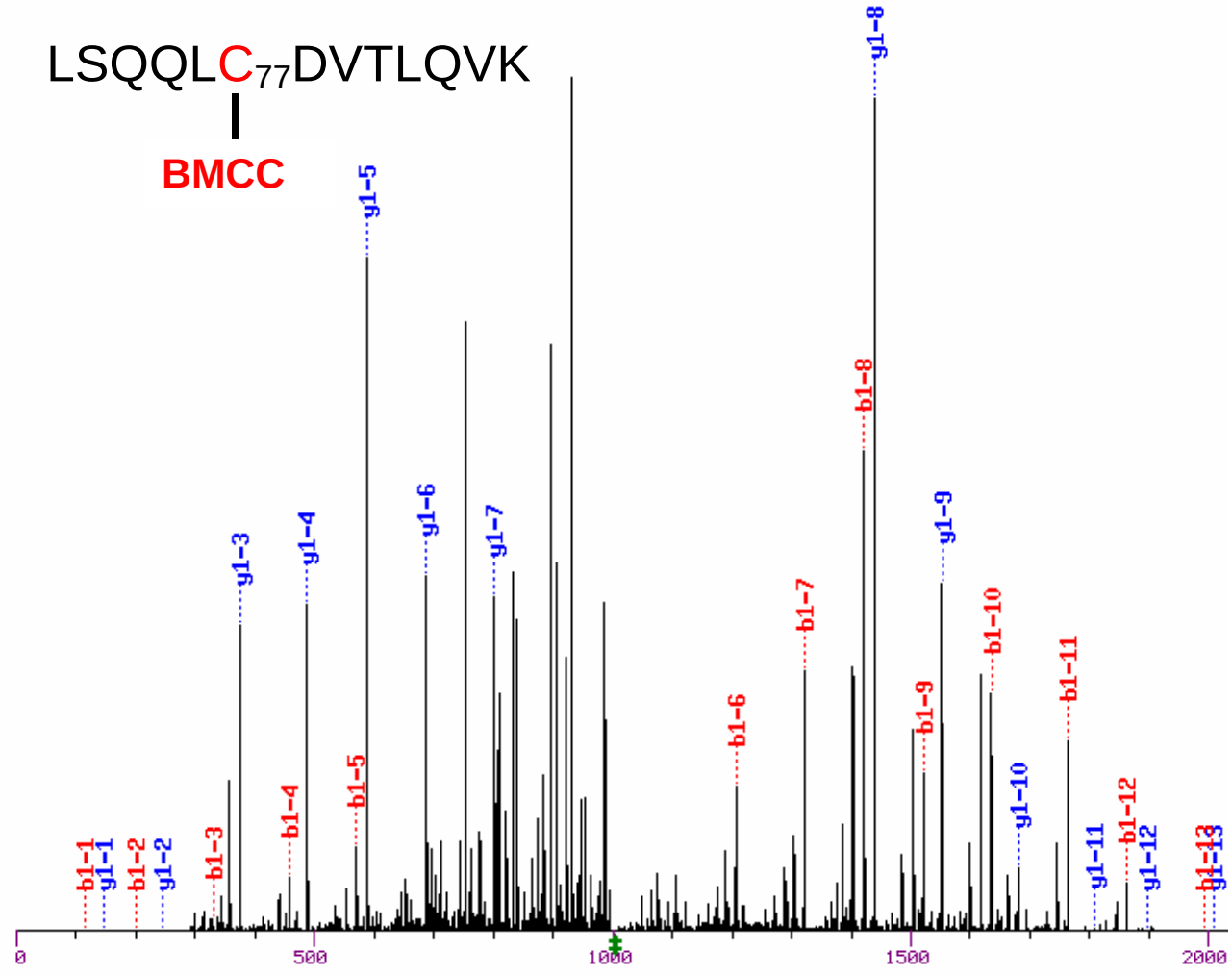


Figure m

Seq	#	b	y	(+1)
A	1	72.0	4177.2	33
C@	2	232.1	4106.2	32
S	3	319.1	3946.2	31
D	4	434.1	3859.2	30
F	5	581.2	3744.1	29
L	6	**694.3**	3597.1	28
V	7	**793.3**	3484.0	27
Q	8	**921.4**	3384.9	26
Q	9	**1049.5**	3256.9	25
L	10	**1162.5**	3128.8	24
D	11	**1277.6**	3015.7	23
P	12	**1374.6**	2900.7	22
S	13	1461.6	2803.6	21
N	14	**1575.7**	2716.6	20
A	15	**1646.7**	2602.6	19
I	16	**1759.8**	2531.5	18
G	17	1816.8	2418.4	17
I	18	**1929.9**	2361.4	16
A	19	2001.0	2248.3	15
N	20	2115.0	2177.3	14
F	21	2262.1	2063.3	13
A	22	2333.1	1916.2	12
E	23	2462.1	1845.1	11
Q	24	2590.2	1716.1	10
I	25	2703.3	1588.0	9
G	26	2760.3	**1475.0**	8
C*	27	3396.8	1417.9	7
V	28	3495.9	781.4	6
E	29	3624.9	**682.4**	5
L	30	3738.0	553.3	4
H	31	3875.1	440.2	3
Q	32	4003.1	303.2	2
R	33	4159.2	175.1	1

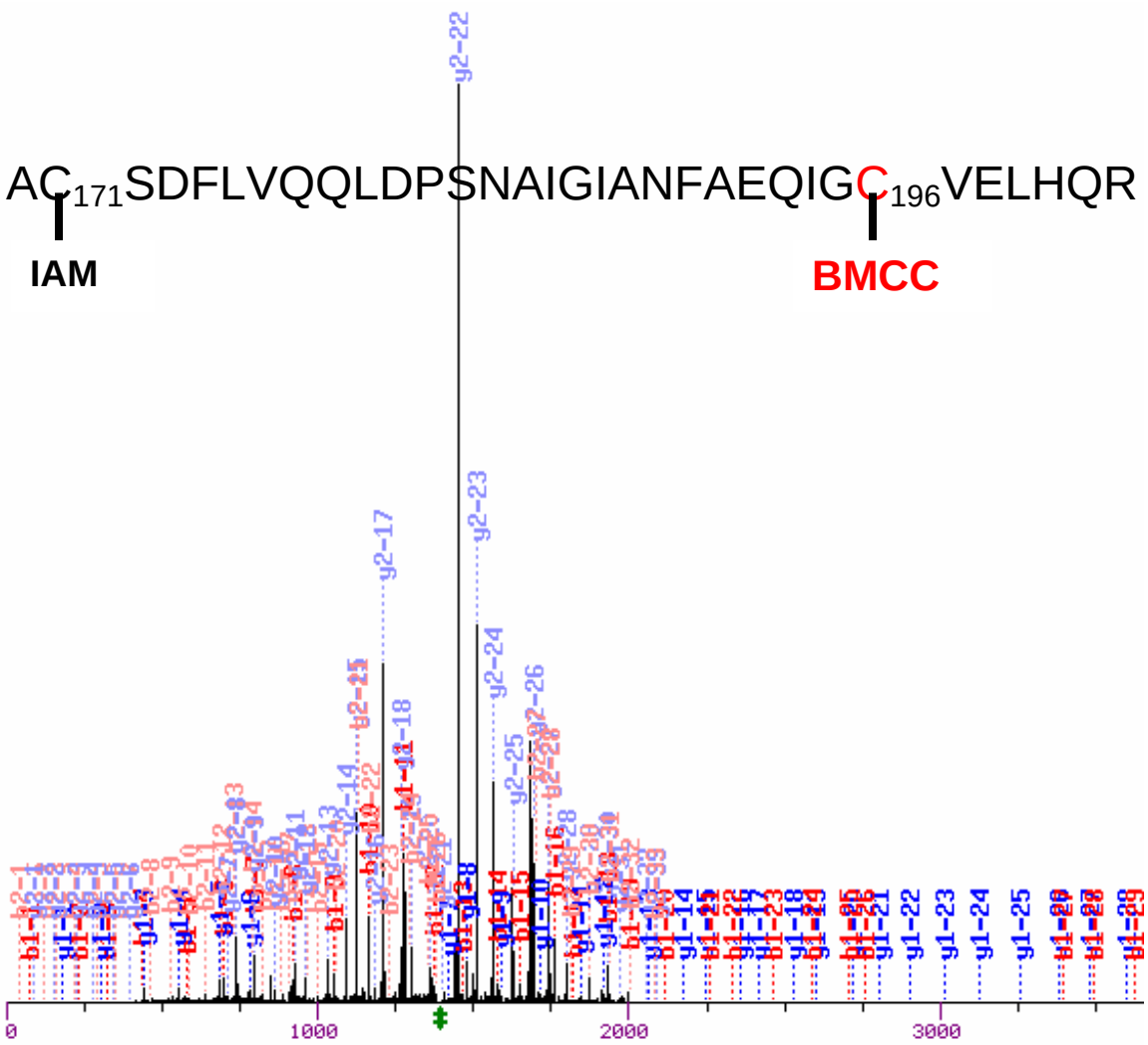


Figure n1

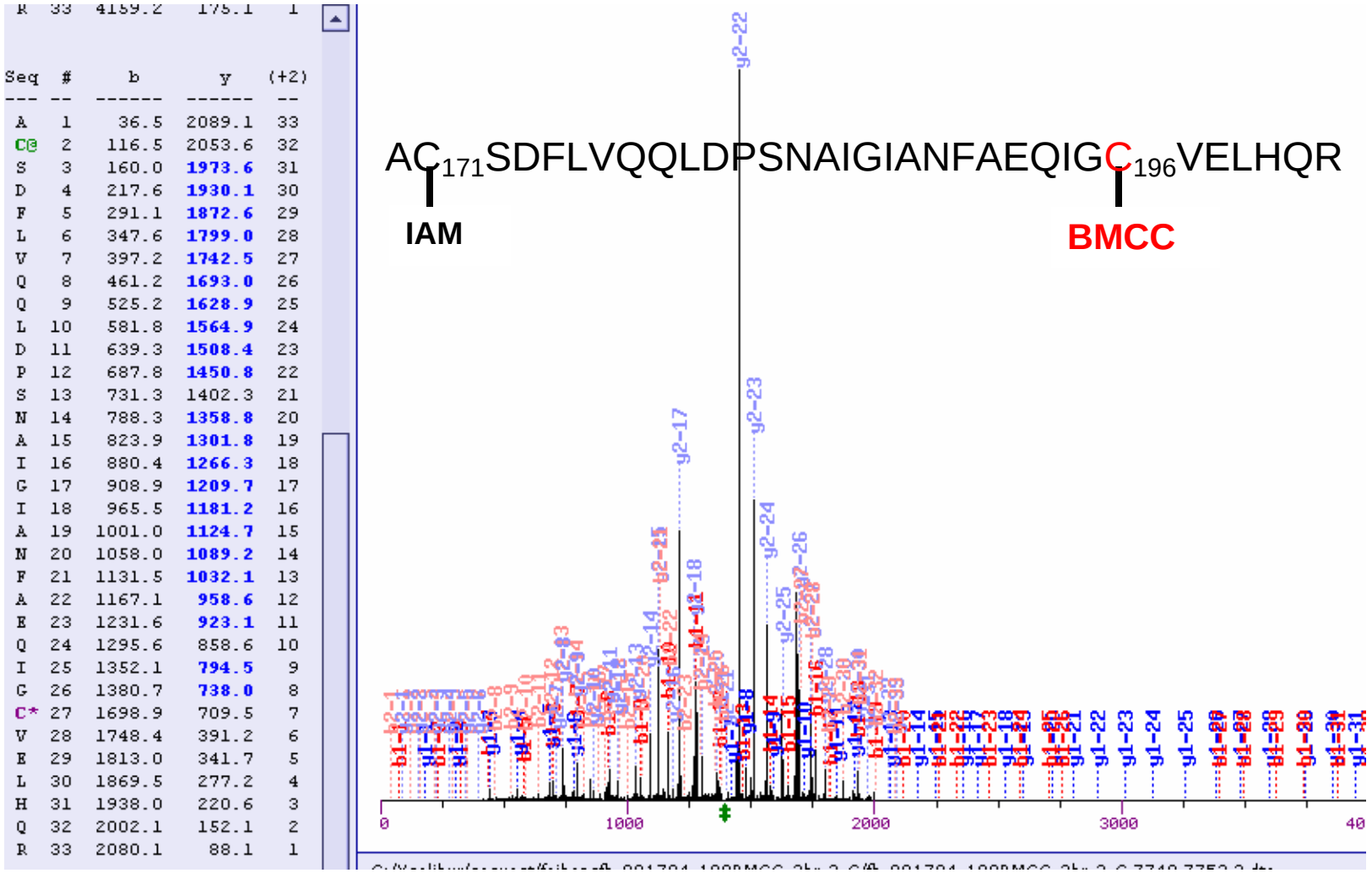

R 33 4189.2 175.1 1
Seq # b y (+2)
A 1 36.5 2089.1 33
C@ 2 116.5 2053.6 32
S 3 160.0 1973.6 31
D 4 217.6 1930.1 30
F 5 291.1 1872.6 29
L 6 347.6 1799.0 28
V 7 397.2 1742.5 27
Q 8 461.2 1693.0 26
Q 9 525.2 1628.9 25
L 10 581.8 1564.9 24
D 11 639.3 1508.4 23
P 12 687.8 1450.8 22
S 13 731.3 1402.3 21
N 14 788.3 1358.8 20
A 15 823.9 1301.8 19
I 16 880.4 1266.3 18
G 17 908.9 1209.7 17
I 18 965.5 1181.2 16
A 19 1001.0 1124.7 15
N 20 1058.0 1089.2 14
F 21 1131.5 1032.1 13
A 22 1167.1 958.6 12
E 23 1231.6 923.1 11
Q 24 1295.6 858.6 10
I 25 1352.1 794.5 9
G 26 1380.7 738.0 8
C* 27 1698.9 709.5 7
V 28 1748.4 391.2 6
E 29 1813.0 341.7 5
L 30 1869.5 277.2 4
H 31 1938.0 220.6 3
Q 32 2002.1 152.1 2
R 33 2080.1 88.1 1
AC171SDFLVQQLDPSNAIGIANFAEQIGC196VELHQR
IAM
BMCC
0
1000
2000
3000

Figure n2

Seq	#	b	y	(+1)
C@	1	161.0	2256.3	14
E	2	**290.1**	2096.2	13
S	3	**377.1**	1967.2	12
E	4	**506.1**	1880.2	11
V	5	**605.2**	1751.1	10
F	6	752.3	1652.1	9
H	7	889.3	**1505.0**	8
A	8	960.4	**1367.9**	7
C#	9	1597.9	**1296.9**	6
I	10	1711.0	659.4	5
N	11	1825.0	**546.3**	4
W	12	2011.1	**432.3**	3
V	13	2110.2	**246.2**	2
K	14	2238.2	147.1	1

Seq	#	b	y	(+2)
C@	1	81.0	1128.6	14
E	2	145.5	**1048.6**	13
S	3	189.0	**984.1**	12
E	4	253.6	**940.6**	11
V	5	303.1	**876.1**	10
F	6	**376.6**	**826.5**	9
H	7	445.2	753.0	8
A	8	480.7	**684.5**	7
C#	9	799.4	**649.0**	6
I	10	856.0	330.2	5
N	11	**913.0**	**273.7**	4
W	12	**1006.0**	216.6	3
V	13	**1055.6**	123.6	2
K	14	1119.6	74.1	1

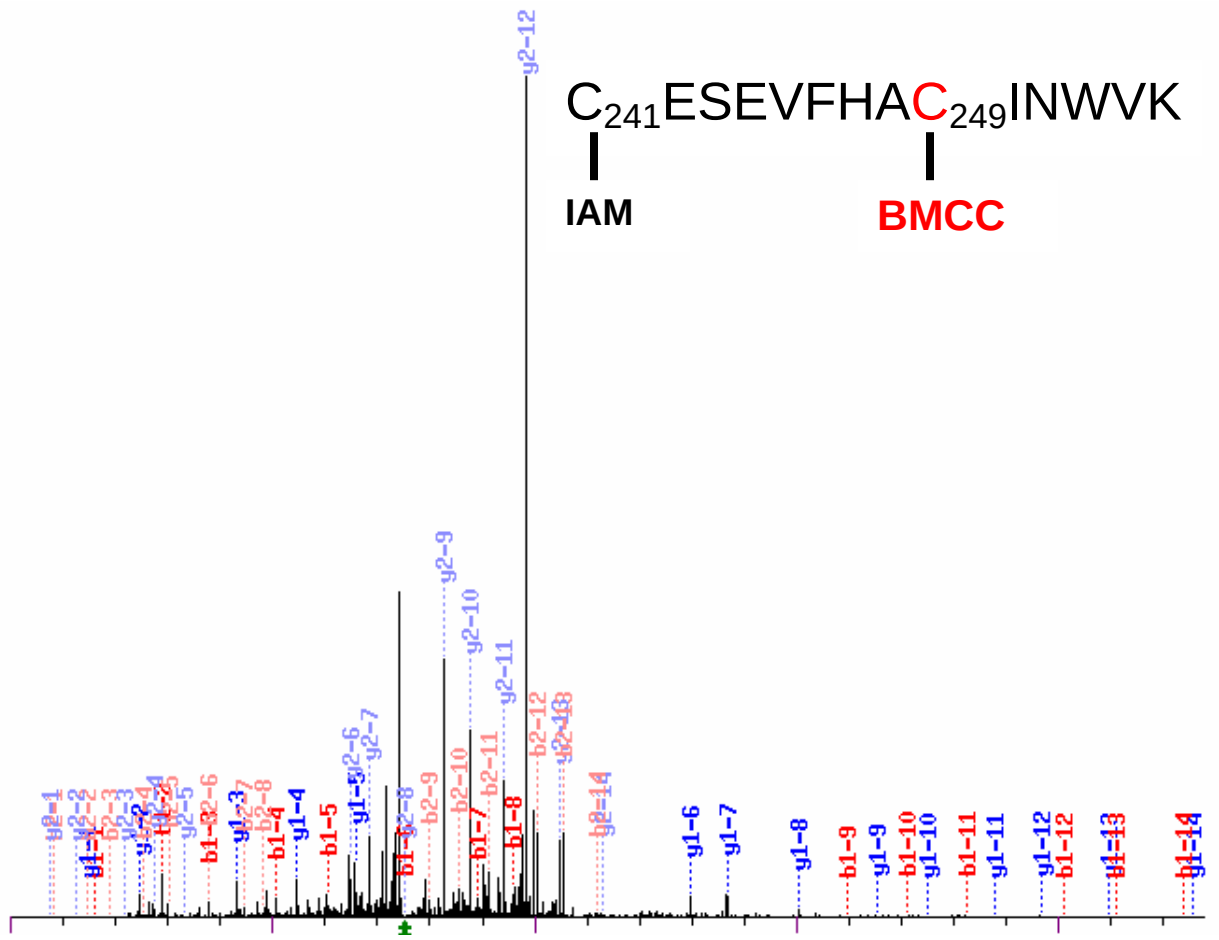


Figure o

S	1	88.0	2182.4	18
G	2	145.1	2095.4	17
L	3	258.1	2038.3	16
A	4	329.2	1925.2	15
G	5	386.2	1854.2	14
C*	6	1022.7	1797.2	13
V	7	1121.8	1160.7	12
V	8	1220.8	1061.6	11
G	9	1277.9	962.5	10
G	10	1334.9	905.5	9
L	11	1448.0	848.5	8
L	12	1561.1	735.4	7
Y	13	1724.1	622.3	6
A	14	1795.2	459.3	5
V	15	1894.2	388.2	4
G	16	1951.3	289.2	3
G	17	2008.3	232.1	2
R	18	2164.4	175.1	1

Seq	#	b	y	(+2)
S	1	44.5	1091.7	18
G	2	73.0	1048.2	17
L	3	129.6	1019.7	16
A	4	165.1	963.1	15
G	5	193.6	927.6	14
C*	6	511.9	899.1	13
V	7	561.4	580.8	12
V	8	610.9	531.3	11
G	9	639.4	481.8	10
G	10	668.0	453.3	9
L	11	724.5	424.8	8
L	12	781.0	368.2	7
Y	13	862.6	311.7	6
A	14	898.1	230.1	5
V	15	947.6	194.6	4
G	16	976.1	145.1	3
G	17	1004.6	116.6	2
R	18	1082.7	88.1	1

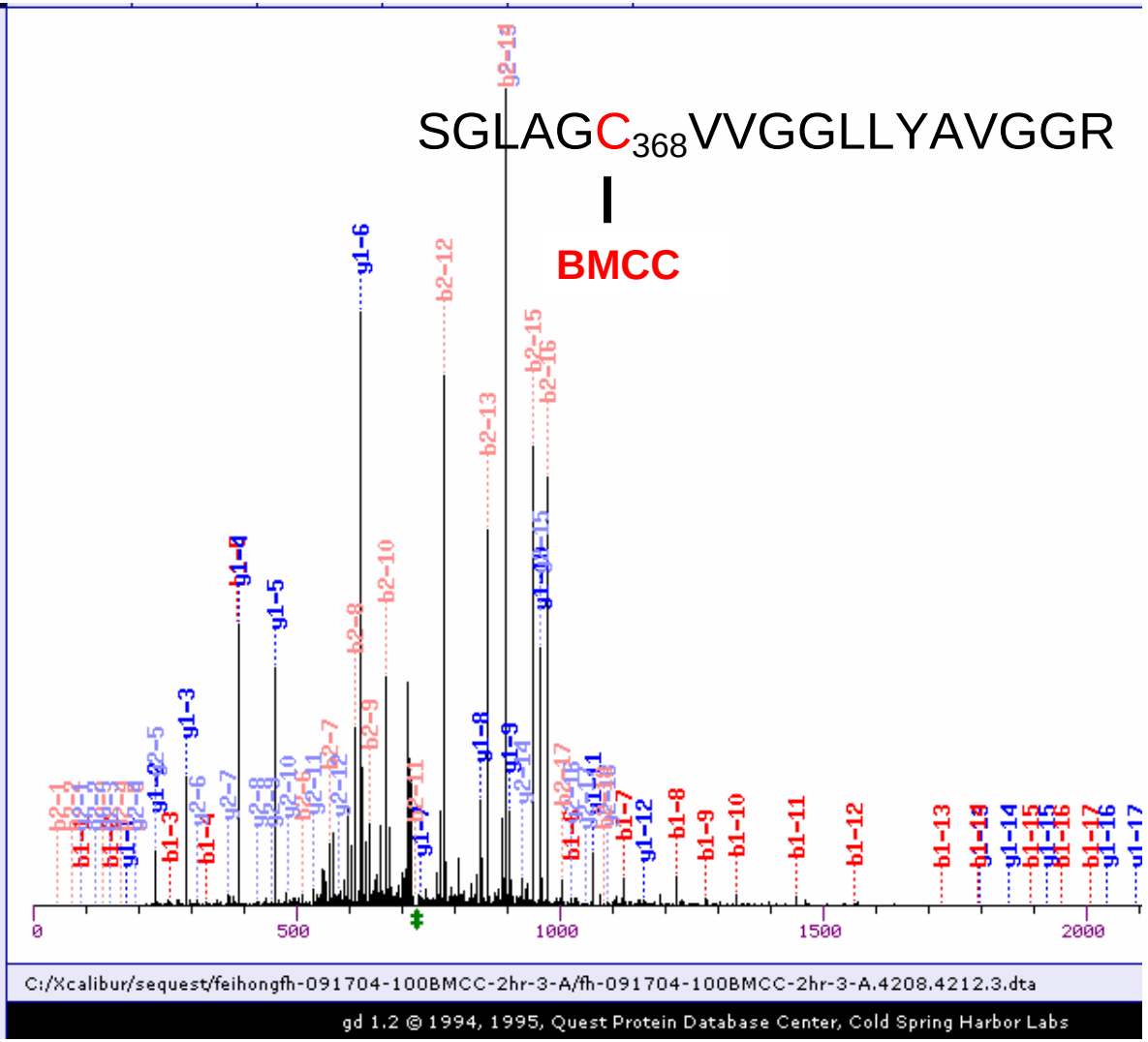


Figure p

Seq	#	b	y	(+1)
L	1	114.1	1878.1	11
N	2	228.1	1765.0	10
S	3	315.2	1651.0	9
A	4	386.2	1563.9	8
E	5	515.2	1492.9	7
C*	6	1151.8	1363.9	6
Y	7	1314.8	727.3	5
Y	8	1477.9	564.3	4
P	9	1574.9	401.2	3
E	10	1704.0	304.2	2
R	11	1860.1	175.1	1

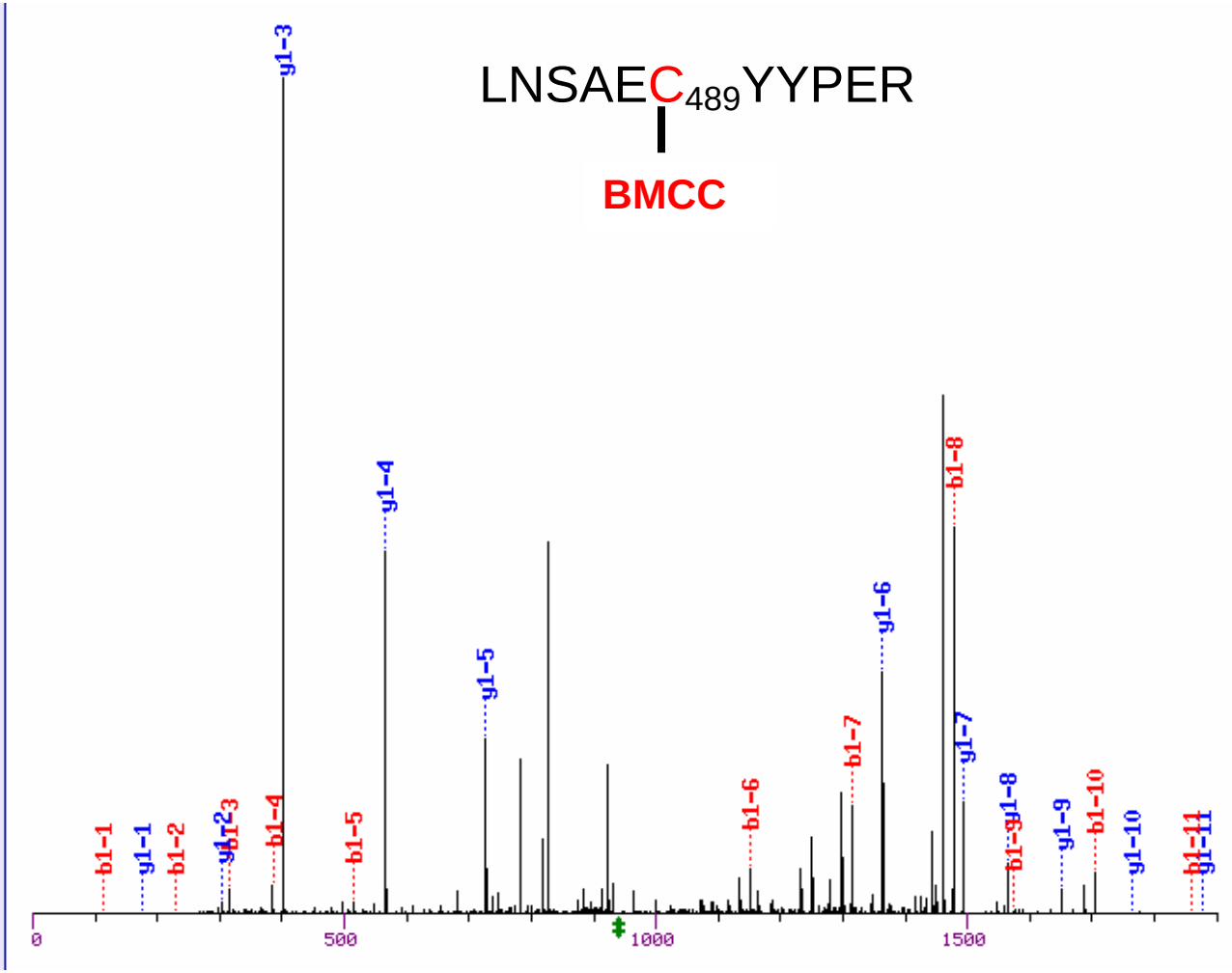
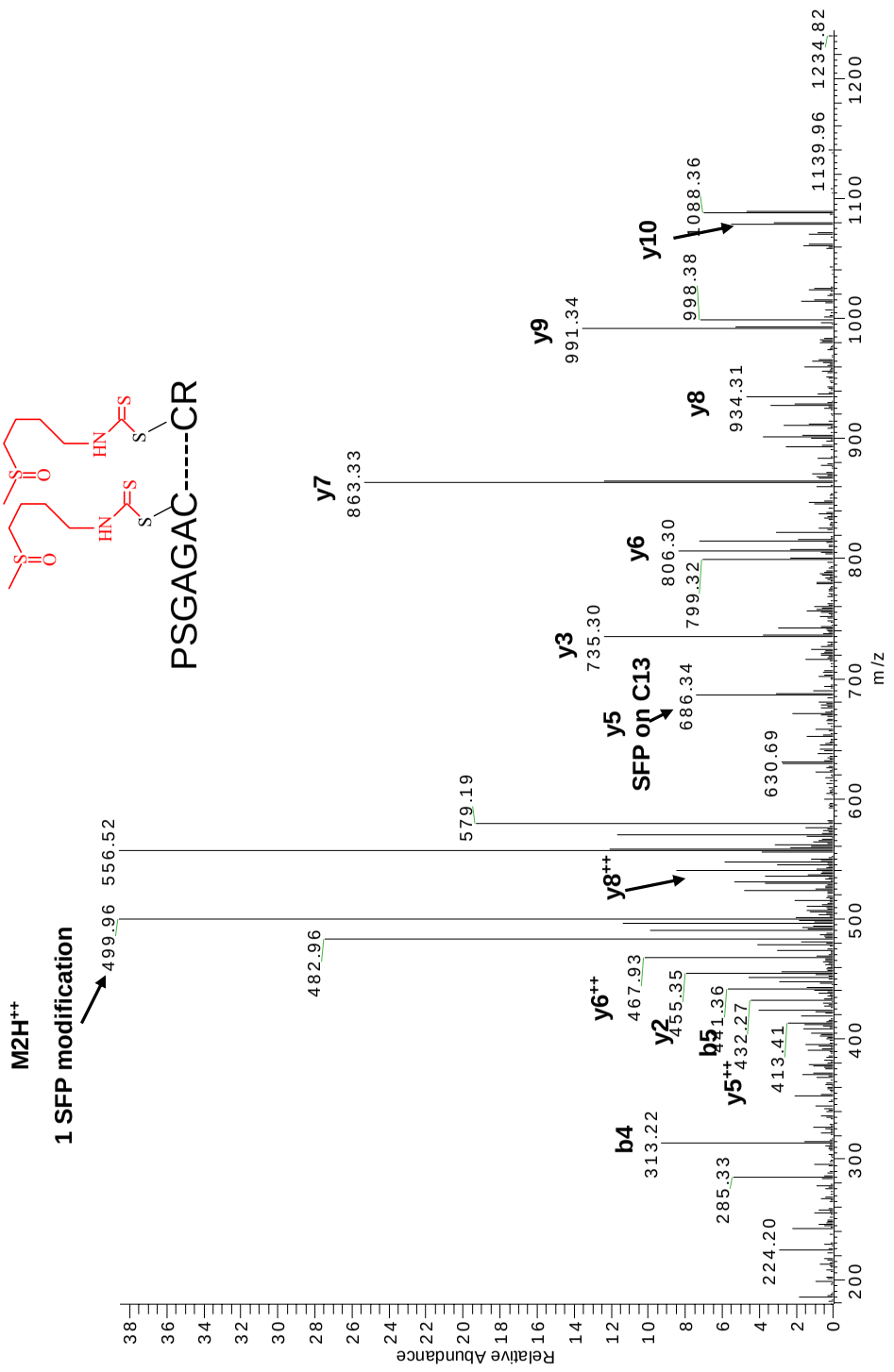
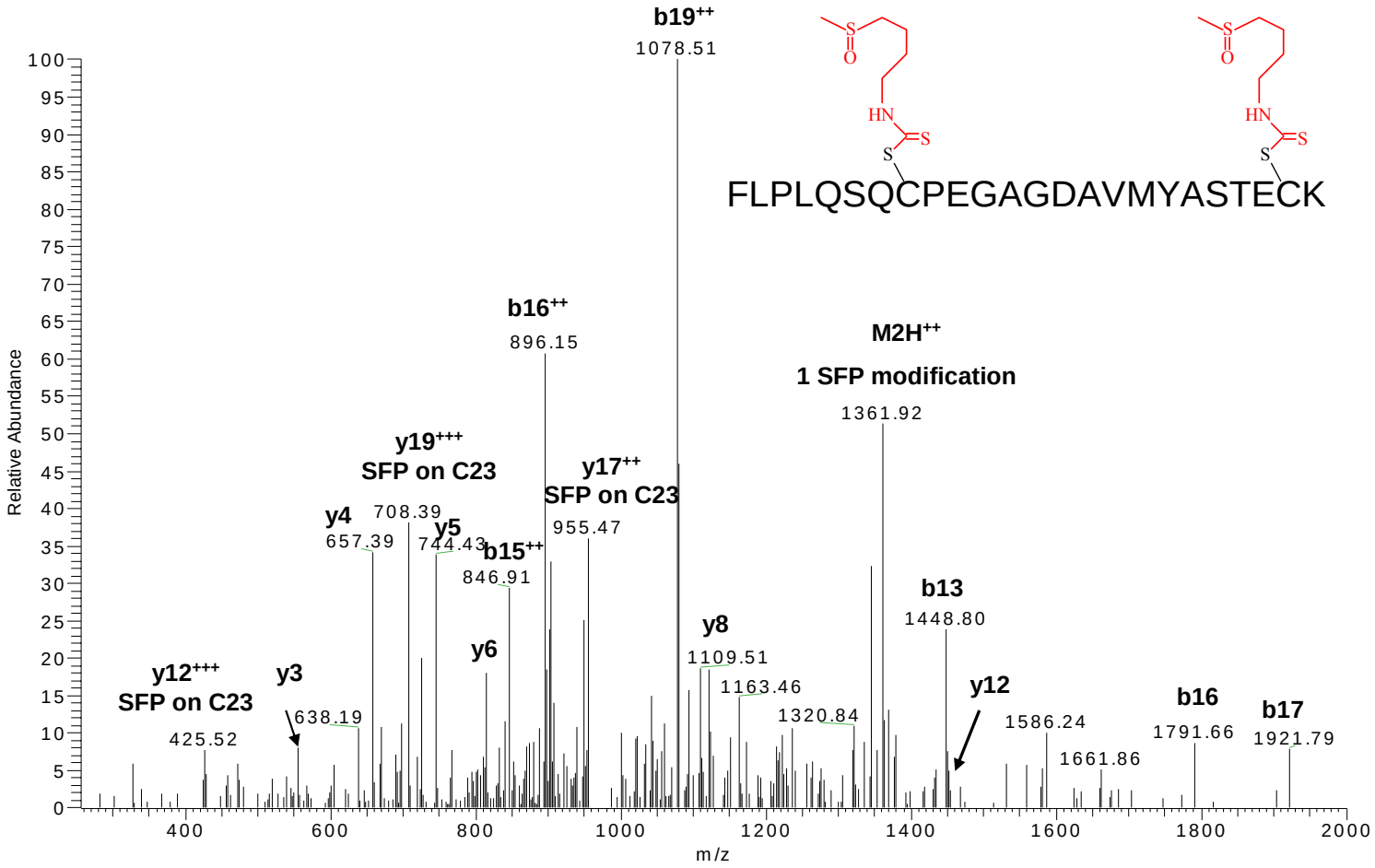
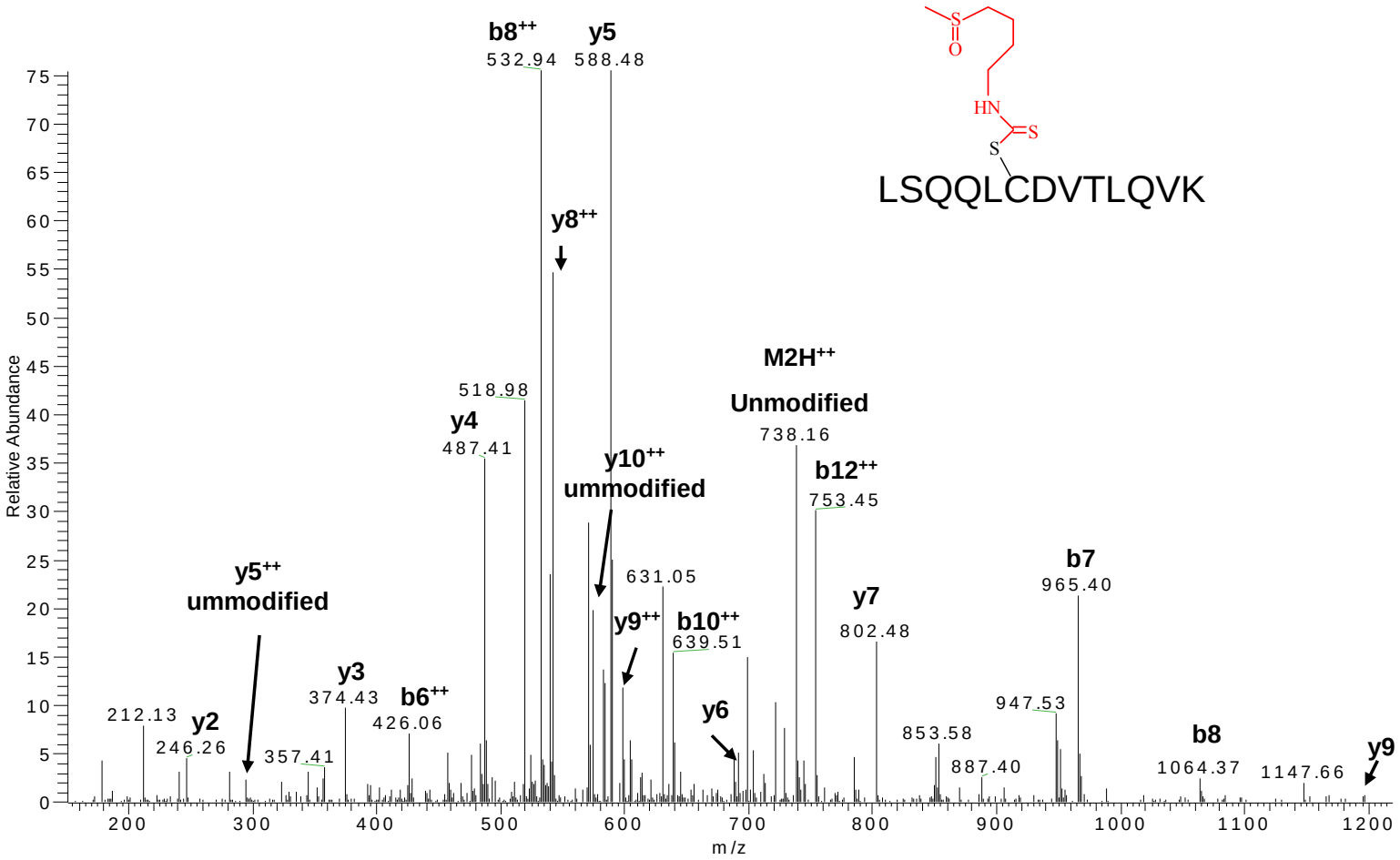


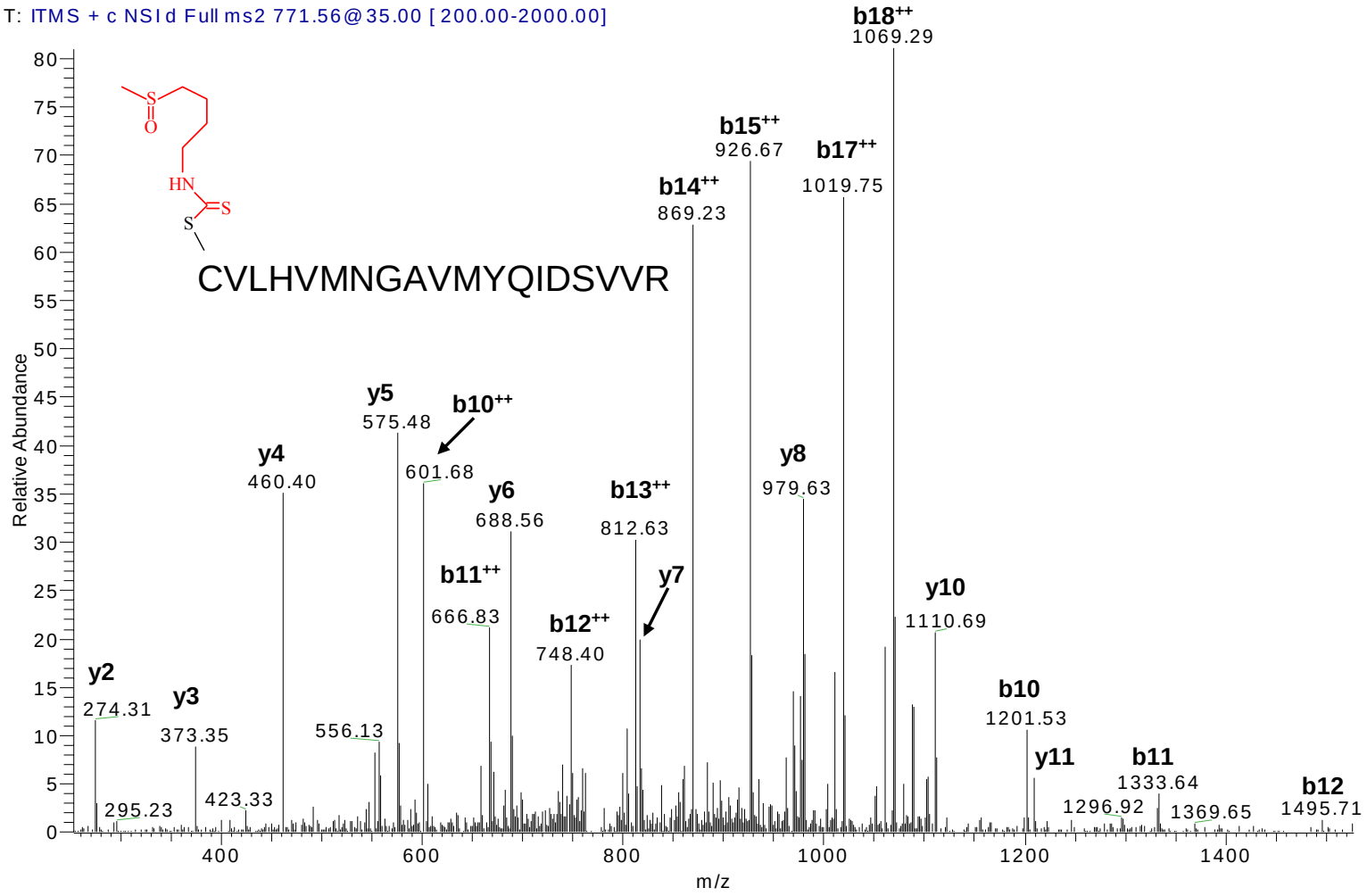
Figure q

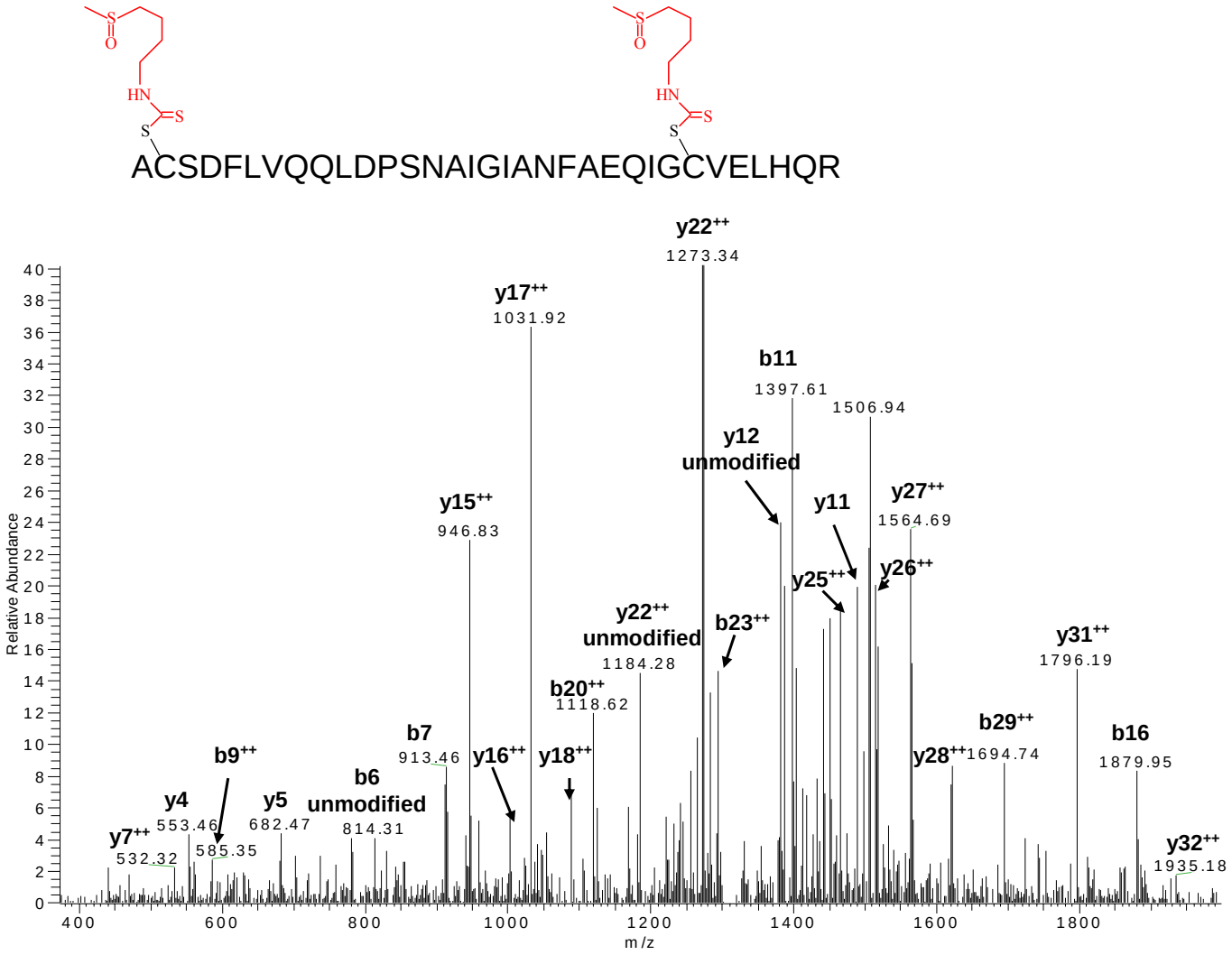
APPENDIX 2

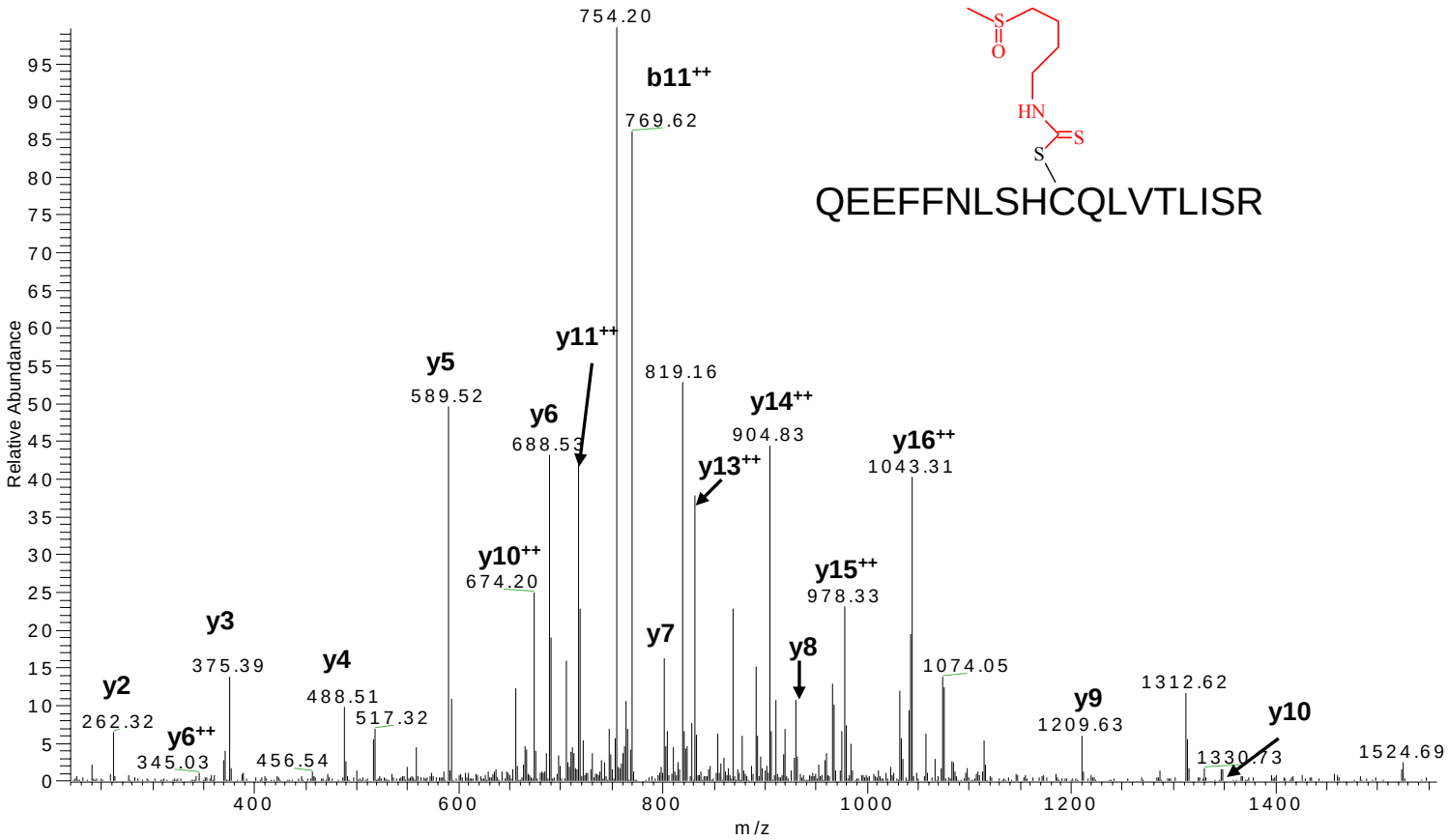

PSGAGAC----CR
M2H++
1 SFP modification
SFP on C13
y10
y9
y8
y7
y6
y5
y3
y8++
y6++
y2
b5
y5++
b4
Relative Abundance
m/z

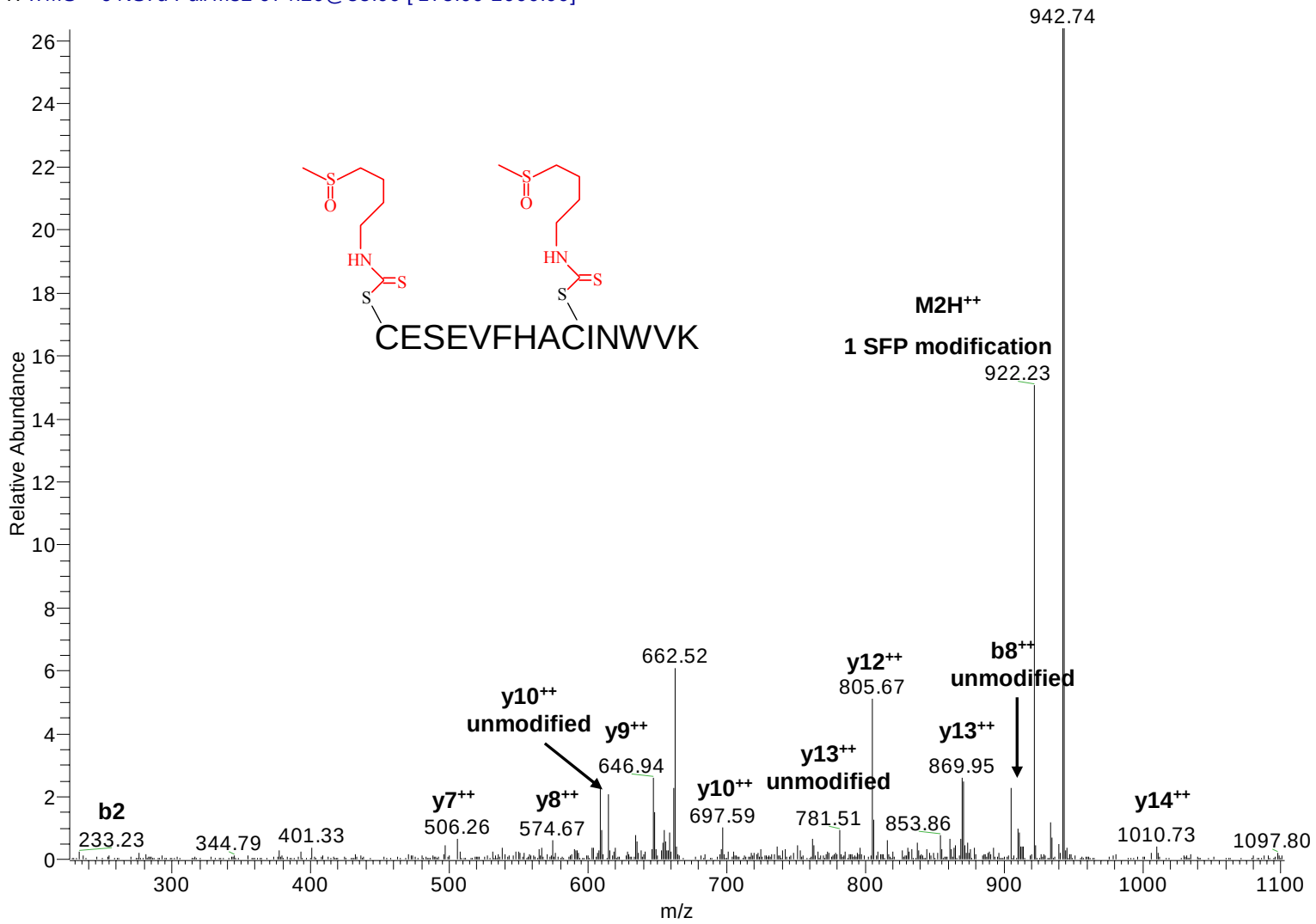

b19++
1078.51
FLPLQSQCPEGAGDAVMYASTECK
b16++
896.15
M2H++
1 SFP modification
1361.92
y19+++
SFP on C23
y17++
SFP on C23
955.47
y4
657.39
708.39
y5
744.43
b15++
846.91
b13
1448.80
y8
1109.51
y6
y3
638.19
y12+++
SFP on C23
425.52
1163.46
1320.84
y12
1586.24
b16
1791.66
b17
1921.79
1661.86
Relative Abundance
m/z

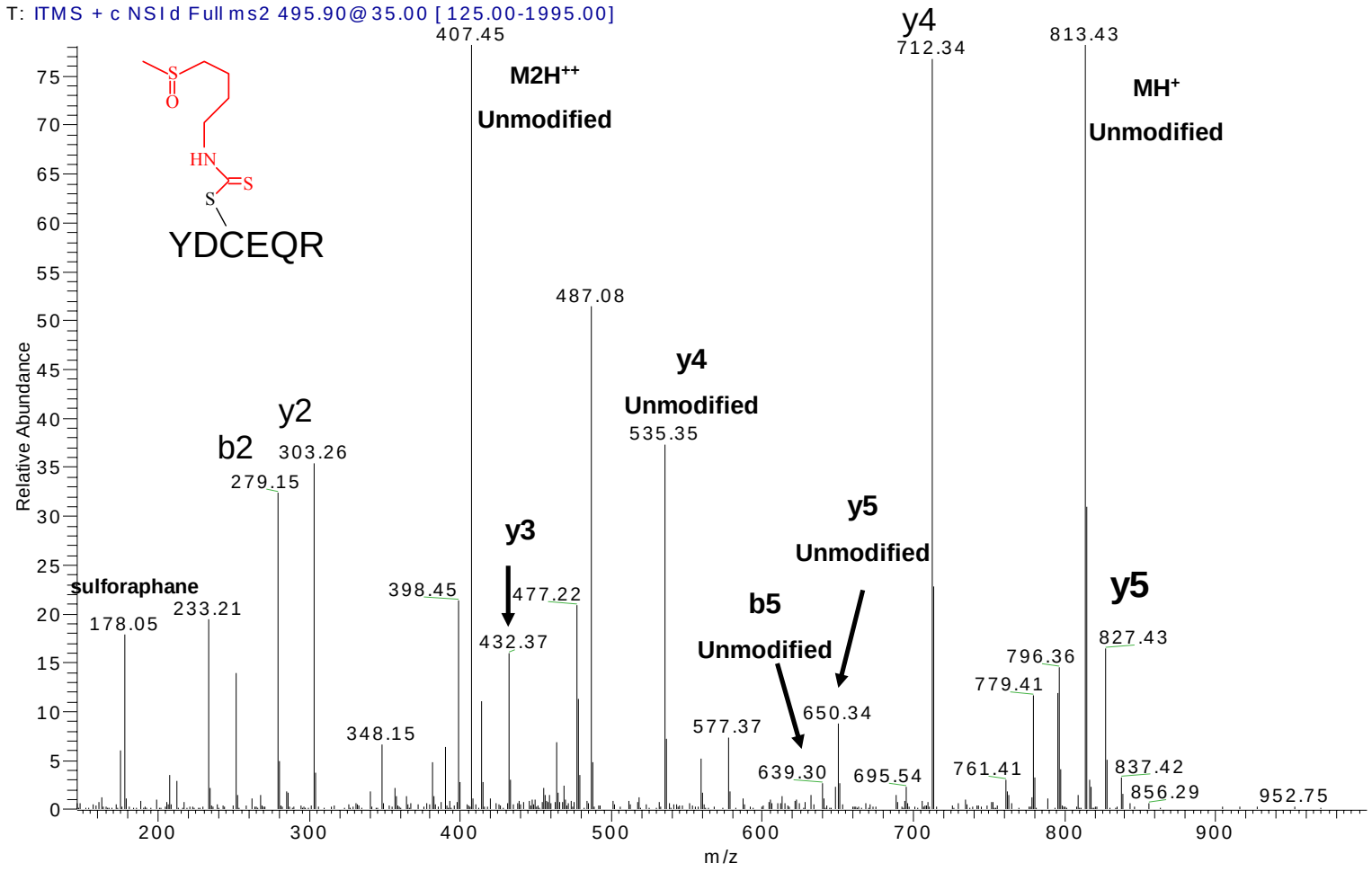

b8++
532.94
y5
588.48
y8++
518.98
y4
487.41
M2H++
Unmodified
738.16
b12++
753.45
y10++
ummodified
631.05
y9++
b10++
639.51
y7
802.48
b7
965.40
y5++
ummodified
y3
374.43
b6++
426.06
212.13
y2
246.26
357.41
y6
947.53
853.58
887.40
b8
1064.37
1147.66
y9
LSQQLCDVTLQVK
Relative Abundance
m/z

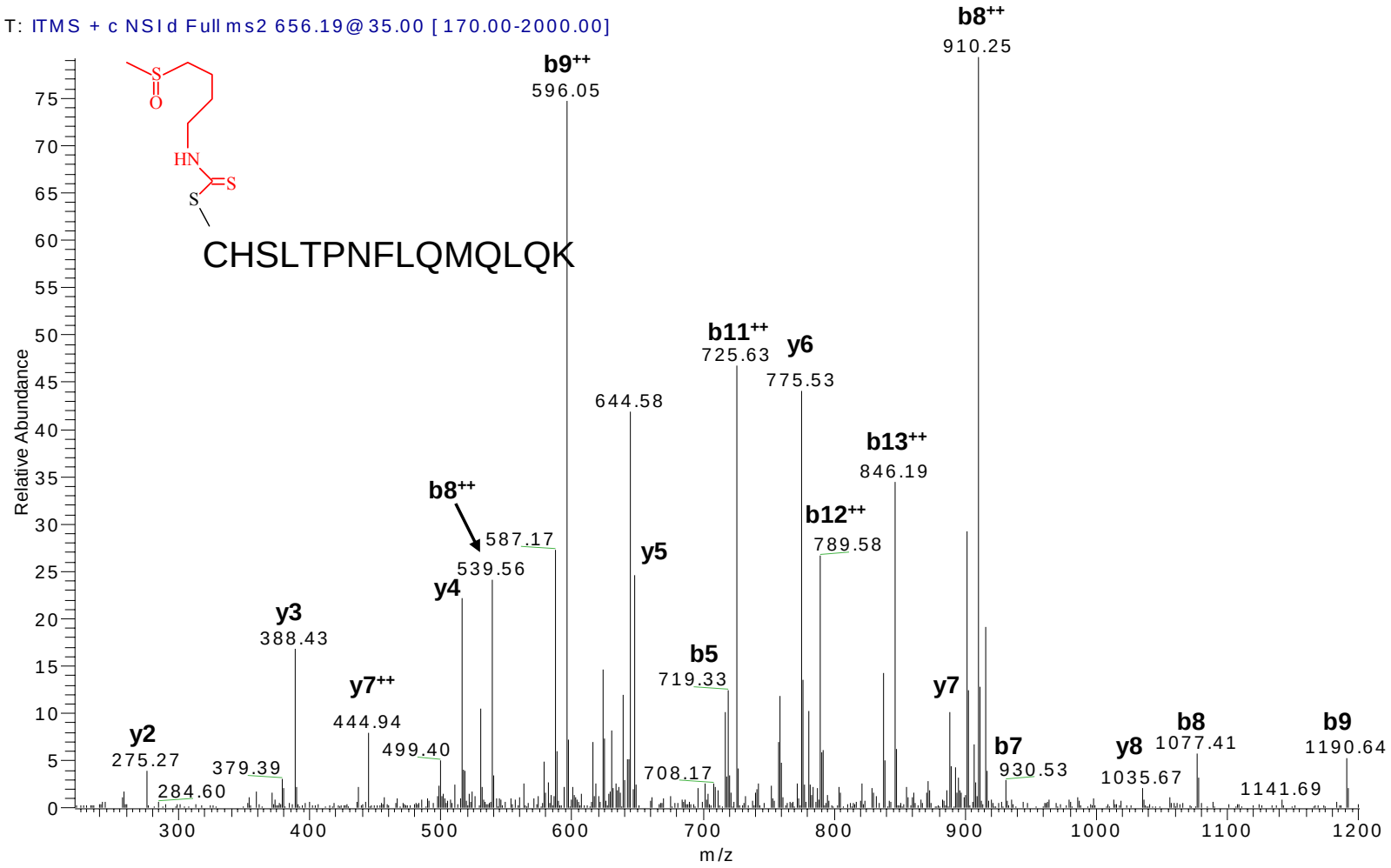

T: ITMS + c NSI d Full ms2 771.56@35.00 [200.00-2000.00]
CVLHVMNGAVMYQIDSVVR
b18++
1069.29
b15++
926.67
b17++
1019.75
b14++
869.23
y5
575.48
b10++
601.68
y4
460.40
y8
979.63
y6
688.56
b13++
812.63
b11++
666.83
y7
b12++
748.40
y10
1110.69
y2
274.31
y3
373.35
556.13
b10
1201.53
y11
b11
1333.64
b12
1495.71
295.23
423.33
1296.92
1369.65
Relative Abundance
m/z

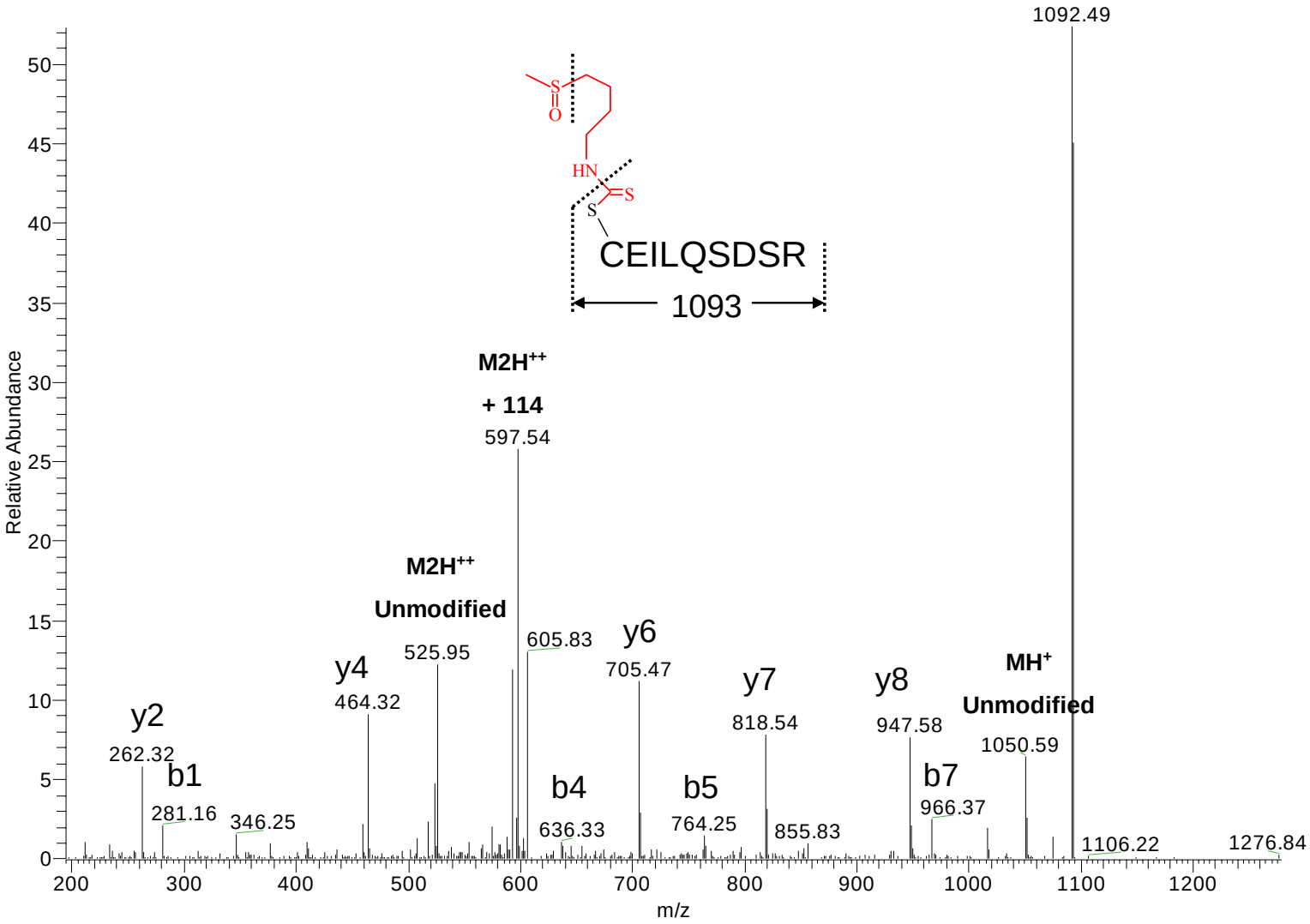

ACSDFLVQQLDPSNAIGIANFAEQIGCVELHQR
y22++
1273.34
y17++
1031.92
b11
1397.61
1506.94
y12
unmodified
y15++
946.83
y11
y27++
1564.69
y25++
y26++
y22++
unmodified
1184.28
b23++
y31++
1796.19
b20++
1118.62
b29++
b16
1879.95
b7
913.46
y16++
b9++
y18++
y28++1694.74
b6
unmodified
814.31
y4
553.46
y5
682.47
y7++
532.32
585.35
y32++
1935.18
Relative Abundance
m/z
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40
400
600
800
1000
1200
1400
1600
1800

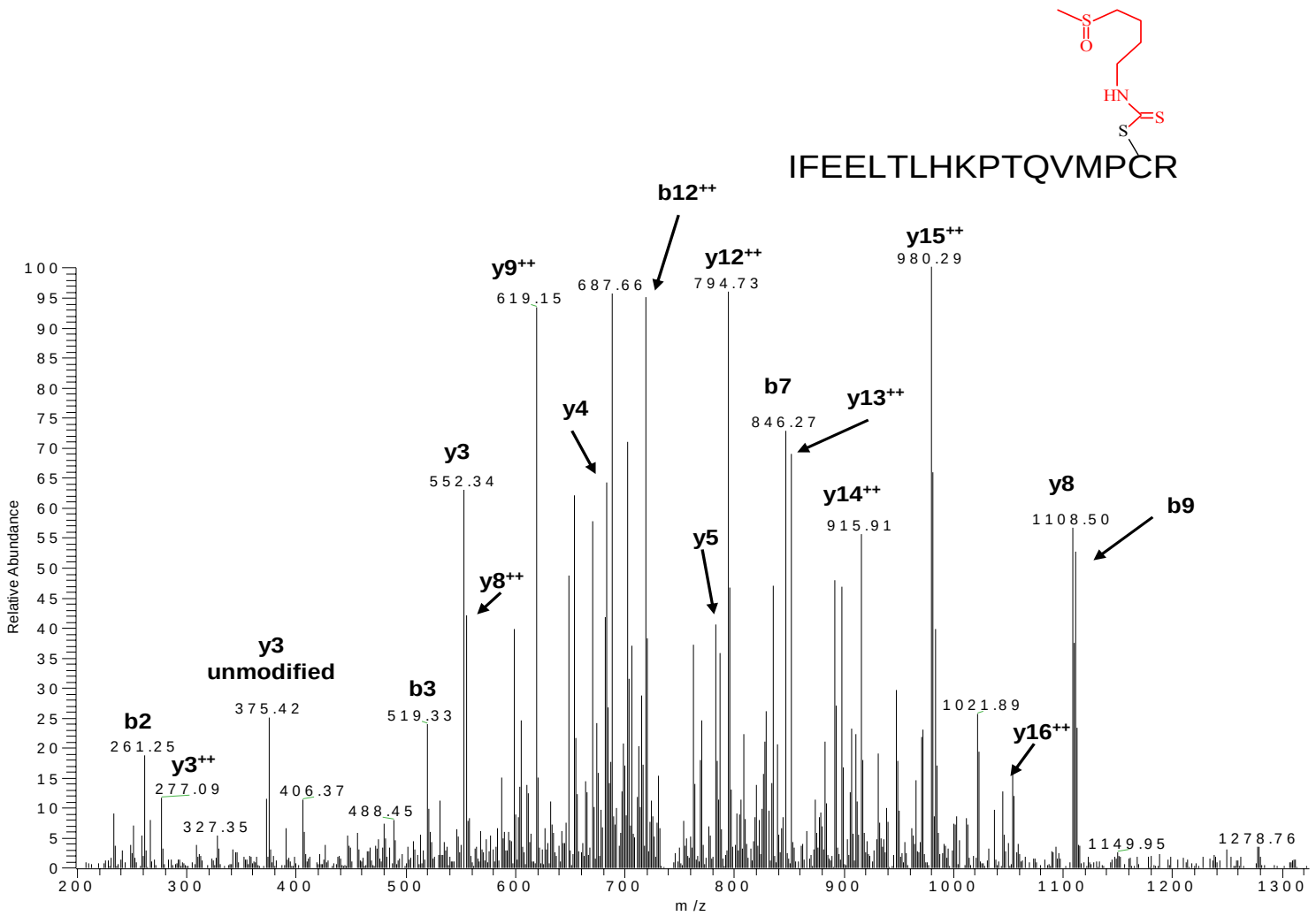

754.20
b11++
769.62
QEEFFNLSHCQLVTLISR
y11++
y5
589.52
819.16
y14++
904.83
y6
688.53
y13++
y16++
1043.31
y10++
674.20
y15++
978.33
y3
375.39
y7
y8
y4
488.51
517.32
1074.05
y2
262.32
y9
1209.63
1312.62
y10
y6++
345.03
456.54
1330.73
1524.69
Relative Abundance
m/z

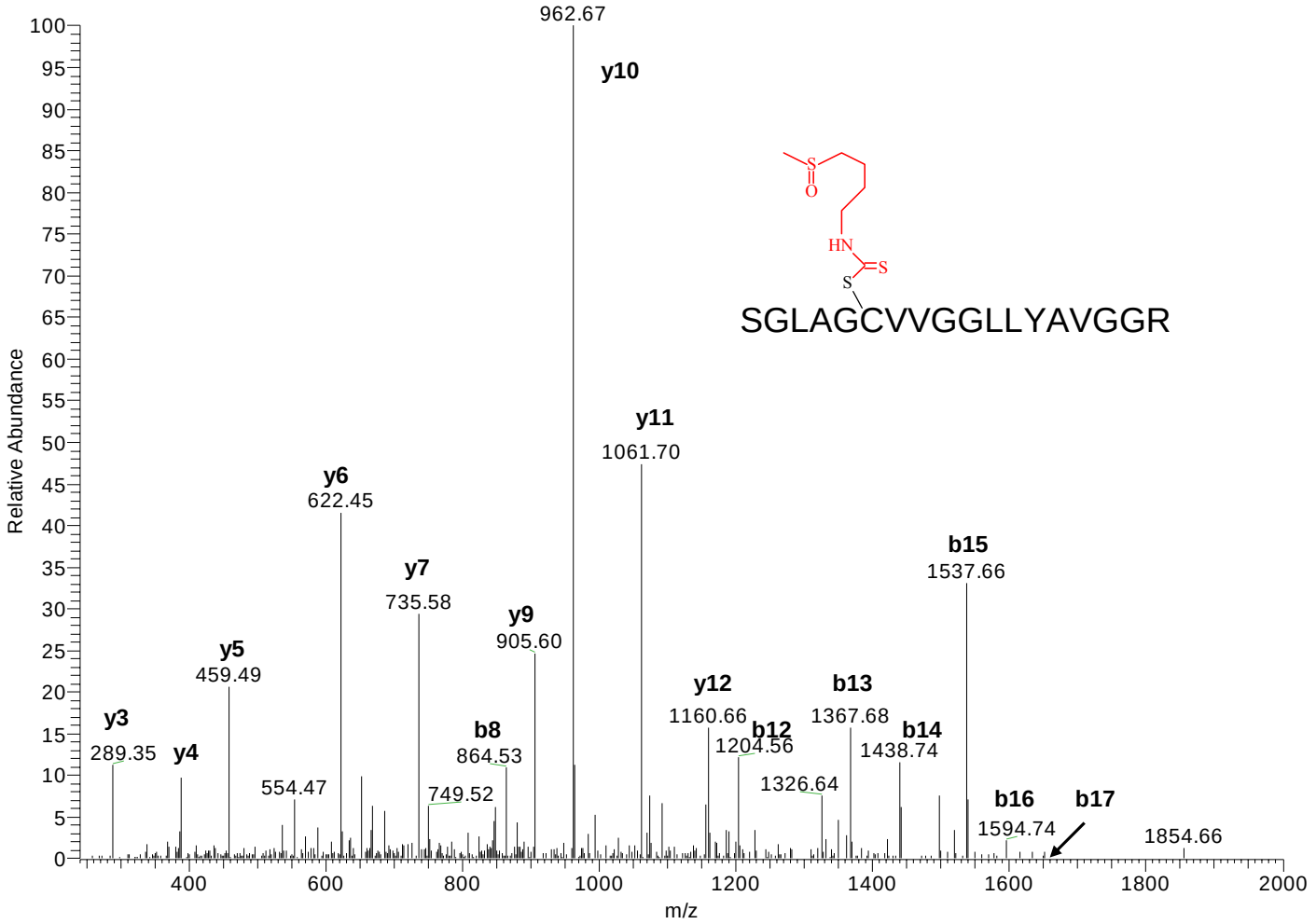

T: ITMS + c NSI d Full ms2 674.20@35.00 [175.00-2000.00]
942.74
CESEVFHACINWVK
M2H++
1 SFP modification
922.23
662.52
y10++
unmodified
y9++
646.94
y12++
805.67
b8++
unmodified
y13++
869.95
y13++
unmodified
781.51
y10++
697.59
y8++
574.67
y7++
506.26
b2
233.23
344.79
401.33
853.86
y14++
1010.73
1097.80
Relative Abundance
m/z

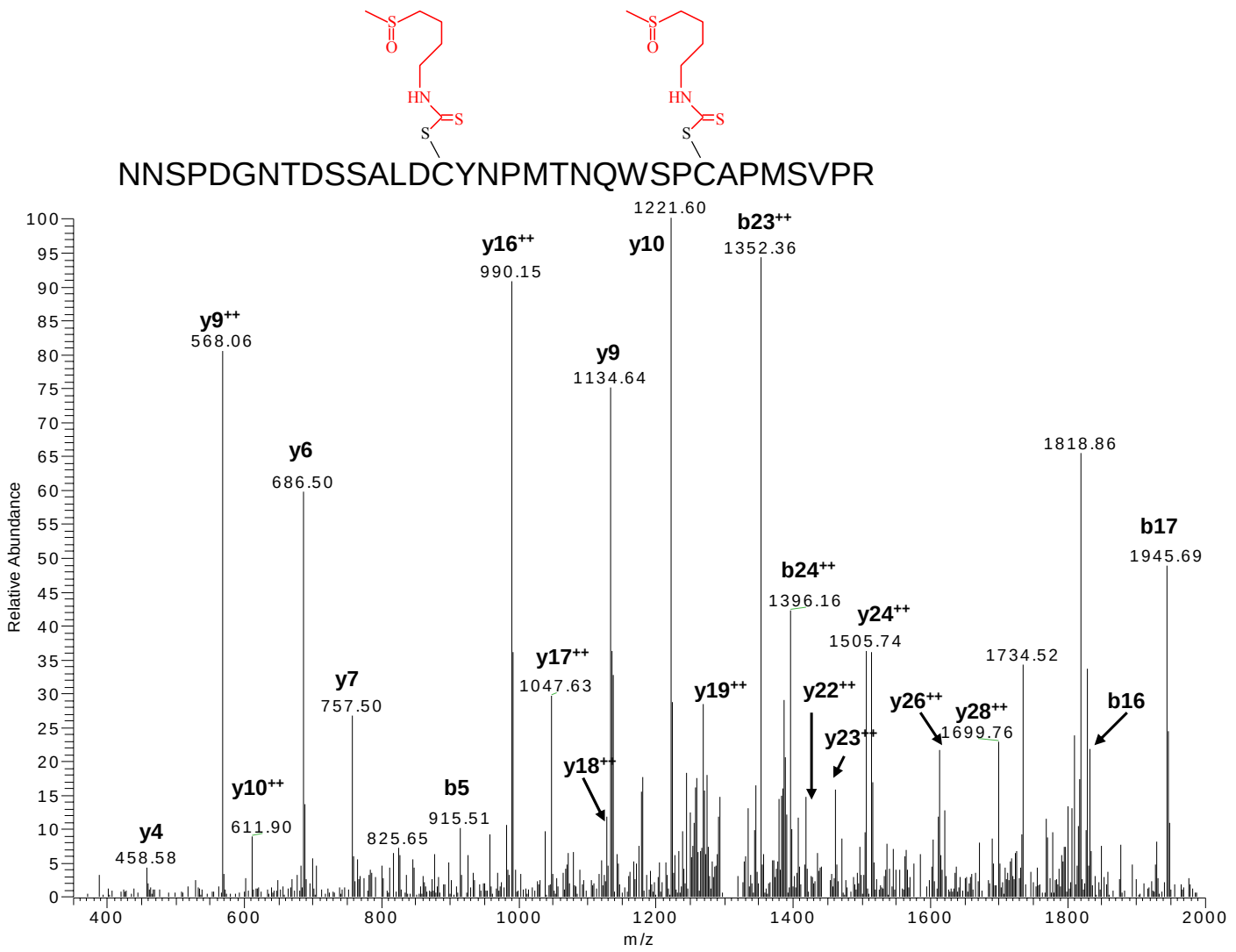

T: ITMS + c NSI d Full ms2 495.90@35.00 [125.00-1995.00]
YDCEQR
sulforaphane
407.45
M2H++
Unmodified
y4
712.34
813.43
MH+
Unmodified
487.08
y4
Unmodified
535.35
y2
303.26
b2
279.15
y3
y5
Unmodified
y5
b5
Unmodified
398.45
477.22
432.37
233.21
178.05
348.15
577.37
650.34
639.30
695.54
761.41
796.36
779.41
827.43
837.42
856.29
952.75
Relative Abundance
m/z

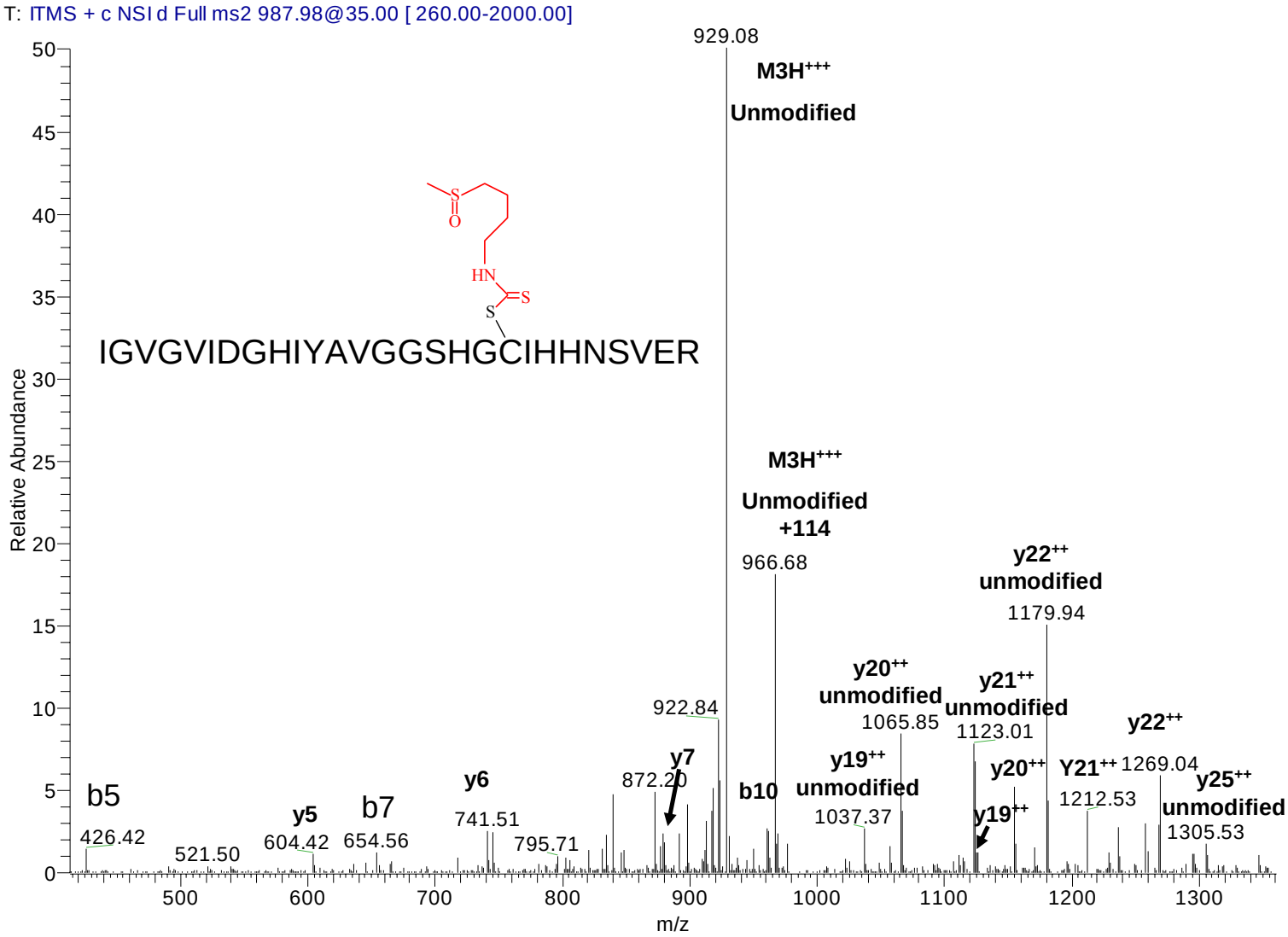

T: ITMS + c NSI d Full ms2 656.19@35.00 [170.00-2000.00]
CHSLTPNFLQMQLQK
b8++
910.25
b9++
596.05
b11++
725.63
y6
775.53
644.58
b13++
846.19
b8++
587.17
539.56
b12++
789.58
y4
y5
y3
388.43
b5
719.33
y7++
444.94
499.40
y7
y2
275.27
379.39
284.60
708.17
b7
930.53
y8
1035.67
b8
1077.41
1141.69
b9
1190.64
Relative Abundance
m/z
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
300
400
500
600
700
800
900
1000
1100
1200

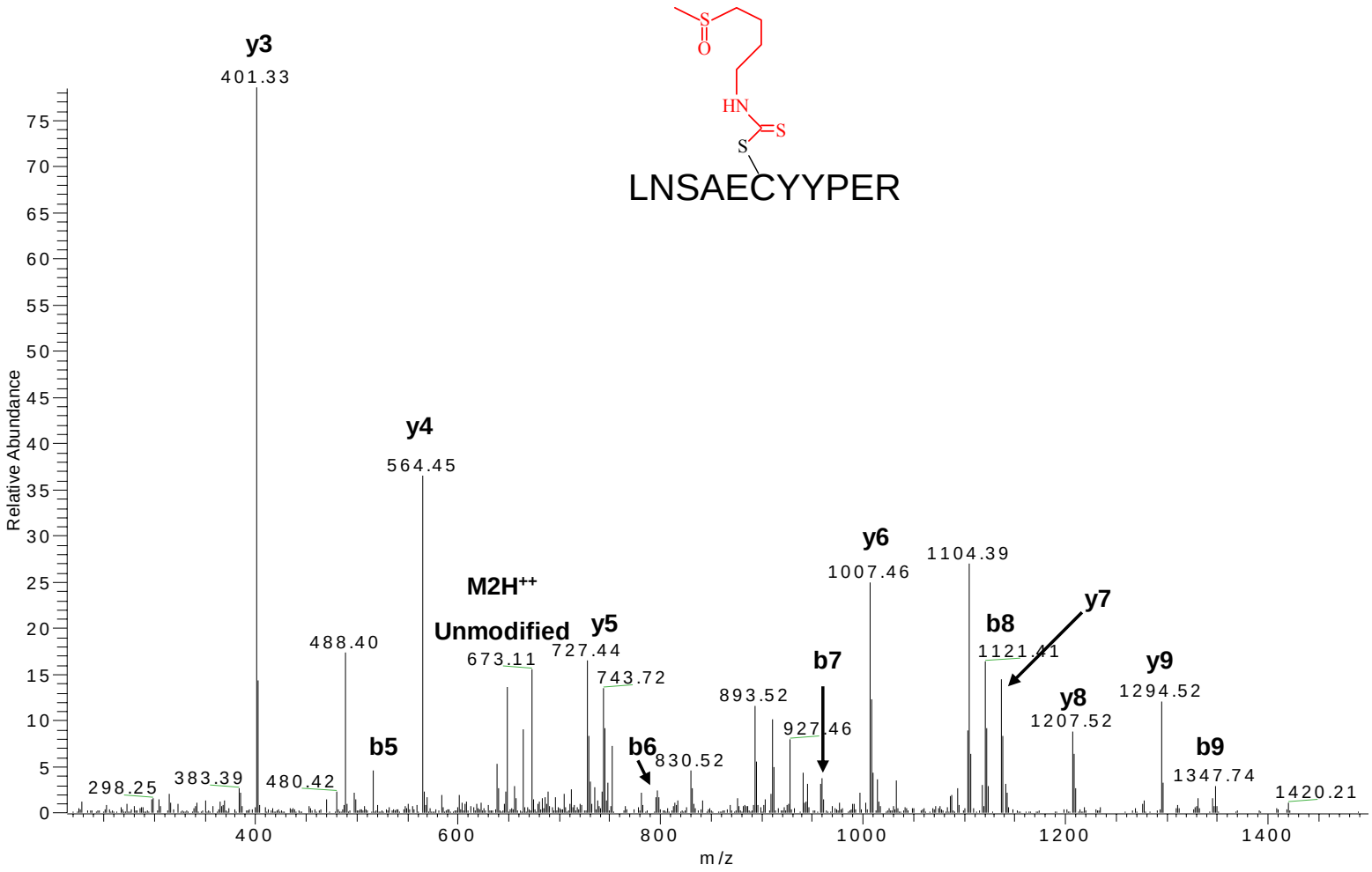

1092.49
CEILQSDSR
1093
M2H++
+ 114
597.54
M2H++
Unmodified
525.95
605.83
y6
705.47
y4
464.32
y7
818.54
y8
947.58
MH+
Unmodified
1050.59
y2
262.32
b1
281.16
346.25
b4
636.33
b5
764.25
855.83
b7
966.37
1106.22
1276.84
Relative Abundance
m/z

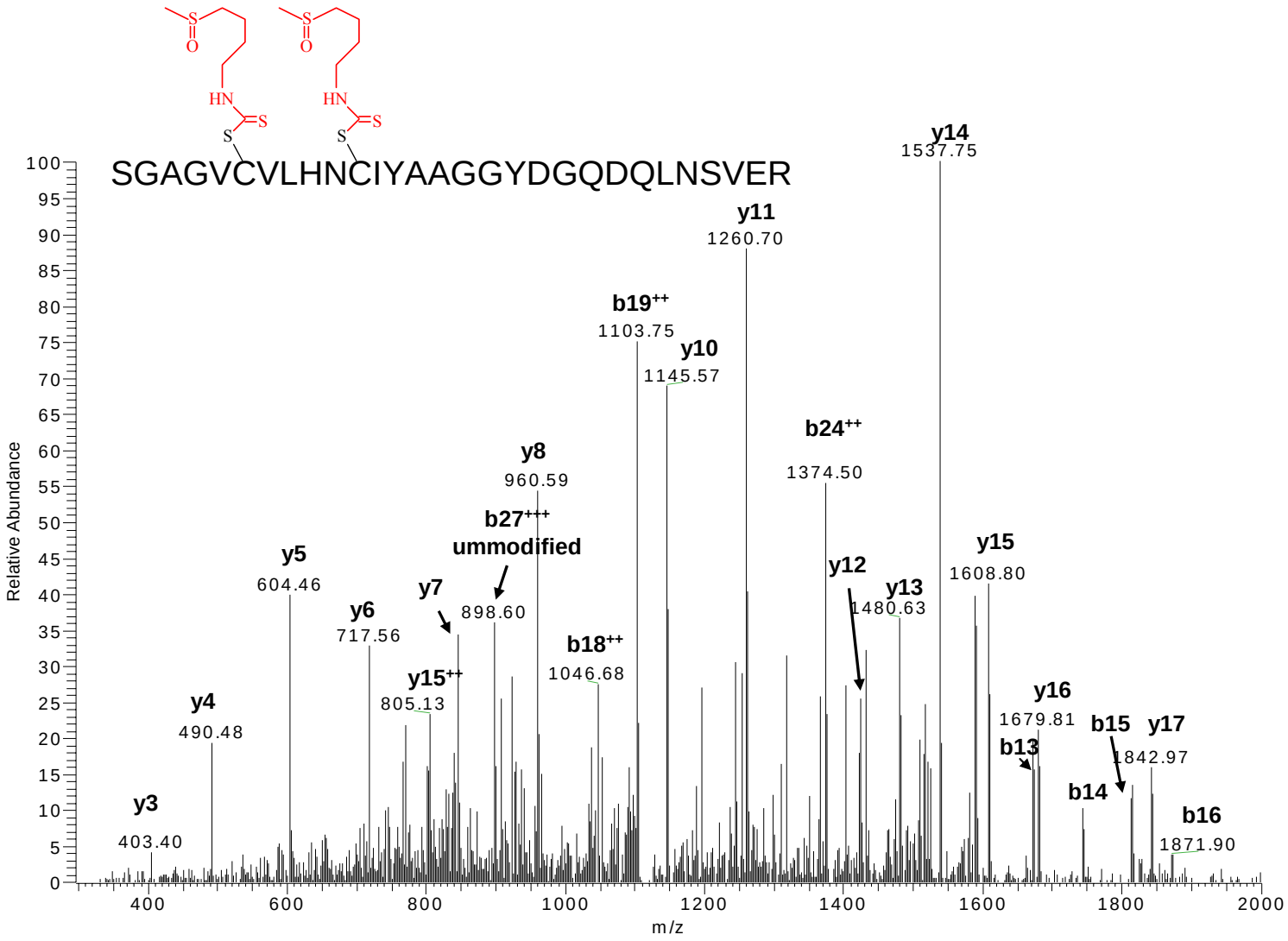

IFEELTLHKPTQVMPCR
b12++
687.66
y9++
619.15
y12++
794.73
y15++
980.29
b7
846.27
y13++
y4
y3
552.34
y14++
915.91
y8
1108.50
b9
y5
y8++
y3
unmodified
375.42
b3
519.33
b2
261.25
y3++
277.09
406.37
488.45
327.35
1021.89
y16++
1149.95
1278.76
Relative Abundance
m/z

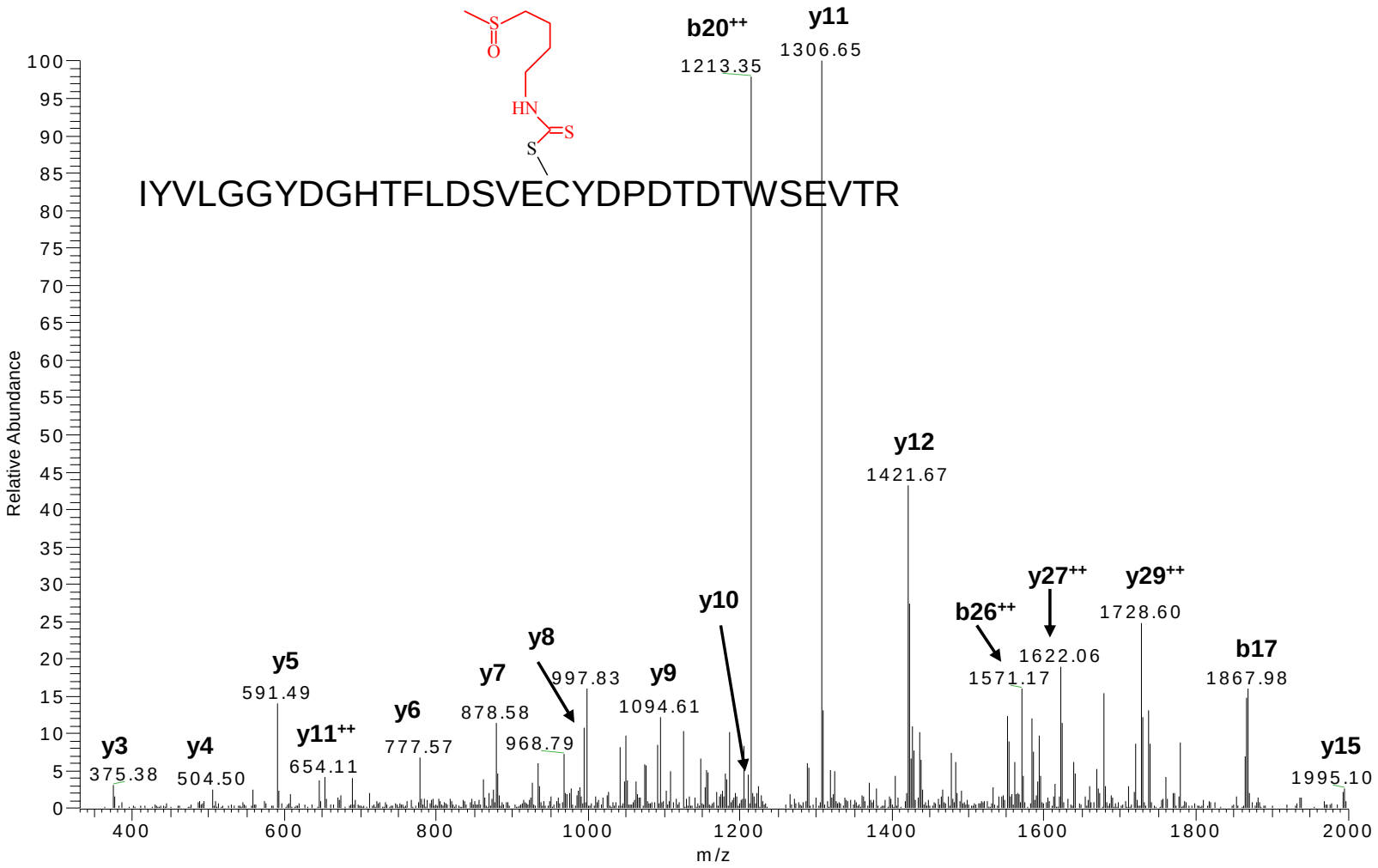

962.67
y10
y11
1061.70
y6
622.45
y7
735.58
y9
905.60
y5
459.49
y3
289.35
y4
554.47
b8
864.53
749.52
y12
1160.66
b12
1204.56
b13
1367.68
1326.64
b14
1438.74
b15
1537.66
b16
1594.74
b17
1854.66
SGLAGCVVGGLLYAVGGR
Relative Abundance
m/z
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
400
600
800
1000
1200
1400
1600
1800
2000

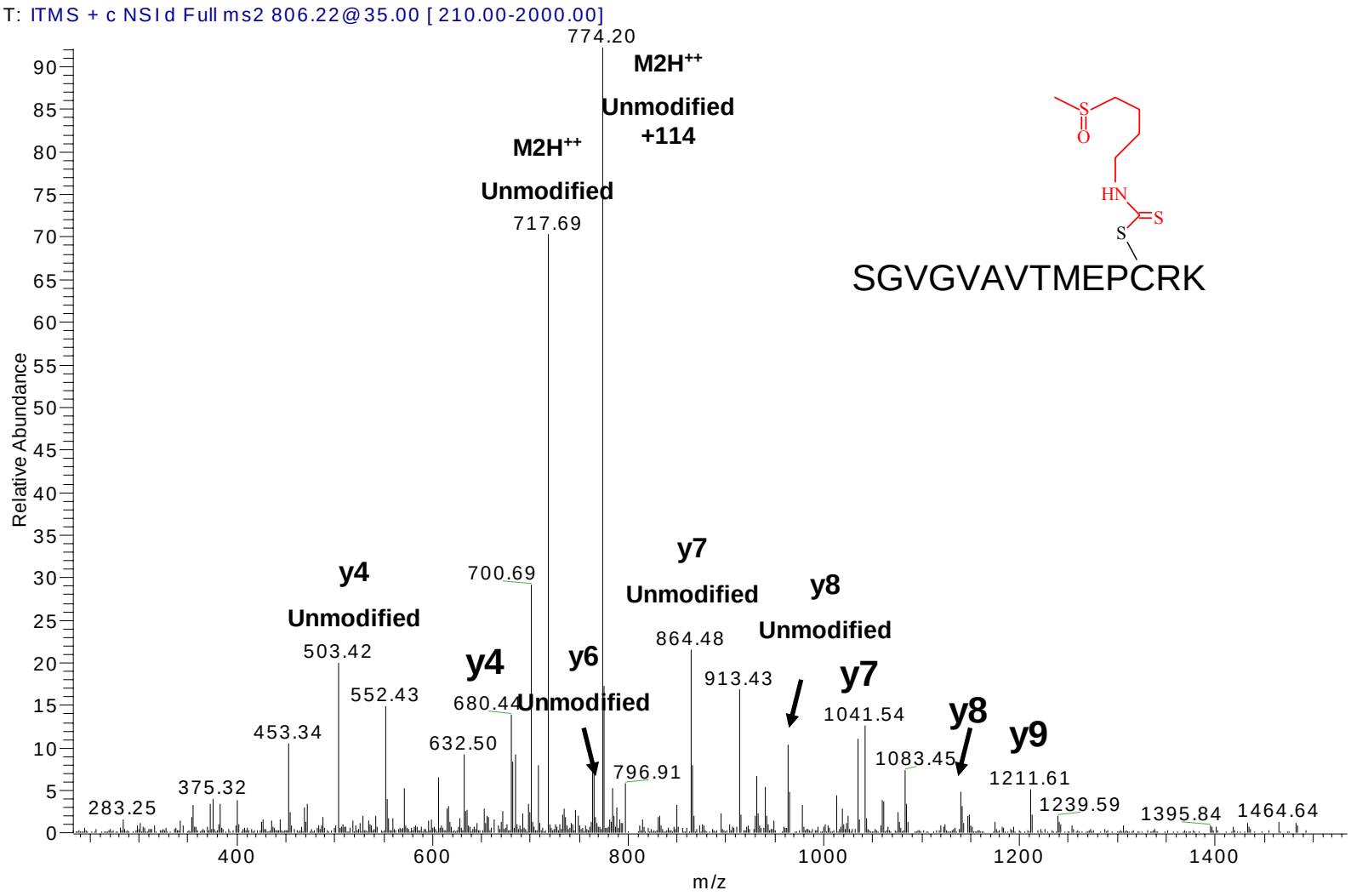

NNSPDGNTDSSALDCYNPMTNQWSPCAPMSVPR
Relative Abundance
m/z
y4
458.58
y9++
568.06
y10++
611.90
y6
686.50
y7
757.50
825.65
b5
915.51
y16++
990.15
y17++
1047.63
y18++
y9
1134.64
1221.60
y10
y19++
b23++
1352.36
b24++
1396.16
y22++
y23++
y24++
1505.74
y26++
y28++
1699.76
1734.52
1818.86
b16
b17
1945.69

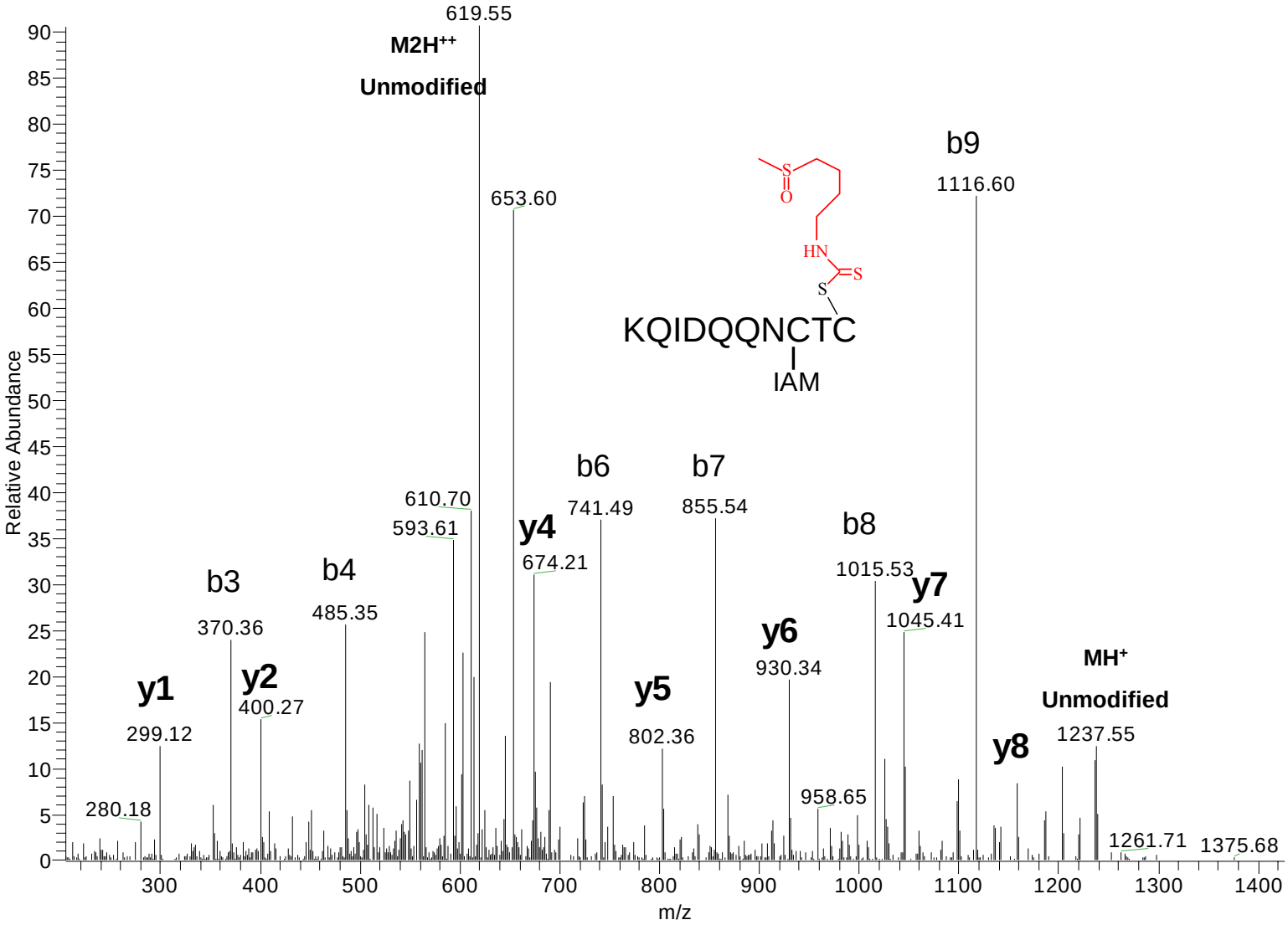

T: ITMS + c NSI d Full ms2 987.98@35.00 [260.00-2000.00]
929.08
M3H+++
Unmodified
IGVGVIDGHIYAVGGSHGCIHHNSVER
M3H+++
Unmodified
+114
966.68
y22++
unmodified
1179.94
y20++
unmodified
1065.85
y21++
unmodified
1123.01
922.84
y7
872.20
y6
741.51
b10
y19++
unmodified
1037.37
y20++
Y21++
1269.04
y22++
y25++
unmodified
1305.53
1212.53
y19++
b5
426.42
521.50
y5
604.42
b7
654.56
795.71
Relative Abundance
m/z
0
5
10
15
20
25
30
35
40
45
50
500
600
700
800
900
1000
1100
1200
1300


LNSAECYYPER
y3
401.33
y4
564.45
M2H++
Unmodified
673.11
y5
727.44
743.72
488.40
b5
298.25
383.39
480.42
b6
830.52
893.52
927.46
b7
y6
1007.46
1104.39
b8
1121.41
y7
y8
1207.52
y9
1294.52
b9
1347.74
1420.21
Relative Abundance
m/z
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
400
600
800
1000
1200
1400


SGAGVCVLHNCIYAAGGYDGQDQLNSVER
Relative Abundance
m/z
y14
1537.75
y11
1260.70
b19++
1103.75
y10
1145.57
b24++
1374.50
y8
960.59
b27+++
ummodified
y5
604.46
y7
898.60
y6
717.56
y12
y13
1480.63
y15
1608.80
b18++
1046.68
y15++
805.13
y4
490.48
y16
1679.81
b15
y17
1842.97
b13
b14
b16
1871.90
y3
403.40


IYVLGGYDGHTFLDSVECYDPDTDTWSEVTR
b20++
1213.35
y11
1306.65
y12
1421.67
y10
y8
997.83
y7
878.58
y9
1094.61
y6
777.57
y5
591.49
y11++
654.11
y4
504.50
y3
375.38
b26++
1571.17
y27++
1622.06
y29++
1728.60
b17
1867.98
y15
1995.10
Relative Abundance
m/z
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
400
600
800
1000
1200
1400
1600
1800
2000


T: ITMS + c NSI d Full ms2 806.22@35.00 [210.00-2000.00]
774.20
M2H++
Unmodified
+114
M2H++
Unmodified
717.69
SGVGVAVTMEPCRK
y4
Unmodified
503.42
y4
700.69
y6
Unmodified
y7
Unmodified
864.48
y8
Unmodified
913.43
y7
1041.54
y8
y9
1083.45
1211.61
1239.59
1395.84
1464.64
552.43
680.44
632.50
453.34
796.91
375.32
283.25
Relative Abundance
m/z


619.55
M2H++
Unmodified
653.60
b9
1116.60
KQIDQQNCTC
IAM
b6
741.49
b7
855.54
610.70
593.61
y4
674.21
b8
1015.53
y7
1045.41
b3
370.36
b4
485.35
y6
930.34
y1
299.12
y2
400.27
y5
802.36
MH+
Unmodified
1237.55
y8
958.65
280.18
1261.71
1375.68
Relative Abundance
m/z

REFERENCES

1. Hong, W. K. and Sporn, M. B. (1997) Recent advances in chemoprevention of cancer, *Science* ***278***, 1073-1077.

2. Steinmetz, K. A. and Potter, J. D. (1996) Vegetables, fruit, and cancer prevention: a review, *J. Am. Diet. Assoc.* ***96***, 1027-1039.

3. Terry, P., Giovannucci, E., Michels, K. B., Bergkvist, L., Hansen, H., Holmberg, L., and Wolk, A. (2001) Fruit, vegetables, dietary fiber, and risk of colorectal cancer, *J. Natl. Cancer Inst.* ***93***, 525-533.

4. Cohen, J. H., Kristal, A. R., and Stanford, J. L. (2000) Fruit and vegetable intakes and prostate cancer risk, *J. Natl. Cancer Inst.* ***92***, 61-68.

5. Surh, Y. J. (2003) Cancer chemoprevention with dietary phytochemicals, *Nat. Rev. Cancer* ***3***, 768-780.

6. Greenwald, P. (1996) Chemoprevention of cancer, *Sci. Am.* ***275***, 96-99.

7. Sporn, M. B. (1993) Chemoprevention of cancer, *Lancet* ***342***, 1211-1213.

8. Chen, C. and Kong, A. N. (2004) Dietary chemopreventive compounds and ARE/EpRE signaling, *Free Radic. Biol. Med.* ***36***, 1505-1516.

9. Miller, E. and Miller JA (1985) Some historical perspectives on the metabolism of xenobiotic chemicals to reactive intermediates, pp 1-28.

10. Wilkinson, J. and Clapper, M. L. (1997) Detoxication enzymes and chemoprevention, *Proc. Soc. Exp. Biol. Med.* ***216***, 192-200.

11. Talalay, P., De Long, M. J., and Prochaska, H. J. (1988) Identification of a common chemical signal regulating the induction of enzymes that protect against chemical carcinogenesis., *Proc. Natl. Acad. Sci. USA* ***85***, 8261-8265.

12. Landi, S. (2000) Mammalian class theta GST and differential susceptibility to carcinogens: a review, *Mutat. Res.* ***463***, 247-283.

13. Prochaska, H. J. and Talalay, P. (1988) Regulatory mechanisms of monofunctional and bifunctional anticarcinogenic enzyme inducers in murine liver, *Cancer Res.* ***48***, 4776-4782.

14. Sidransky, H., Ito, N., and Verney, E. (1966) Influence of alpha-naphthyl-isothiocyanate on liver tumorigenesis in rats ingesting ethionine and N-2-fluorenylacetamide, *J. Natl. Cancer Inst.* ***37***, 677-686.

15. Verhoeven, D. T., Verhagen, H., Goldbohm, R. A., Van den Brandt, P. A., and van, P. G. (1997) A review of mechanisms underlying anticarcinogenicity by brassica vegetables, *Chem. Biol. Interact.* ***103***, 79-129.

16. Hecht, S. S. (1999) Chemoprevention of cancer by isothiocyanates, modifiers of carcinogen metabolism, *J. Nutr.* ***129***, 768S-774S.

17. Conaway, C. C., Yang, Y. M., and Chung, F. L. (2002) Isothiocyanates as cancer chemopreventive agents: their biological activities and metabolism in rodents and humans, *Curr. Drug Metab* ***3***, 233-255.

18. Fahey, J. W., Zalcmann, A. T., and Talalay, P. (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants, *Phytochemistry* ***56***, 5-51.

19. Lin, H. J., Probst-Hensch, N. M., Louie, A. D., Kau, I. H., Witte, J. S., Ingles, S. A., Frankl, H. D., Lee, E. R., and Haile, R. W. (1998) Glutathione transferase null genotype, broccoli, and lower prevalence of colorectal adenomas, *Cancer Epidemiol. Biomarkers Prev.* ***7***, 647-652.

20. Chung, F. L., Conaway, C. C., Rao, C. V., and Reddy, B. S. (2000) Chemoprevention of colonic aberrant crypt foci in Fischer rats by sulforaphane and phenethyl isothiocyanate, *Carcinogenesis* ***21***, 2287-2291.

21. Chiao, J. W., Chung, F. L., Kancherla, R., Ahmed, T., Mittelman, A., and Conaway, C. C. (2002) Sulforaphane and its metabolite mediate growth arrest and apoptosis in human prostate cancer cells, *Int. J. Oncol.* ***20***, 631-636.

22. Fahey, J. W., Haristoy, X., Dolan, P. M., Kensler, T. W., Scholtus, I., Stephenson, K. K., Talalay, P., and Lozniewski, A. (2002) Sulforaphane inhibits extracellular, intracellular, and antibiotic-resistant strains of Helicobacter pylori and prevents benzo[a]pyrene-induced stomach tumors, *Proc. Natl. Acad. Sci. U. S. A* ***99***, 7610-7615.

23. Talalay, P., Fahey, J. W., Holtzclaw, W. D., Prestera, T., and Zhang, Y. (1995) Chemoprotection against cancer by phase 2 enzyme induction, *Toxicol. Lett.* ***82-83***, 173-179.

24. Talalay, P., Dinkova-Kostova, A. T., and Holtzclaw, W. D. (2003) Importance of phase 2 gene regulation in protection against electrophile and reactive oxygen toxicity and carcinogenesis, *Adv. Enzyme Regul.* ***43***, 121-134.

25. Rushmore, T. H. and Pickett, C. B. (1990) Transcriptional regulation of the rat glutathione S-transferase Ya subunit gene. Characterization of a xenobiotic-responsive element controlling inducible expression by phenolic antioxidants, *J. Biol. Chem.* ***265***, 14648-14653.

26. Telakowski-Hopkins, C. A., King, R. G., and Pickett, C. B. (1988) Glutathione S-transferase Ya subunit gene: identification of regulatory elements required for basal level and inducible expression, *Proc. Natl. Acad. Sci. U. S. A* ***85***, 1000-1004.

27. Friling, R. S., Bensimon, A., Tichauer, Y., and Daniel, V. (1990) Xenobiotic-inducible expression of murine glutathione S-transferase Ya subunit gene is controlled by an electrophile-responsive element, *Proc. Natl. Acad. Sci. U. S. A* ***87***, 6258-6262.

28. Moinova, H. R. and Mulcahy, R. T. (1998) An electrophile responsive element (EpRE) regulates beta-naphthoflavone induction of the human gamma-glutamylcysteine synthetase regulatory subunit gene. Constitutive expression is mediated by an adjacent AP-1 site, *J. Biol. Chem.* ***273***, 14683-14689.

29. Prestera, T., Talalay, P., Alam, J., Ahn, Y. I., Lee, P. J., and Choi, A. M. (1995) Parallel induction of heme oxygenase-1 and chemoprotective phase 2 enzymes by electrophiles and antioxidants: regulation by upstream antioxidant-responsive elements (ARE), *Mol. Med.* ***1***, 827-837.

30. Primiano, T., Gastel, J. A., Kensler, T. W., and Sutter, T. R. (1996) Isolation of cDNAs representing dithiolethione-responsive genes, *Carcinogenesis* ***17***, 2297-2303.

31. Li, Y. and Jaiswal, A. K. (1992) Regulation of human NAD(P)H:quinone oxidoreductase gene. Role of AP1 binding site contained within human antioxidant response element, *J. Biol. Chem.* ***267***, 15097-15104.

32. Okuda, A., Imagawa, M., Maeda, Y., Sakai, M., and Muramatsu, M. (1989) Structural and functional analysis of an enhancer GPEI having a phorbol 12-O-tetradecanoate 13-acetate responsive element-like sequence found in the rat glutathione transferase P gene, *J. Biol. Chem.* ***264***, 16919-16926.

33. Wasserman, W. W. and Fahl, W. E. (1997) Functional antioxidant responsive elements, *Proc. Natl. Acad. Sci. U. S. A* ***94***, 5361-5366.

34. Zhang, Y., Gonzalez, V., and Xu, M. J. (2002) Expression and regulation of glutathione S-transferase P1-1 in cultured human epidermal cells, *J. Dermatol. Sci.* ***30***, 205-214.

35. Nioi, P., McMahon, M., Itoh, K., Yamamoto, M., and Hayes, J. D. (2003) Identification of a novel Nrf2-regulated antioxidant response element (ARE) in the mouse NAD(P)H:quinone oxidoreductase 1 gene: reassessment of the ARE consensus sequence, *Biochem. J.* ***374***, 337-348.

36. Friling, R. S., Bergelson, S., and Daniel, V. (1992) Two adjacent AP-1-like binding sites form the electrophile-responsive element of the murine glutathione S-transferase Ya subunit gene, *Proc. Natl. Acad. Sci. U. S. A* ***89***, 668-672.

37. Xie, T., Belinsky, M., Xu, Y., and Jaiswal, A. K. (1995) ARE- and TRE-mediated regulation of gene expression. Response to xenobiotics and antioxidants, *J. Biol. Chem.* ***270***, 6894-6900.

38. Venugopal, R. and Jaiswal, A. K. (1996) Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate the human antioxidant response element-mediated expression of NAD(P)H:quinone oxidoreductase1 gene, *Proc. Natl. Acad. Sci. U. S. A* ***93***, 14960-14965.

39. Wilkinson, J., Radjendirane, V., Pfeiffer, G. R., Jaiswal, A. K., and Clapper, M. L. (1998) Disruption of c-Fos leads to increased expression of NAD(P)H:quinone oxidoreductase1 and glutathione S-transferase, *Biochem. Biophys. Res. Commun.* ***253***, 855-858.

40. Chan, J. Y., Han, X. L., and Kan, Y. W. (1993) Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast, *Proc. Natl. Acad. Sci. U. S. A* ***90***, 11371-11375.

41. Moi, P., Chan, K., Asunis, I., Cao, A., and Kan, Y. W. (1994) Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region, *Proc. Natl. Acad. Sci. U. S. A* ***91***, 9926-9930.

42. Chan, K., Lu, R., Chang, J. C., and Kan, Y. W. (1996) NRF2, a member of the NFE2 family of transcription factors, is not essential for murine erythropoiesis, growth, and development, *Proc. Natl. Acad. Sci. U. S. A* ***93***, 13943-13948.

43. Kobayashi, A., Ito, E., Toki, T., Kogame, K., Takahashi, S., Igarashi, K., Hayashi, N., and Yamamoto, M. (1999) Molecular cloning and functional characterization of a new Cap'n' collar family transcription factor Nrf3, *J. Biol. Chem.* ***274***, 6443-6452.

44. Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J. D., and Yamamoto, M. (1999) Keap1 represses nuclear activation of antioxidant

responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain, *Genes Dev.* ***13***, 76-86.

45. Katoh, Y., Itoh, K., Yoshida, E., Miyagishi, M., Fukamizu, A., and Yamamoto, M. (2001) Two domains of Nrf2 cooperatively bind CBP, a CREB binding protein, and synergistically activate transcription, *Genes Cells* ***6***, 857-868.

46. Ishii, T., Itoh, K., Takahashi, S., Sato, H., Yanagawa, T., Katoh, Y., Bannai, S., and Yamamoto, M. (2000) Transcription factor Nrf2 coordinately regulates a group of oxidative stress-inducible genes in macrophages, *J. Biol. Chem.* ***275***, 16023-16029.

47. Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., Hatayama, I., Yamamoto, M., and Nabeshima, Y. (1997) An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements, *Biochem. Biophys. Res. Commun.* ***236***, 313-322.

48. Moinova, H. R. and Mulcahy, R. T. (1999) Up-regulation of the human gamma-glutamylcysteine synthetase regulatory subunit gene involves binding of Nrf-2 to an electrophile responsive element, *Biochem. Biophys. Res. Commun.* ***261***, 661-668.

49. Wild, A. C., Moinova, H. R., and Mulcahy, R. T. (1999) Regulation of gamma-glutamylcysteine synthetase subunit gene expression by the transcription factor Nrf2, *J. Biol. Chem.* ***274***, 33627-33636.

50. Alam, J., Stewart, D., Touchard, C., Boinapally, S., Choi, A. M., and Cook, J. L. (1999) Nrf2, a Cap'n'Collar transcription factor, regulates induction of the heme oxygenase-1 gene, *J. Biol. Chem.* ***274***, 26071-26078.

51. McMahon, M., Itoh, K., Yamamoto, M., Chanas, S. A., Henderson, C. J., McLellan, L. I., Wolf, C. R., Cavin, C., and Hayes, J. D. (2001) The Cap'n'Collar basic leucine zipper transcription factor Nrf2 (NF-E2 p45-related factor 2) controls both constitutive and inducible expression of intestinal detoxification and glutathione biosynthetic enzymes, *Cancer Res.* ***61***, 3299-3307.

52. Kwak, M. K., Itoh, K., Yamamoto, M., Sutter, T. R., and Kensler, T. W. (2001) Role of transcription factor Nrf2 in the induction of hepatic phase 2 and antioxidative enzymes in vivo by the cancer chemoprotective agent, 3H-1, 2-dimethiole-3-thione, *Mol. Med.* ***7***, 135-145.

53. Ramos-Gomez, M., Kwak, M. K., Dolan, P. M., Itoh, K., Yamamoto, M., Talalay, P., and Kensler, T. W. (2001) Sensitivity to carcinogenesis is increased and chemoprotective efficacy of enzyme inducers is lost in nrf2 transcription factor-deficient mice, *Proc. Natl. Acad. Sci. U. S. A* ***98***, 3410-3415.

54. Thimmulappa, R. K., Mai, K. H., Srisuma, S., Kensler, T. W., Yamamoto, M., and Biswal, S. (2002) Identification of Nrf2-regulated genes induced by the chemopreventive agent sulforaphane by oligonucleotide microarray, *Cancer Res.* ***62***, 5196-5203.

55. Chan, K. and Kan, Y. W. (1999) Nrf2 is essential for protection against acute pulmonary injury in mice, *Proc. Natl. Acad. Sci. U. S. A* ***96***, 12731-12736.

56. Chan, K., Han, X. D., and Kan, Y. W. (2001) An important function of Nrf2 in combating oxidative stress: detoxification of acetaminophen, *Proc. Natl. Acad. Sci. U. S. A* ***98***, 4611-4616.

57. Enomoto, A., Itoh, K., Nagayoshi, E., Haruta, J., Kimura, T., O'Connor, T., Harada, T., and Yamamoto, M. (2001) High sensitivity of Nrf2 knockout mice to acetaminophen hepatotoxicity associated with decreased expression of ARE-regulated drug metabolizing enzymes and antioxidant genes, *Toxicol. Sci.* ***59***, 169-177.

58. Cho, H. Y., Jedlicka, A. E., Reddy, S. P., Zhang, L. Y., Kensler, T. W., and Kleeberger, S. R. (2002) Linkage analysis of susceptibility to hyperoxia. Nrf2 is a candidate gene, *Am. J. Respir. Cell Mol. Biol.* ***26***, 42-51.

59. Cho, H. Y., Jedlicka, A. E., Reddy, S. P., Kensler, T. W., Yamamoto, M., Zhang, L. Y., and Kleeberger, S. R. (2002) Role of NRF2 in protection against hyperoxic lung injury in mice, *Am. J. Respir. Cell Mol. Biol.* ***26***, 175-182.

60. Ramos-Gomez, M., Dolan, P. M., Itoh, K., Yamamoto, M., and Kensler, T. W. (2003) Interactive effects of nrf2 genotype and oltipraz on benzo[a]pyrene-DNA adducts and tumor yield in mice, *Carcinogenesis* ***24***, 461-467.

61. Aoki, Y., Sato, H., Nishimura, N., Takahashi, S., Itoh, K., and Yamamoto, M. (2001) Accelerated DNA adduct formation in the lung of the Nrf2 knockout mouse exposed to diesel exhaust, *Toxicol. Appl. Pharmacol.* ***173***, 154-160.

62. Kwak, M. K., Wakabayashi, N., Itoh, K., Motohashi, H., Yamamoto, M., and Kensler, T. W. (2002) Modulation of gene expression by cancer chemopreventive dithiolethiones through the Keap1-Nrf2 pathway:

Identification of novel gene clusters for cell survival, *J. Biol. Chem.* ***278***, 8135-8145.

63. Lee, J. M., Calkins, M. J., Chan, K., Kan, Y. W., and Johnson, J. A. (2003) Identification of the NF-E2-related factor 2-dependent genes conferring protection against oxidative stress in primary cortical astrocytes using oligonucleotide microarray analysis, *J. Biol. Chem.* ***278***, 37948-37956.

64. Shih, A. Y., Johnson, D. A., Wong, G., Kraft, A. D., Jiang, L., Erb, H., Johnson, J. A., and Murphy, T. H. (2003) Coordinate regulation of glutathione biosynthesis and release by Nrf2-expressing glia potently protects neurons from oxidative stress, *J. Neurosci.* ***23***, 3394-3406.

65. Xue, F. and Cooley, L. (1993) kelch encodes a component of intercellular bridges in Drosophila egg chambers, *Cell* ***72***, 681-693.

66. Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., O'Connor, T., and Yamamoto, M. (2003) Keap1 regulates both cytoplasmic-nuclear shuttling and degradation of Nrf2 in response to electrophiles, *Genes Cells* ***8***, 379-391.

67. Zipper, L. M. and Mulcahy, R. T. (2002) The Keap1 BTB/POZ dimerization function is required to sequester Nrf2 in cytoplasm, *J. Biol. Chem.* ***277***, 36544-36552.

68. Wakabayashi, N., Itoh, K., Wakabayashi, J., Motohashi, H., Noda, S., Takahashi, S., Imakado, S., Kotsuji, T., Otsuka, F., Roop, D. R., Harada, T., Engel, J. D., and Yamamoto, M. (2003) Keap1-null mutation leads to postnatal lethality due to constitutive Nrf2 activation, *Nat. Genet.* ***35***, 238-245.

69. Dinkova-Kostova, A. T., Holtzclaw, W. D., Cole, R. N., Itoh, K., Wakabayashi, N., Katoh, Y., Yamamoto, M., and Talalay, P. (2002) Direct evidence that sulfhydryl groups of Keap1 are the sensors regulating induction of phase 2 enzymes that protect against carcinogens and oxidants, *Proc. Natl. Acad. Sci. U. S. A* ***99***, 11908-11913.

70. Wakabayashi, N., Dinkova-Kostova, A. T., Holtzclaw, W. D., Kang, M. I., Kobayashi, A., Yamamoto, M., Kensler, T. W., and Talalay, P. (2004) Protection against electrophile and oxidant stress by induction of the phase 2 response: Fate of cysteines of the Keap1 sensor modified by inducers, *Proc. Natl. Acad. Sci. U. S. A* ***101***, 2040-2045.

71. Sekhar, K. R., Soltaninassab, S. R., Borrelli, M. J., Xu, Z. Q., Meredith, M. J., Domann, F. E., and Freeman, M. L. (2000) Inhibition of the 26S proteasome

induces expression of GLCLC, the catalytic subunit for gamma-glutamylcysteine synthetase, *Biochem. Biophys. Res. Commun.* ***270***, 311-317.

72. Kwak, M. K., Itoh, K., Yamamoto, M., and Kensler, T. W. (2002) Enhanced expression of the transcription factor Nrf2 by cancer chemopreventive agents: role of antioxidant response element-like sequences in the nrf2 promoter, *Mol. Cell Biol.* ***22***, 2883-2892.

73. McMahon, M., Itoh, K., Yamamoto, M., and Hayes, J. D. (2003) Keap1-dependent proteasomal degradation of transcription factor Nrf2 contributes to the negative regulation of antioxidant response element-driven gene expression, *J. Biol. Chem.* ***278,*** 21592-21600.

74. Zhang, D. D. and Hannink, M. (2003) Distinct cysteine residues in Keap1 are required for Keap1-dependent ubiquitination of Nrf2 and for stabilization of Nrf2 by chemopreventive agents and oxidative stress, *Mol Cell Biol.* ***23***, 8137-8151.

75. Zhang, D. D., Lo, S. C., Cross, J. V., Templeton, D. J., and Hannink, M. (2004) Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex, *Mol. Cell Biol.* ***24***, 10941-10953.

76. Cullinan, S. B., Gordan, J. D., Jin, J., Harper, J. W., and Diehl, J. A. (2004) The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: oxidative stress sensing by a Cul3-Keap1 ligase, *Mol. Cell Biol.* ***24***, 8477-8486.

77. Furukawa, M. and Xiong, Y. (2005) BTB protein Keap1 targets antioxidant transcription factor Nrf2 for ubiquitination by the Cullin 3-Roc1 ligase, *Mol. Cell Biol.* ***25***, 162-171.

78. Kobayashi, A., Kang, M. I., Okawa, H., Ohtsuji, M., Zenke, Y., Chiba, T., Igarashi, K., and Yamamoto, M. (2004) Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2, *Mol. Cell Biol.* ***24***, 7130-7139.

79. Yu, R., Tan, T. H., and Kong, A. T. (1997) Butylated hydroxyanisole and its metabolite tert-butylhydroquinone differentially regulate mitogen-activated protein kinases. The role of oxidative stress in the activation of mitogen-activated protein kinases by phenolic antioxidants, *J. Biol. Chem.* ***272***, 28962-28970.

80. Yu, R., Lei, W., Mandlekar, S., Weber, M. J., Der, C. J., Wu, J., and Kong, A. T. (1999) Role of a mitogen-activated protein kinase pathway in the induction

of phase II detoxifying enzymes by chemicals, *J. Biol. Chem.* ***274***, 27545-27552.

81. Yu, R., Chen, C., Mo, Y. Y., Hebbar, V., Owuor, E. D., Tan, T. H., and Kong, A. N. (2000) Activation of mitogen-activated protein kinase pathways induces antioxidant response element-mediated gene expression via a Nrf2-dependent mechanism, *J. Biol. Chem.* ***275***, 39907-39913.

82. Kang, K. W., Ryu, J. H., and Kim, S. G. (2000) The essential role of phosphatidylinositol 3-kinase and of p38 mitogen-activated protein kinase activation in the antioxidant response element-mediated rGSTA2 induction by decreased glutathione in H4IIE hepatoma cells, *Mol. Pharmacol.* ***58***, 1017-1025.

83. Kang, K. W., Cho, M. K., Lee, C. H., and Kim, S. G. (2001) Activation of phosphatidylinositol 3-kinase and Akt by tert-butylhydroquinone is responsible for antioxidant response element-mediated rGSTA2 induction in H4IIE cells, *Mol. Pharmacol.* ***59***, 1147-1156.

84. Kang, K. W., Lee, S. J., Park, J. W., and Kim, S. G. (2002) Phosphatidylinositol 3-kinase regulates nuclear translocation of NF-E2-related factor 2 through actin rearrangement in response to oxidative stress, *Mol. Pharmacol.* ***62***, 1001-1010.

85. Lee, J. M., Hanson, J. M., Chu, W. A., and Johnson, J. A. (2001) Phosphatidylinositol 3-kinase, not extracellular signal-regulated kinase, regulates activation of the antioxidant-responsive element in IMR-32 human neuroblastoma cells, *J. Biol. Chem.* ***276***, 20011-20016.

86. Li, J., Lee, J. M., and Johnson, J. A. (2002) Microarray analysis reveals an antioxidant responsive element-driven gene set involved in conferring protection from an oxidative stress-induced apoptosis in IMR-32 cells, *J. Biol. Chem.* ***277***, 388-394.

87. Huang, H. C., Nguyen, T., and Pickett, C. B. (2000) Regulation of the antioxidant response element by protein kinase C-mediated phosphorylation of NF-E2-related factor 2, *Proc. Natl. Acad. Sci. U. S. A* ***97,*** 12475-12480.

88. Huang, H. C., Nguyen, T., and Pickett, C. B. (2002) Phosphorylation of Nrf2 at Ser-40 by protein kinase C regulates antioxidant response element-mediated transcription, *J. Biol. Chem.* ***277***, 42769-42774.

89. Bloom, D. A. and Jaiswal, A. K. (2003) Phosphorylation of Nrf2 at Ser40 by protein kinase C in response to antioxidants leads to the release of Nrf2 from INrf2, but is not required for Nrf2 stabilization/accumulation in the nucleus

and transcriptional activation of antioxidant response element-mediated NAD(P)H:quinone oxidoreductase-1 gene expression, *J. Biol. Chem.* ***278***, 44675-44682.

90. Kensler, T. W., Egner, P. A., Dolan, P. M., Groopman, J. D., and Roebuck, B. D. (1987) Mechanism of protection against aflatoxin tumorigenicity in rats fed 5-(2-pyrazinyl)-4-methyl-1,2-dithiol-3-thione (oltipraz) and related 1,2-dithiol-3-thiones and 1,2-dithiol-3-ones, *Cancer Res.* ***47***, 4271-4277.

91. Kensler, T. W., Groopman, J. D., Eaton, D. L., Curphey, T. J., and Roebuck, B. D. (1992) Potent inhibition of aflatoxin-induced hepatic tumorigenesis by the monofunctional enzyme inducer 1,2-dithiole-3-thione, *Carcinogenesis* ***13***, 95-100.

92. Rao, C. V., Rivenson, A., Katiwalla, M., Kelloff, G. J., and Reddy, B. S. (1993) Chemopreventive effect of oltipraz during different stages of experimental colon carcinogenesis induced by azoxymethane in male F344 rats, *Cancer Res.* ***53***, 2502-2506.

93. Wattenberg, L. W. and Bueding, E. (1986) Inhibitory effects of 5-(2-pyrazinyl)-4-methyl-1,2-dithiol-3-thione (Oltipraz) on carcinogenesis induced by benzo[a]pyrene, diethylnitrosamine and uracil mustard, *Carcinogenesis* ***7***, 1379-1381.

94. Kwak, M. K., Egner, P. A., Dolan, P. M., Ramos-Gomez, M., Groopman, J. D., Itoh, K., Yamamoto, M., and Kensler, T. W. (2001) Role of phase 2 enzyme induction in chemoprotection by dithiolethiones, *Mutat. Res.* ***480-481,*** 305-315.

95. Rushmore, T. H., King, R. G., Paulson, K. E., and Pickett, C. B. (1990) Regulation of glutathione S-transferase Ya subunit gene expression: identification of a unique xenobiotic-responsive element controlling inducible expression by planar aromatic compounds, *Proc. Natl. Acad. Sci. U. S. A* ***87***, 3826-3830.

96. Wattenberg, L. W. (1973) Inhibition of chemical carcinogen-induced pulmonary neoplasia by butylated hydroxyanisole, *J. Natl. Cancer Inst.* ***50***, 1541-1544.

97. Weisburger, E. K. and Evarts, R. P. (1977) Inhibitory effect of butylated hydroxytoluene (BHT) on intestinal carcinogenesis in rats by azoxymethane, *Food Cosmet. Toxicol.* ***15***, 139-141.

98. Gelboin, H. V., Miller, J. A., and Miller, E. C. (1958) Studies on hepatic protein-bound dye formation in rats given single large doses of 3'methyl-4-dimethylaminoazobenzene, *Cancer Res.* **18**, 608-617.

99. Skipper, P. L. and Tannenbaum, S. R. (1990) Protein adducts in the molecular dosimetry of chemical carcinogens, *Carcinogenesis* ***11***, 507-518.

100. Weisburger, E. K., Weisburger, J. H., and MORRIS, H. P. (1953) Studies on the metabolism of 2-acetylaminofluorene-9-C14, *Arch. Biochem. Biophys.* ***43***, 474-484.

101. WIEST, W. G. and Heidelberger, C. (1953) The interaction of carcinogenic hydrocarbons with tissue constituents. II. 1,2,5,6-Dibenzanthracene-9,10-C14 in skin, *Cancer Res.* ***13***, 250-254.

102. Pohl, L. R. (1993) An immunochemical approach of identifying and characterizing protein targets of toxic reactive metabolites, *Chem. Res. Toxicol.* ***6***, 786-793.

103. Farmer, P. B. (1994) Carcinogen adducts: use in diagnosis and risk assessment, *Clin. Chem.* ***40***, 1438-1443.

104. Melikian, A. A., Sun, P., Pierpont, C., Coleman, S., and Hecht, S. S. (1997) Gas chromatographic-mass spectrometric determination of benzo[a]pyrene and chrysene diol epoxide globin adducts in humans, *Cancer Epidemiol. Biomarkers Prev.* ***6***, 833-839.

105. Yeowell-O'Connell, K., Jin, Z., and Rappaport, S. M. (1996) Determination of albumin and hemoglobin adducts in workers exposed to styrene and styrene oxide, *Cancer Epidemiol. Biomarkers Prev.* ***5***, 205-215.

106. Liebler, D. C. (2001) *Introduction to Proteomics: Tools for the New Biology* Humana Press, Totowa, NJ.

107. Whitehouse, C. M., Dreyer, R. N., Yamashita, M., and Fenn, J. B. (1985) Electrospray interface for liquid chromatographs and mass spectrometers, *Anal. Chem.* ***57,*** 675-679.

108. Karas, M. and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons, *Anal. Chem.* ***60***, 2299-2301.

109. Gaskell, S. J. (1997) Electrospray: Principles and Practice, *J. Mass Spectrom.* **32** 677-688.

110. Yates, J. R. (1996) Protein structure analysis by mass spectrometry., *Methods. Enzymol.* ***271***, 351-377.

111. Clauser, K. R., Baker, P., and Burlingame, A. L. (1999) Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching, *Anal. Chem.* ***71***, 2871-2882.

112. Mann, M., Hojrup, P., and Roepstorff, P. (1993) Use of mass spectrometric molecular weight information to identify proteins in sequence databases, *Biol. Mass Spectrom.* ***22***, 338-345.

113. Pappin, D. J., Hojrup, P., and Bleasby, A. J. (1993) Rapid identification of proteins by peptide-mass fingerprinting, *Curr. Biol.* ***3***, 327-332.

114. Hunt, D. F., Bone W.M., and Shabanowitz, J. (1981) Sequence-Analysis of Oligopeptides by Secondary Ion-Collison Activated Dissociation Mass-Spectrometry, *Anal. Chem.* **53** 1704-1706.

115. Opiteck, G. J. and Jorgenson JW (1997) Two-dimensional SEC/RPLC coupled to mass spectrometry for the analysis of peptides, *Anal. Chem.* **69** 2283-2291.

116. Yates, J. R., Eng, J. K., and McCormack, A. L. (1995) Mining genomes: correlating tandem mass spectra of modified and unmodified peptides to sequences in nucleotide databases., *Anal. Chem.* ***67***, 3202-3210.

117. Yates, J. R., Eng, J. K., McCormack, A. L., and Schieltz, D. (1995) Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database., *Anal. Chem.* ***67***, 1426-1436.

118. Liebler, D. C., Hansen, B. T., Davey, S. W., Tiscareno, L., and Mason, D. E. (2002) Peptide sequence motif analysis of tandem MS data with the SALSA algorithm, *Anal. Chem.* ***74***, 203-210.

119. Hansen, B. T., Jones, J. A., Mason, D. E., and Liebler, D. C. (2001) SALSA: a pattern recognition algorithm to detect electrophile-adducted peptides by automated evaluation of CID spectra in LC-MS-MS analyses, *Anal. Chem.* ***73***, 1676-1683.

120. Hansen, B. T., Davey, S. W., Ham, A. J., and Liebler, D. C. (2005) P-Mod: An Algorithm and Software To Map Modifications To Peptide Sequences Using Tandem MS Data, *J. Proteome Res.* ***4***, 358-368.

121. Mason, D. E. and Liebler, D. C. (2000) Characterization of benzoquinone-peptide adducts by electrospray mass spectrometry, *Chem. Res Toxicol.* ***13***, 976-982.

122. Rice, R. H., Means, G. E., and Brown, W. D. (1977) Stabilization of bovine trypsin by reductive methylation, *Biochim. Biophys. Acta* ***492***, 316-321.

123. Sorensen, S. B., Sorensen, T. L., and Breddam, K. (1991) Fragmentation of proteins by S. aureus strain V8 protease. Ammonium bicarbonate strongly inhibits the enzyme but does not improve the selectivity for glutamic acid, *FEBS Lett.* ***294***, 195-197.

124. Houmard, J. and Drapeau, G. R. (1972) Staphylococcal protease: a proteolytic enzyme specific for glutamoyl bonds, *Proc. Natl. Acad. Sci. U. S. A* ***69***, 3506-3509.

125. Drapeau, G. R. (1978) The primary structure of staphylococcal protease, *Can. J. Biochem.* ***56***, 534-544.

126. Jekel, P. A., Weijer, W. J., and Beintema, J. J. (1983) Use of endoproteinase Lys-C from Lysobacter enzymogenes in protein sequence analysis, *Anal. Biochem.* ***134***, 347-354.

127. Drapeau, G. R. (1980) Substrate specificity of a proteolytic enzyme isolated from a mutant of Pseudomonas fragi, *J. Biol. Chem.* ***255***, 839-840.

128. Noreau, J. and Drapeau, G. R. (1979) Isolation and properties of the protease from the wild-type and mutant strains of Pseudomonas fragi, *J. Bacteriol.* ***140***, 911-916.

129. Mitchell, W. M. (1977) Cleavage at arginine residues by clostripain, *Methods Enzymol.* ***47***, 165-170.

130. Mitchell, W. M. and Harrington, W. F. (1968) Purification and properties of clostridiopeptidase B (Clostripain), *J. Biol. Chem.* ***243***, 4683-4692.

131. Hayashi, R. (1976) Carboxypeptidase Y, *Methods Enzymol.* ***45***, 568-587.

132. Hayashi, R., Moore, S., and Stein, W. H. (1973) Carboxypeptidase from yeast. Large scale preparation and the application to COOH-terminal analysis of peptides and proteins, *J. Biol. Chem.* ***248***, 2296-2302.

133. Cao, P. and Stults, J. T. (1999) Phosphopeptide analysis by on-line immobilized metal-ion affinity chromatography-capillary electrophoresis-electrospray ionization mass spectrometry, *J. Chromatogr. A* ***853***, 225-235.

134. Davis, M. T., Stahl, D. C., Hefta, S. A., and Lee, T. D. (1995) A microscale electrospray interface for on-line, capillary liquid chromatography/tandem mass spectrometry of complex peptide mixtures, *Anal. Chem.* ***67***, 4549-4556.

135. Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., III (1999) Direct analysis of protein complexes using mass spectrometry, *Nat. Biotechnol.* ***17***, 676-682.

136. Dhakshinamoorthy, S. and Jaiswal, A. K. (2001) Functional characterization and role of INrf2 in antioxidant response element-mediated expression and antioxidant induction of NAD(P)H:quinone oxidoreductase1 gene, *Oncogene* ***20***, 3906-3917.

137. Velichkova, M. and Hasson, T. (2005) Keap1 regulates the oxidation-sensitive shuttling of Nrf2 into and out of the nucleus via a Crm1-dependent nuclear export mechanism, *Mol. Cell Biol.* ***25***, 4501-4513.

138. Kensler, T. W. (1997) Chemoprevention by inducers of carcinogen detoxication enzymes, *Environ. Health Perspect.* ***105 Suppl 4***, 965-970.

139. Talalay, P. (2000) Chemoprotection against cancer by induction of phase 2 enzymes, *Biofactors* ***12***, 5-11.

140. Hayes, J. D. and McMahon, M. (2001) Molecular basis for the contribution of the antioxidant responsive element to cancer chemoprevention, *Cancer Lett.* ***174***, 103-113.

141. Nguyen, T., Sherratt, P. J., and Pickett, C. B. (2003) Regulatory mechanisms controlling gene expression mediated by the antioxidant response element, *Annu. Rev. Pharmacol. Toxicol.* ***43***, 233-260.

142. Dinkova-Kostova, A. T., Massiah, M. A., Bozak, R. E., Hicks, R. J., and Talalay, P. (2001) Potency of Michael reaction acceptors as inducers of enzymes that protect against carcinogenesis depends on their reactivity with sulfhydryl groups, *Proc. Natl. Acad. Sci. U. S. A* ***98***, 3404-3409.

143. Itoh, K., Igarashi, K., Hayashi, N., Nishizawa, M., and Yamamoto, M. (1995) Cloning and characterization of a novel erythroid cell-derived CNC family transcription factor heterodimerizing with the small Maf family proteins, *Mol. Cell Biol.* ***15***, 4184-4193.

144. Kang, M. I., Kobayashi, A., Wakabayashi, N., Kim, S. G., and Yamamoto, M. (2004) Scaffolding of Keap1 to the actin cytoskeleton controls the function of Nrf2 as key regulator of cytoprotective phase 2 genes, *Proc. Natl. Acad. Sci. U. S. A* ***101***, 2046-2051.

145. Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S. P. (2003) A proteomics approach to understanding protein ubiquitination, *Nat. Biotechnol.* ***21***, 921-926.

146. Zhang, D. D. and Hannink, M. (2003) Distinct cysteine residues in Keap1 are required for Keap1-dependent ubiquitination of Nrf2 and for stabilization of Nrf2 by chemopreventive agents and oxidative stress, *Mol. Cell Biol.* ***23***, 8137-8151.

147. Cope, G. A. and Deshaies, R. J. (2003) COP9 signalosome: a multifunctional regulator of SCF and other cullin-based ubiquitin ligases, *Cell* ***114***, 663-671.

148. Wolf, D. A., Zhou, C., and Wee, S. (2003) The COP9 signalosome: an assembly and maintenance platform for cullin ubiquitin ligases?, *Nat. Cell Biol.* ***5***, 1029-1033.

149. Barral, Y., Jentsch, S., and Mann, C. (1995) G1 cyclin turnover and nutrient uptake are controlled by a common pathway in yeast, *Genes Dev.* ***9***, 399-409.

150. Galan, J. M. and Peter, M. (1999) Ubiquitin-dependent degradation of multiple F-box proteins by an autocatalytic mechanism, *Proc. Natl. Acad. Sci. U. S. A* ***96***, 9124-9129.

151. Wirbelauer, C., Sutterluty, H., Blondel, M., Gstaiger, M., Peter, M., Reymond, F., and Krek, W. (2000) The F-box protein Skp2 is a ubiquitylation target of a Cul1-based core ubiquitin ligase complex: evidence for a role of Cul1 in the suppression of Skp2 expression in quiescent fibroblasts, *EMBO J.* ***19***, 5362-5375.

152. Geyer, R., Wee, S., Anderson, S., Yates, J., and Wolf, D. A. (2003) BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases, *Mol. Cell* ***12***, 783-790.

153. Block, G., Patterson, B., and Subar, A. (1992) Fruit, vegetables, and cancer prevention: a review of the epidemiological evidence, *Nutr. Cancer* ***18***, 1-29.

154. Michaud, D. S., Spiegelman, D., Clinton, S. K., Rimm, E. B., Curhan, G. C., Willett, W. C., and Giovannucci, E. L. (1999) Fluid intake and the risk of bladder cancer in men, *N. Engl. J. Med.* ***340***, 1390-1397.

155. Michaud, D. S., Spiegelman, D., Clinton, S. K., Rimm, E. B., Willett, W. C., and Giovannucci, E. L. (1999) Fruit and vegetable intake and incidence of bladder cancer in a male prospective cohort, *J. Natl. Cancer Inst.* ***91***, 605-613.

156. Verhoeven, D. T., Goldbohm, R. A., van, P. G., Verhagen, H., and Van den Brandt, P. A. (1996) Epidemiological studies on brassica vegetables and cancer risk, *Cancer Epidemiol. Biomarkers Prev.* ***5***, 733-748.

157. Fahey, J. W., Zhang, Y., and Talalay, P. (1997) Broccoli sprouts: an exceptionally rich source of inducers of enzymes that protect against chemical carcinogens, *Proc. Natl. Acad. Sci. U. S. A* ***94***, 10367-10372.

158. Prochaska, H. J., Santamaria, A. B., and Talalay, P. (1992) Rapid detection of inducers of enzymes that protect against carcinogens, *Proc. Natl. Acad. Sci. U. S. A* ***89***, 2394-2398.

159. Zhang, Y., Talalay, P., Cho, C. G., and Posner, G. H. (1992) A major inducer of anticarcinogenic protective enzymes from broccoli: Isolation and elucidation of structure., *Proc. Natl. Acad. Sci. USA* ***89***, 2399-2403.

160. Gerhauser, C., You, M., Liu, J., Moriarty, R. M., Hawthorne, M., Mehta, R. G., Moon, R. C., and Pezzuto, J. M. (1997) Cancer chemopreventive potential of sulforamate, a novel analogue of sulforaphane that induces phase 2 drug-metabolizing enzymes, *Cancer Res.* ***57***, 272-278.

161. Chung, F. L., Conaway, C. C., Rao, C. V., and Reddy, B. S. (2000) Chemoprevention of colonic aberrant crypt foci in Fischer rats by sulforaphane and phenethyl isothiocyanate, *Carcinogenesis* ***21***, 2287-2291.

162. Rushmore, T. H., Morton, M. R., and Pickett, C. B. (1991) The antioxidant response element. Activation by oxidative stress and identification of the DNA concensus sequence required for functional activity., *J. Biol. Chem.* ***266***, 11632-11639.

163. Favreau, L. V. and Pickett, C. B. (1991) Transcriptional regulation of the rat NAD(P)H:quinone reductase gene. Identification of regulatory elements controlling basal level expression and inducible expression by planar aromatic compounds and phenolic antioxidants, *J. Biol. Chem.* ***266***, 4556-4561.

164. Keum, Y. S., Owuor, E. D., Kim, B. R., Hu, R., and Kong, A. N. (2003) Involvement of Nrf2 and JNK1 in the activation of antioxidant responsive element (ARE) by chemopreventive agent phenethyl isothiocyanate (PEITC), *Pharm. Res.* ***20***, 1351-1356.

165. Hong, F., Sekhar, K. R., Freeman, M. L., and Liebler, D. C. (2005) Specific patterns of electrophile adduction trigger Keap1 ubiquitination and Nrf2 activation, *J. Biol. Chem. in press*.

166. McWalter, G. K., Higgins, L. G., McLellan, L. I., Henderson, C. J., Song, L., Thornalley, P. J., Itoh, K., Yamamoto, M., and Hayes, J. D. (2004) Transcription factor Nrf2 is essential for induction of NAD(P)H:quinone oxidoreductase 1, glutathione S-transferases, and glutamate cysteine ligase by broccoli seeds and isothiocyanates, *J. Nutr.* ***134***, 3499S-3506S.

167. Gao, X. and Talalay, P. (2004) Induction of phase 2 genes by sulforaphane protects retinal pigment epithelial cells against photooxidative damage, *Proc. Natl. Acad. Sci. U. S. A* ***101***, 10446-10451.

168. Kraft, A. D., Johnson, D. A., and Johnson, J. A. (2004) Nuclear factor E2-related factor 2-dependent antioxidant response element activation by tert-butylhydroquinone and sulforaphane occurring preferentially in astrocytes conditions neurons against oxidative insult, *J. Neurosci.* ***24***, 1101-1112.

169. Jones, J. A. and Liebler, D. C. (2000) Tandem MS Analysis of Model Peptide Adducts from Reactive Metabolites of the Hepatotoxin 1,1-Dichloroethylene, *Chem. Res Toxicol.* ***13***, 1302-1312.

170. Hecht, S. S. (2000) Inhibition of carcinogenesis by isothiocyanates, *Drug Metab Rev.* ***32***, 395-411.

171. Dinkova-Kostova, A. T., Holtzclaw, W. D., and Wakabayashi, N. (2005) Keap1, the sensor for electrophiles and oxidants that regulates the phase 2 response, is a zinc metalloprotein, *Biochem.* ***44***, 6889-6899.

172. Numazawa, S., Ishikawa, M., Yoshida, A., Tanaka, S., and Yoshida, T. (2003) Atypical protein kinase C mediates activation of NF-E2-related factor 2 in response to oxidative stress, *Am. J. Physiol Cell Physiol* ***285***, C334-C342.

173. Numazawa, S. and Yoshida, T. (2004) Nrf2-dependent gene expressions: a molecular toxicological aspect, *J. Toxicol. Sci.* ***29***, 81-89.

174. Lee, S. R., Kwon, K. S., Kim, S. R., and Rhee, S. G. (1998) Reversible inactivation of protein-tyrosine phosphatase 1B in A431 cells stimulated with epidermal growth factor, *J. Biol. Chem.* ***273***, 15366-15372.

175. Levonen, A. L., Landar, A., Ramachandran, A., Ceasar, E. K., Dickinson, D. A., Zanon, G., Morrow, J. D., and Darley-Usmar, V. M. (2004) Cellular mechanisms of redox cell signaling: role of cysteine modification in controlling antioxidant defenses in response to electrophilic lipid oxidation products, *Biochem. J.* ***378***, 373-382.

176. Braun, S., Hanselmann, C., Gassmann, M. G., auf dem, K. U., Born-Berclaz, C., Chan, K., Kan, Y. W., and Werner, S. (2002) Nrf2 transcription factor, a

novel target of keratinocyte growth factor action which regulates gene expression and inflammation in the healing skin wound, *Mol. Cell Biol.* ***22***, 5492-5505.

177. Dhakshinamoorthy, S. and Porter, A. G. (2004) Nitric oxide-induced transcriptional up-regulation of protective genes by Nrf2 via the antioxidant response element counteracts apoptosis of neuroblastoma cells, *J. Biol. Chem.* ***279***, 20096-20107.