

Center on Proteolytic Pathways

TRP 4: Proteolysis MAP

Fundamental Questions:

- 1. What is the magnitude of a proteolytic event?*
- 2. Where is the event localized?*
- 3. What are the substrates?*
- 4. Which proteases are involved?*
- 5. What products are produced?*
- 6. What pathways precede and follow a proteolytic event?*

Proteolytic signature = what is cleaved?



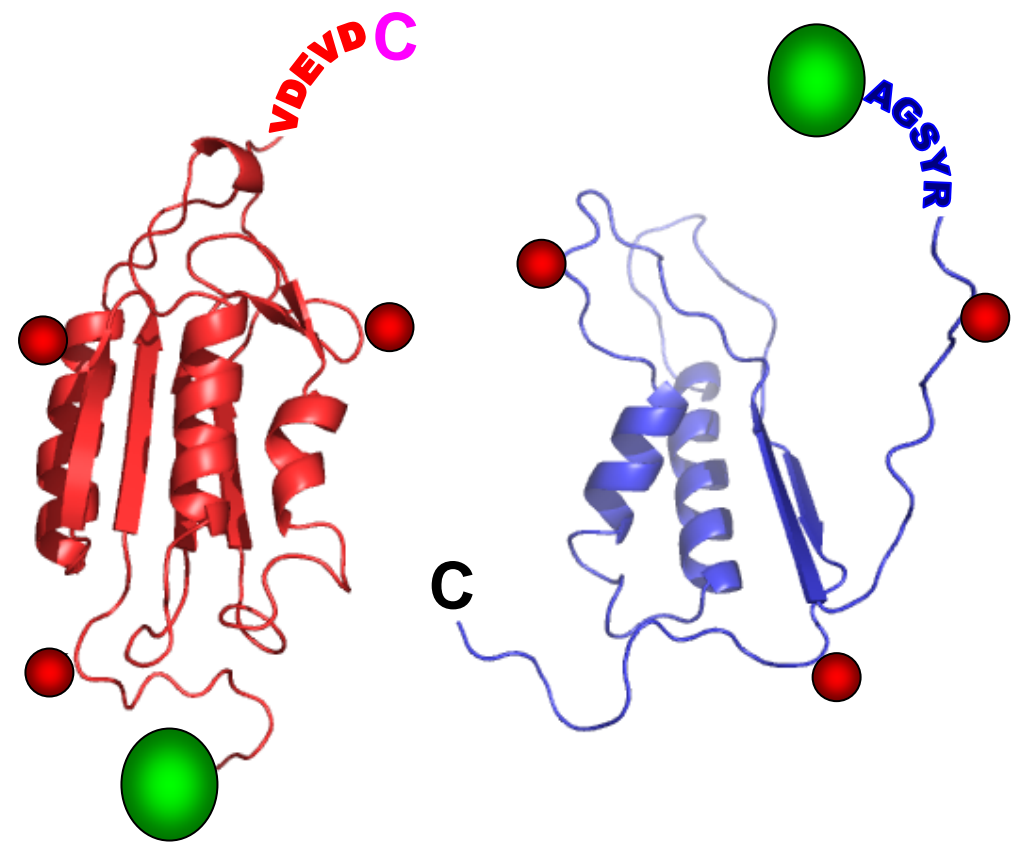
**Reported cleavage
Predicted motif**

Protease



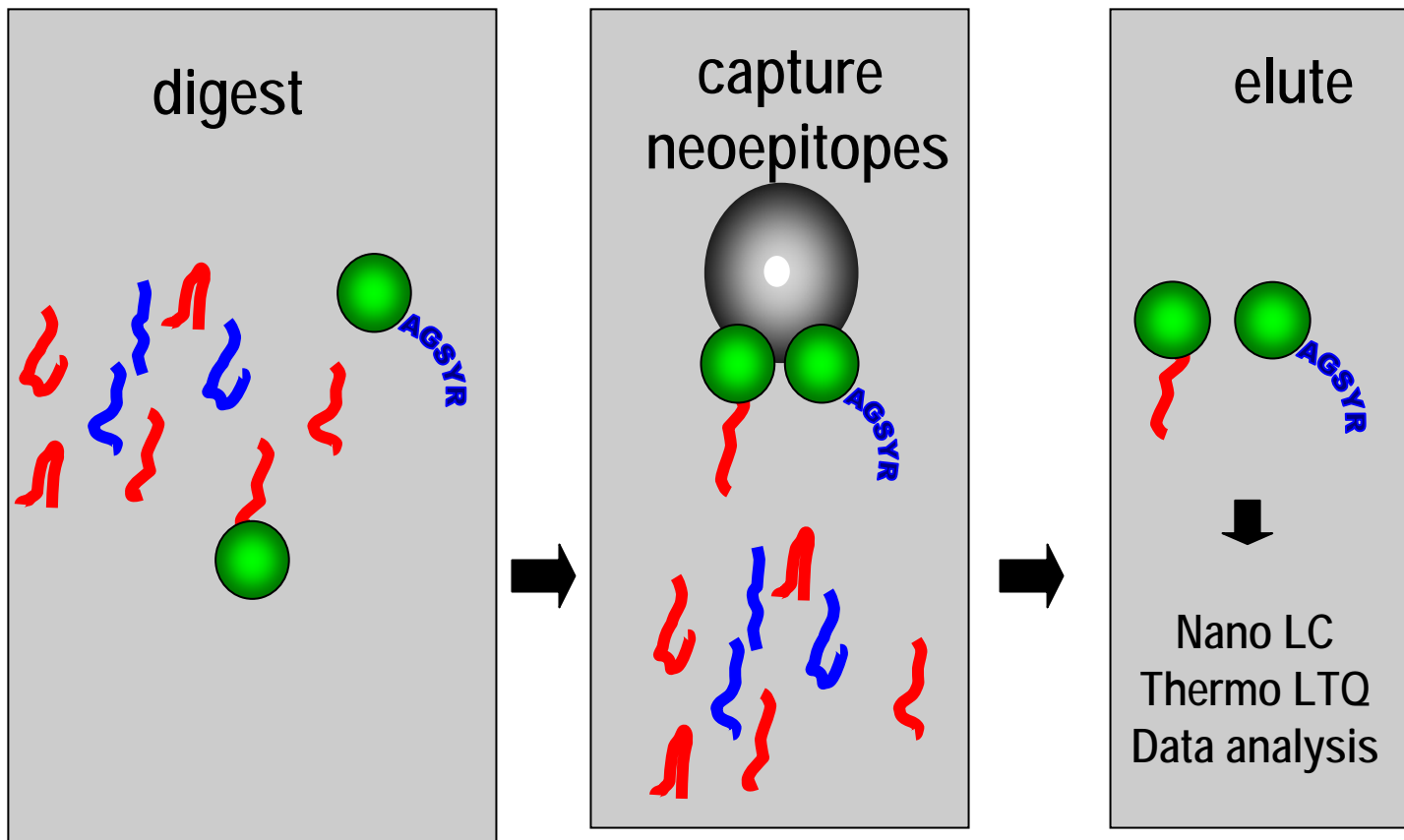
**Specific inhibitors
Activity-based probes
Follow the proteolytic cascades**

Protease pathways



●	Block Lys side-chains	●	Affinity label N-terminus
---	-----------------------	---	---------------------------

One N-terminal only for most proteins
 2-3 N-terminals for internally processed proteins



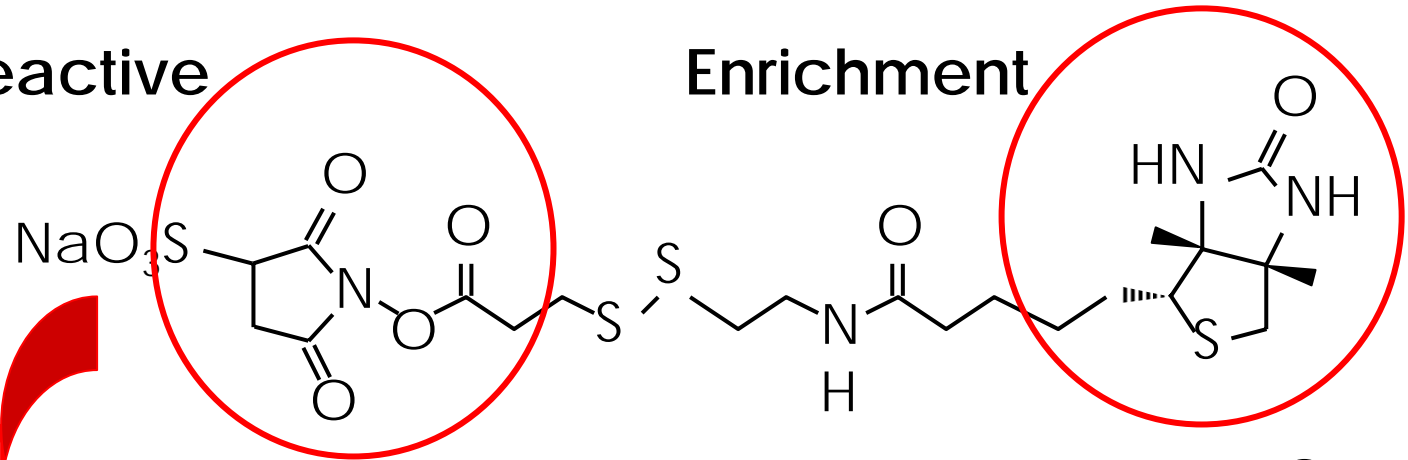
Select and enrich for N-terminal peptides
Simplified proteome!

N-terminal affinity labeling

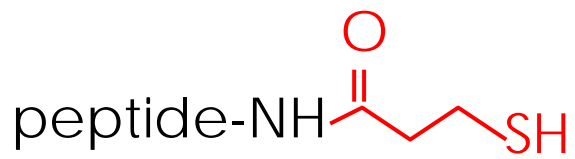
Amine-reactive

Enrichment

protein-NH₂



digestion
avidin capture
thiol reduction



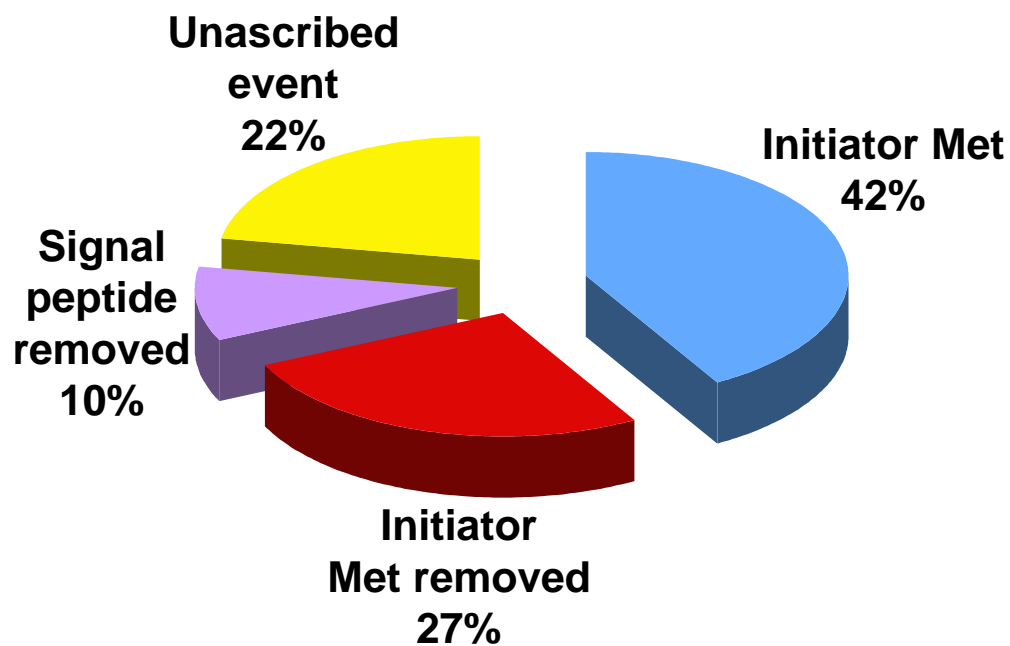
How do we know it's real?

- *E. coli* model system
- Efficient guanidination procedure
-specific N-terminal labeling
- Search only for labeled N-terminals
- SEQUEST Probability score cut-off 0.8
- SEQUEST Xcorr cut-off 2.0

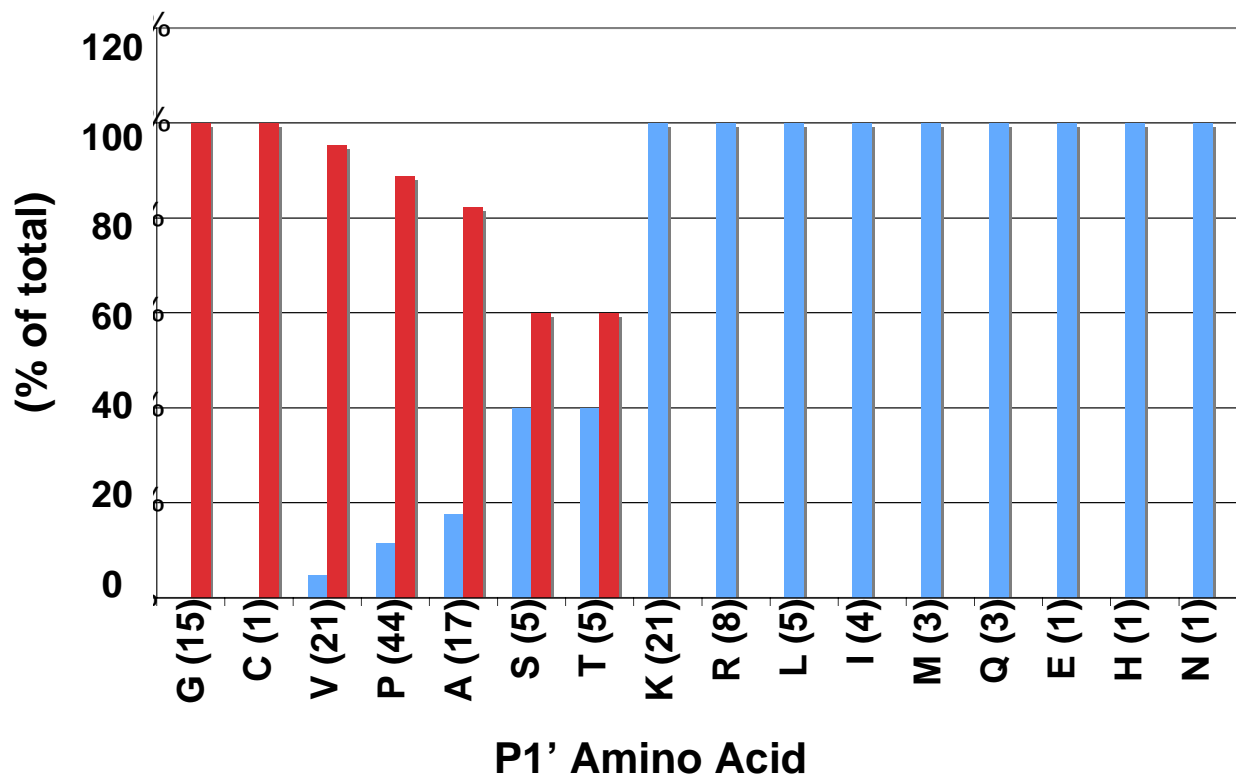
Proteolytic products in soluble *E. coli* proteins

SwissProt ID	Peptide sequence	Spectral counts	Cut site	Obs P1	Annotation	Obs enzyme	Spectral counts/enzyme
ALF1_ECOLI	-.TDIAQLLGK#.D	50	(N-term)-TDIA	0	MET REMOVED	Trypsin	50
ALKH_ECOLI	-.MK#NWK#TSAE.S	12	(N-term)-MKNW	0	MET INTACT	Trypsin, Glu-C	9, 3
AMPC_ECOLI	A.APOQINDIVHR.T	3	STFA-APOQ	19	SIGNAL (Annotated)	Trypsin	3
AMPM_ECOLI	M.AISIK#TPE.D	13	(N-term)M-AISI	1	MET REMOVED	Glu-C, Trypsin	1, 12
AMPN_ECOLI	L.SLPVHVADAFR.A	17	GOPL-SLPV	588	UNASCRIBED	Trypsin	17
ARNT_ECOLI	M.PLLFFSVAK#.G	2	WTIM-PLL	300	UNASCRIBED	Trypsin	2
AROG_ECOLI	-.MNYQNDLRIK#E.I	8	(N-term)-MNYQ	0	MET INTACT	Trypsin, Glu-C	2, 6
ASNA_ECOLI	R.LSPLHSVYVDQWDWER.V	3	DEDR-LSPL	105	UNASCRIBED	Trypsin	3
ASNA_ECOLI	-.MK#TAYIAK#QR.Q	6	(N-term)-MKTA	0	MET INTACT	Trypsin	6
ASPG1_ECOLI	-.MQK#K#SIYVAYTGGTIGMORSE.Q	2	(N-term)-MQKK	0	MET INTACT	Glu-C	2
ATCU_ECOLI	-.SQTIDLTL DGLSCGHCVK#.R	9	(N-term)-SQT	0	MET REMOVED	Trypsin	9
ATP6_ECOLI	M.ASENMTPODYIGHHLNQLDLR.T	7	(N-term)M-ASEN	1	MET REMOVED	Trypsin	7
ATPA_ECOLI	R.IAQFNVSEAHNEGTVSVSDGVIR.I	2	IKQR-IAQF	15	UNASCRIBED	Trypsin	2
ATPA_ECOLI	-.MQLNSTEISELIK#.Q	9	(N-term)-MQLN	0	MET INTACT	Trypsin	9
BCCP_ECOLI	E.ISGHIVRSPMVGTFYRTPSPD.A	2	AAAE-ISGH	77	UNASCRIBED	Glu-C	2
BCP_ECOLI	-.MNPLK#AGDIAPK#.F	23	(N-term)-MNPL	0	MET INTACT	Trypsin	23
BFR_ECOLI	R.SDLALELDGAK#NLR.E	3	EMLR-SDLA	88	UNASCRIBED	Trypsin	3
BFR_ECOLI	-.MK#GDTK#VINYLK#LLGNE.L	6	(N-term)-MKGD	0	MET INTACT	Trypsin, Glu-C	2, 4
BFR_ECOLI	R.EAIGYADSVHDYVSR.D	8	KNLR-EAIG	102	UNASCRIBED	Trypsin	8
BGLX_ECOLI	A.DDLFGNHPLTPEAR.D	4	PALA-DDLF	20	SIGNAL (Potential)	Trypsin	4
CAN_ECOLI	-.MK#DIDTLISNNALWSK#.M	3	(N-term)-MKDI	0	MET INTACT	Trypsin	3

Proteolytic products in soluble *E. coli* proteins



Removal of Met in *E. coli*



Confirm the *in vivo* activity of eMAP

Signal Peptide Processing

Secreted or periplasmic proteins have signal peptides

Signal peptides contain:

Positively charged N-terminus

Hydrophobic stretch

A-X-A motif (cleaved after last A)

E. coli periplasmic proteins - signal peptide removal

A. **AERPTLPIPDLLTTDAR**.N

CUEO_ECOLI Blue copper oxidase precursor.

MQRRDFLKYSVALGVASALPLWSRAVFA**AERPTLPIPDLLTTDAR**NRIQLTIGAGQSTFGGKTATTWGYN
 GNLLGPAVKLQRGKAVTVDIYNQLTEETTLHWHGLEVPGEVDGGPQGIIPPGGKRSVTLNVDQPAATCWF
 HPHQHGKTGRQVAMGLAGLVVIEDDEILKLMLPKQWGIDDVPVIVQDKKFSADGQIDYQLDVMATAAVGWF
 GDTLLTNGAIYPQHAAPRGWLRRLRLLNGCNARSLNFATSDNRPLYVIASDGGLLPEPVKVSELPVLMGER
 FEVLVEVNDNKPFDLVTLTPVSQMGMAIAPFDKPHPVMRIQPIAISASGALPDTLSSLPALPSLEGLTVRK
 LQLSMDPMLDMMGMQMLMEKYGDQAMAGMDHSQMMGHMGMHGNMNMNHGGKFDFFHANKNGQAFDMNKPM
 FAAAKGQYERWVISGVGDMMLHPFHIGTQFRILSENGKPPAAHRAGWKDTVKVEGNVSEVLVKFNHNDAP
 KEHAYMAHCHLLEHEDTGMMLGFTV

A. **AEVPSGTVLAEKQELVR**.H

MPPA_ECOLI Periplasmic murein peptide-binding protein precursor.

MKHSVSVTCCALLVSSISLSYA**AEVPSGTVLAEKQELVR**HIKDEPASLDPKAVGLPEIQVIRDLFEGLV
 NQNEKGEIVPGVATQWKSNDNRIWTFTLRDNKADGTPVTAQDFVYSWQRLVDPKTLSPFAWFAALAGI
 NNAQAIIDGKATPDQLGVTAVDAHTLKIQLDKPLPWFVNLNANFAFFPVQKANVESGKEWTKPGNLIGNG
 AYVLKERVVNEKLVVVPNTHYWDNAKTVLQKVTFLPINQESAATKRYLAGDIDITESFPKNMYQKLLKDI
 PGQVYTPPQLGTYYYAFNTQKGPTADQVRRLALSMTIDRRLMTEKVLGTGEKPAWHFTPDVDTAGFTPEPS
 PFEQMSQEELNAQAKTLLSAAGYGPQKPLKLTLLYNTSENHQKIAIAVASMWKKNLGVDVLQNQEWKTYI
 DSRNTGNFDVIRASWVG DYNEPSTFLTLLTSTHSGNISRFNNPAYDKVLAQASTENTVKARNADYNAAEK
 ILMEQAPIAPIYQYTNGRLIKPWLKGYPINNPEDVAYSRTMYIVKH

E.coli model system

Efficient blocking of Lys side chains

Low incidence of side reactions (0.5-3% labeled)

Detect what is known:

Validate previously suggested MAP specificity

Observe processing of known signal peptides

Unascribed events

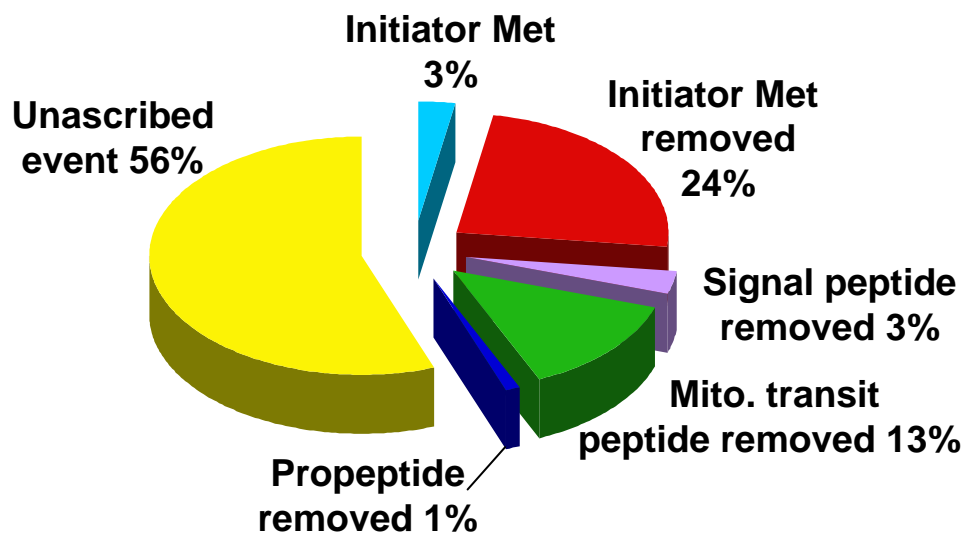
- Trimming of N-terminal
- Observation of predicted cleavage site
- Unknown proteolytic cleavage site

SwissProt ID	Peptide sequence	Spectral counts	Cut site	Obs P1	Annotation	Obs enzyme	Spectral counts/enzyme
6PGL_ECOLI	E.SQQIHVVWNLNHE.G	28	ASPE-SQQI	11	UNASCRIBED	Trypsin, Glu-C	9, 19
ACCC_ECOLI	L.SIK#QEEVHVR.G	1	GOPL-SIKO	321	UNASCRIBED	Trypsin	1
ACEA_ECOLI	E.SK#K#GYINSLGALTGGQALQQAK#AGIE.A	26	LHGE-SKKG	60	UNASCRIBED	Trypsin, Glu-C	15, 11
ADHE_ECOLI	Y.SSGK#PAIVGAGNTRVWIDETADIK#FLA	10	KAAV-SSGK	202	UNASCRIBED	Trypsin	10
ADHE_ECOLI	R.AVASVLM#SK#TF.D	1	DIKR-AVAS	228	UNASCRIBED	Trypsin	1

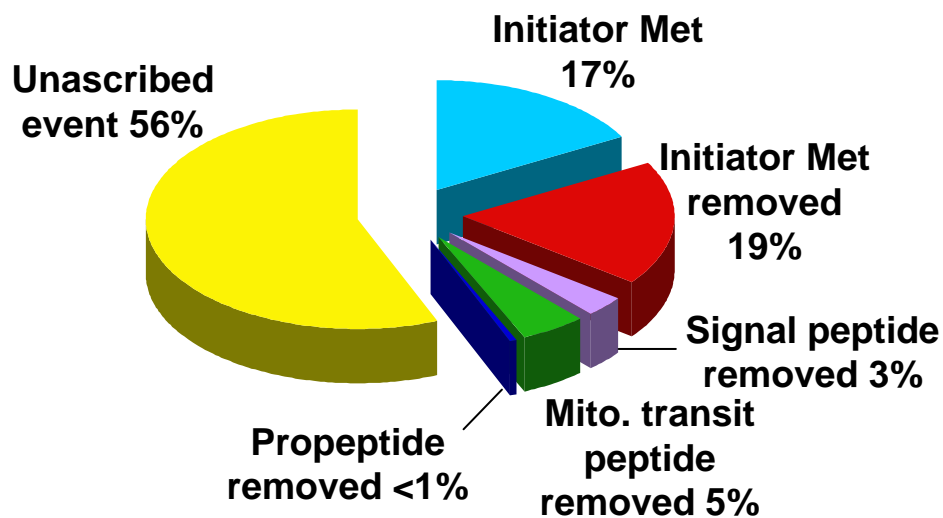
ABUNDANT PROTEIN? site
DIFFERENTIAL LABELING WITH LIGHT AND HEAVY ISOTOPES

- which protease match substrate specificity?
 • Side chain localization CHECK, Protic-DB
 • Proteolysis, unblocked Lys residue

Mouse tissue



Human 293A cells



Mitochondrial transit peptides

MOUSE

SwissProt ID	Peptide sequence	Spectral counts	Cut site	Obs.	Ann.	Description
KBL_MOUSE	A.HSAALQGLI	1	GLMA-AAL	11		NOVEL
THIL_MOUSE	Y.ASK#PTLNE.V	46	ERSY-ASKP	30		ANNOTATED
THIL_MOUSE	V.SA#TTLN#QVTR	1	KLPA-AAL	1		TRANSIT + TRIMMING
ALDH2_MOUSE	L.SAAATSAPVAPNHQPE.V	15	SRLL-SAAA	19		ANNOTATED
ALDH2_MOUSE	L.AATTAAPVAPNHQPE.V	2	PLAS-MAIT	30	19	TRANSIT + TRIMMING
ATPA_MOUSE	L.QK#TGTAEMSSILEER.I	18	NTRL-QKTG	43		ANNOTATED
ATPA_MOUSE	Q.K#TGTAEMSSILEER.I	12	TRLQ-KTGT	44	43	TRANSIT + TRIMMING
CISY_MOUSE	H.ASASSTNLK#DVLSNLIPIK#EQAR.I	5	AARH-ASAS	25	27	NEW TRANSIT
CP27A_MOUSE	A.K#ATIPAAQAGESTEGPETSQDPRRL	7	GAPL-KNT	30	19	NEW TRANSIT
ETHE1_MOUSE	L.SQQSASGAPVLLR.Q	1	GRRL-SQQS	13	21	NEW TRANSIT
GCDH_MOUSE	V.SSAAPVAPNHQPE.V	1	PLAS-SAAA	26	44	NEW TRANSIT
HMGCL_MOUSE +1	V.STSSMGTLPK#QVK#.I	1	LRAV-STSS	21	27	NEW TRANSIT
NDUB8_MOUSE	A.FHMTK#DMLRGSYPR.T	1	ARRA-FHMT	29	28	TRANSIT + TRIMMING
MAON_MOUSE	T.LEERLQLGIHGLIPPCFLSQDVQLLR.I	1	MAFT-LEER	66		TRANSIT @ 50 + MIP @ 58 + MIP
OTC_MOUSE	H.FMLCK#PQQVQLK#GR.	2	MLKQ-FWCG	24	32	NEW TRANSIT
SARDH_MOUSE	L.ATAPRTEK#SYDPYQR.T	4	LOPL-AIEA	38		NOVEL TRANSIT @ 20 + MIP
DHSA_MOUSE	K.ASAK#VSDAISTQYPVVDHE.F	5	ENKK-ASAK	42	43	NEW TRANSIT
DHSA_MOUSE	L.SAK#VSDAISTQYPVVDHE.F	1	NKRA-SAKV	43	43	ANNOTATED
DHSA_MOUSE	S.AK#VSDAISTQYPVVDHE.F	1	KKAS-AKVS	44	43	TRANSIT + TRIMMING
PRDX3_MOUSE	H.TPAVTQHAPYFK#.G	1	SSFH-TPAV	62	63	NEW TRANSIT

Mitochondrial processing peptidase (MPP)

Cleaves upon translocation to mito matrix

Arg in P3 or P2, P1' usually aromatic aa

Mitochondria intermediate peptidase (MIP)

Remove eight N-terminal residues.

Inner membrane protease 1 and 2 (IMP1, IMP2)

N-terminal processing upon translocation from matrix to intermembrane space

MPP:

Arg in P3 or P2, P1' usually aromatic aa

Mitochondria intermediate peptidase (MIP)

Remove eight N-terminal residues.

HUMAN

SwissProt ID	Peptide sequence	Spectral counts	Cut site	Obs. P1	Ann. P1	Description
ACPM_HUMAN	M.PPLTLEGIQDR.V	1	YS DM-PPLT	71	68	TRANSIT + TRIMMING
AASS_HUMAN	G.LHHK#AVLAVR.R	1	LSKG-LHHK	19	32	NEW TRANSIT
ATPO_HUMAN	P.FAK#LVRPPVQVYGI EGR.Y	1	VVRP-FAKL	23	23	ANNOTATED
ATPO_HUMAN	F.AK#LVRPPVQVYGI EGR.Y	2	VRPF-AKLV	24	23	TRANSIT + TRIMMING
CABC1_HUMAN +3	F.FHQDQSPVGG LTAEDIEK#AR.Q	3	QRRF-FHQD	162		NEW TRANSIT
CISY_HUMAN	H.ASASSTNLK#DILADLIPK#.E	28	AARH-ASAS	25	27	NEW TRANSIT
ECH1_HUMAN	L.TGSSAQEEASGVALGEAPDHSYESLR.V	1	SLRL-TGSS	33		NOVEL
DNJA3_HUMAN	A.K#EDYYQILGVPR.N	17	APLA-KEDY	90		NOVEL TRANSIT @ 74 + MIP @ 82 + MIP
EFTS_HUMAN	S.ASASSK#ELLMK#LR.R	2	PRLS-ASAS	41	45	NEW TRANSIT
PRDX5_HUMAN	M.APIK#VGDAIPAVEVFEGEPGNK#VNLAE LFK#.G	5	AAAM-APIK	53		NOVEL TRANSIT @ 45 + MIP
DHSA_HUMAN +1	G.FHFTVDGNK#R.A	2	GTRG-FHFT	32	43	NEW TRANSIT
ODPA_HUMAN +3	L.SGASQKPASRVLVASRN FAND.A	1	SRVL-SGAS	12	29	NEW TRANSIT

MPP:

Arg in P3 or P2, P1' usually aromatic aa

Mitochondria intermediate peptidase (MIP)

Remove eight N-terminal residues.

- Detect a **proteolytic signature** in *E. coli*, yeast, mouse and human samples
- Confirm *in vivo* cleavage events of many biological substrates
- Validate and / or correct predicted protease cleavage sites
- Assemble database of observed and / or previously unobserved protease substrates

Challenges for the Future

Should we use one-hit wonders?

Different cut-offs?

How many peptides do we need to
define a proteolytic signature?

Validation of unascribed events

Quantitation

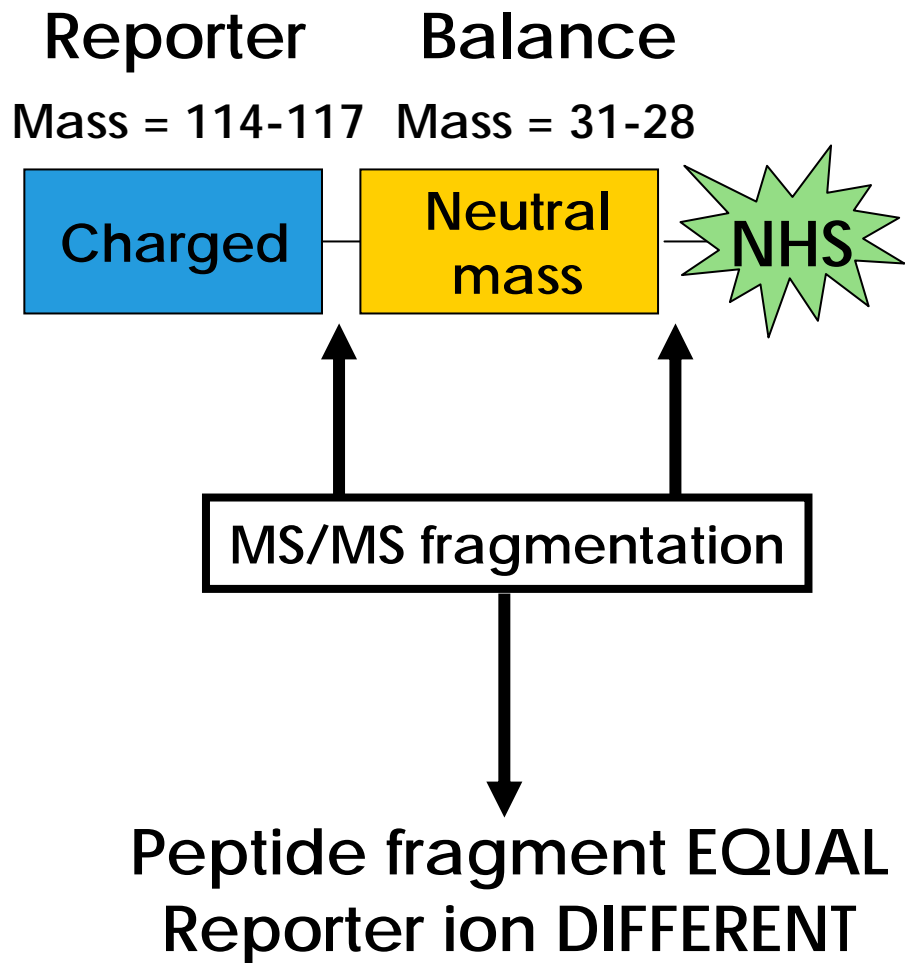
- Spectral counts

- Isotope coding (differential labeling of parallel samples)

 - cleavable biotin-tag

 - iTRAQ

iTRAQ - Isobaric tag (total mass=145)

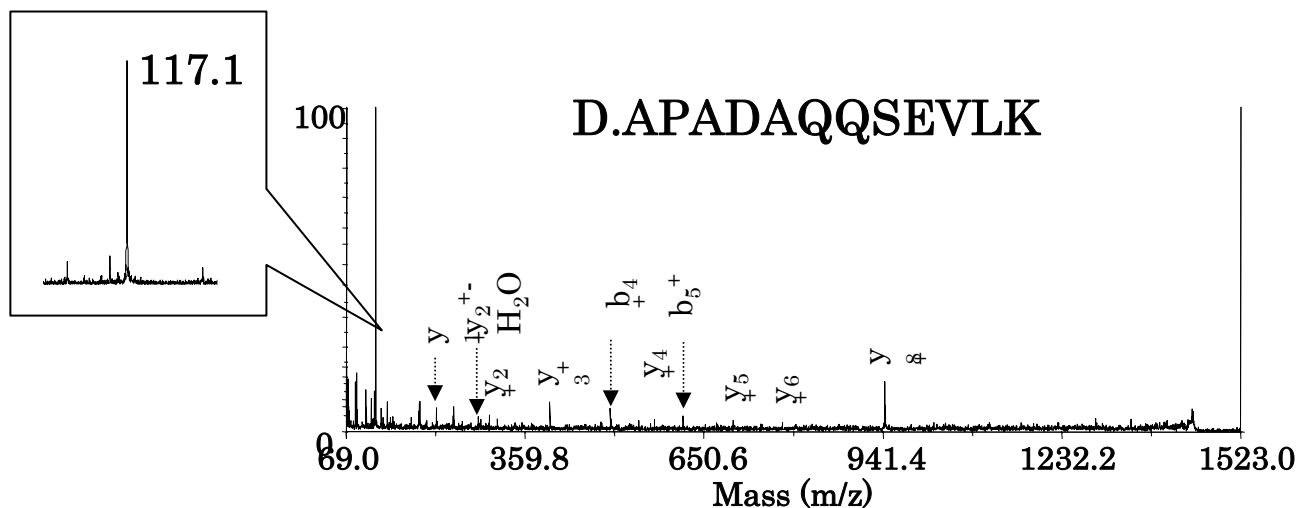
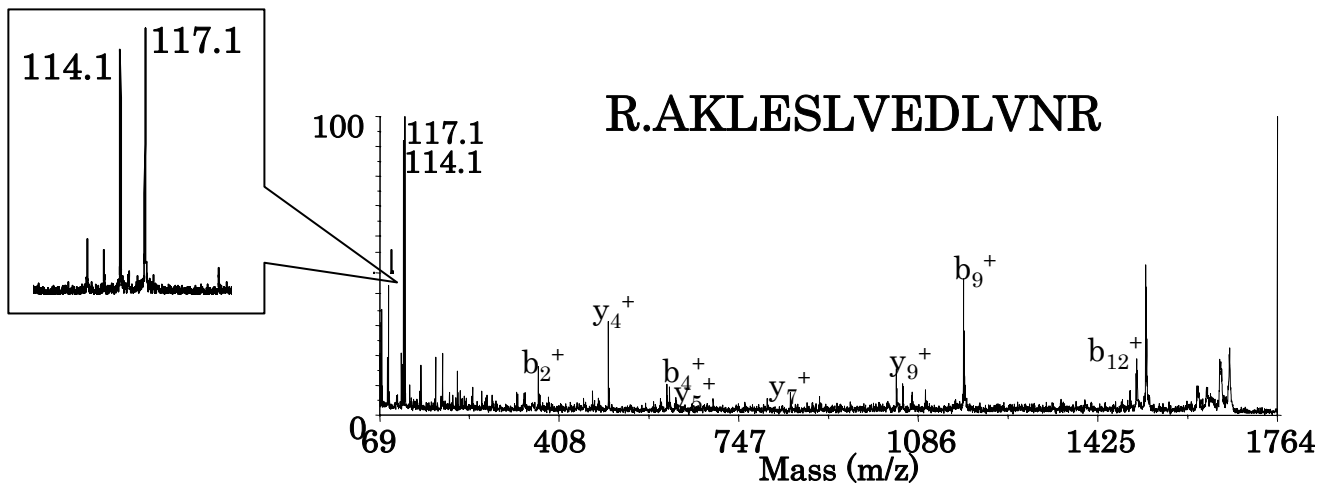


1. Denature/reduce/alkylate
2. Block Lys side chains
3. Label with heavy and light iTRAQ
4. Mix and digest
5. MS SURVEY
6. Data-dependent MS/MS:

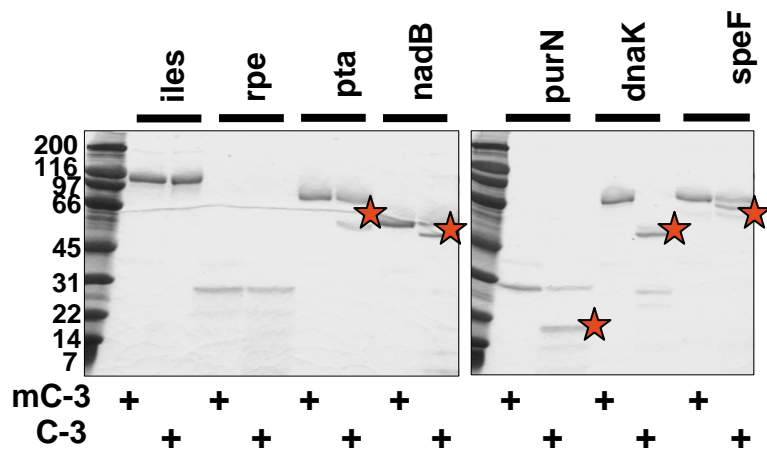
114 and 117 signature ions
 - EQUAL OR
 DIFFERENTIAL
 LABELING?

Data-dependant MS/MS

Equal distribution of peptides vs. sample- specific labeling



Caspase-3 cleavage of *E. coli* proteins

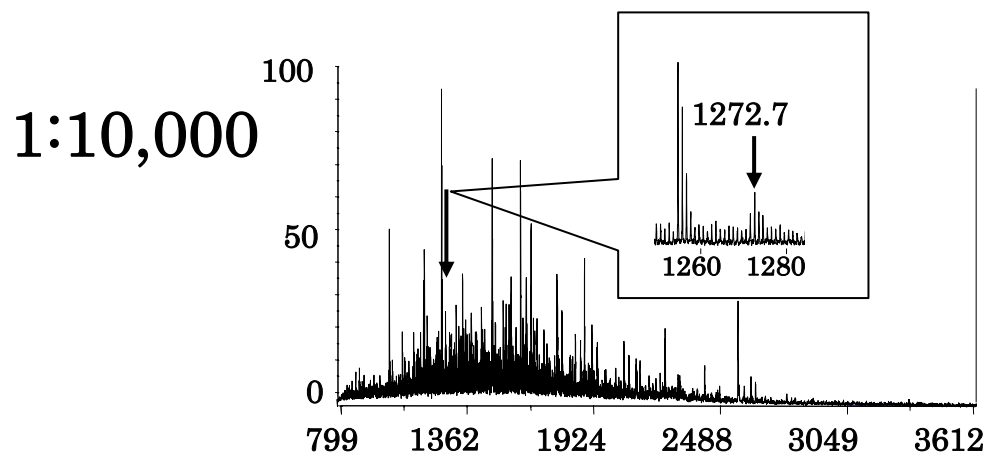


Protein	Full-length (kDa)	Cleavage fragments (kDa)		Putative cleavage site
iles	105	45	59	DEAD.QYVK
rpe	26	5	21	LEVD.GGVK
★ pta	79	22	57	DEVD.GLVS
★ nadB	62	54	8	DETD.SIDS
★ purN	25	7	17	DELD.GGPV
★ dnaK	70	46	24	DEVD.GEKT
★ speF	84	77	7	DEID.ATGY

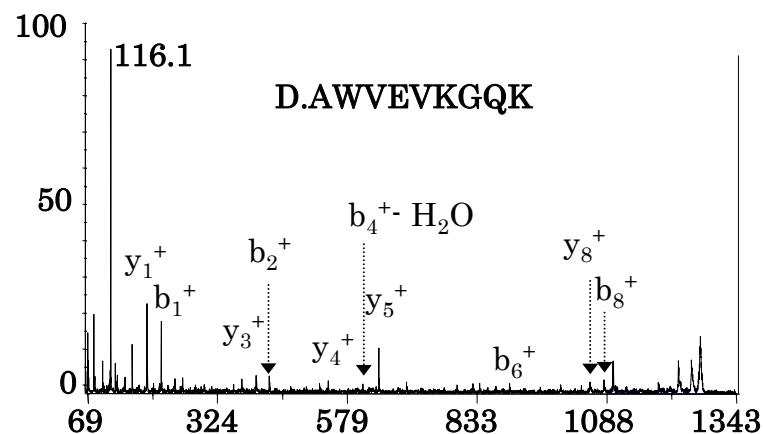
iTRAQ-labeled peptides

Protein	Peptide Sequence	Ratio (117/114)	Detection by SDS-PAGE
dnaK	DNGD.AWVEVK.G	>50	ND
	DNGD.AWVEVKGQK.M	>50	ND
	YTQD.GETLVGQPAKR.Q	>50	ND
	YGKK.VAEFFGKEPR.K	1.3	-
	DVVD.AEFEEVKDKK.-	>50	ND
	KVTR.AKLESLVEDLVNR.S	1.2	-
	ALQD.AGLSVSDIDDVILVGGQTR.M	>50	yes
	DEVD.GEKTFEVLATNGDTHLGGEDFDSR.L	>50	yes
ileS	DYND.APADAQQSEVLK.G	>50	ND
nadB	DETD.SIDSHVEDTLIAGAGICDR.H	>50	yes
pta	LMID.GPLQYDAAVMADVAK.S	>50	ND
purN	DELD.GGPVILQAK.V	>50	yes
speF	DEID.ATGYGIPVFIATENQER.V	>50	yes
rpe	ND	-	-

Caspase-3 cleaved dnaK spiked into 293 cell lysate



Data-dependent MS/MS



SUMMARY

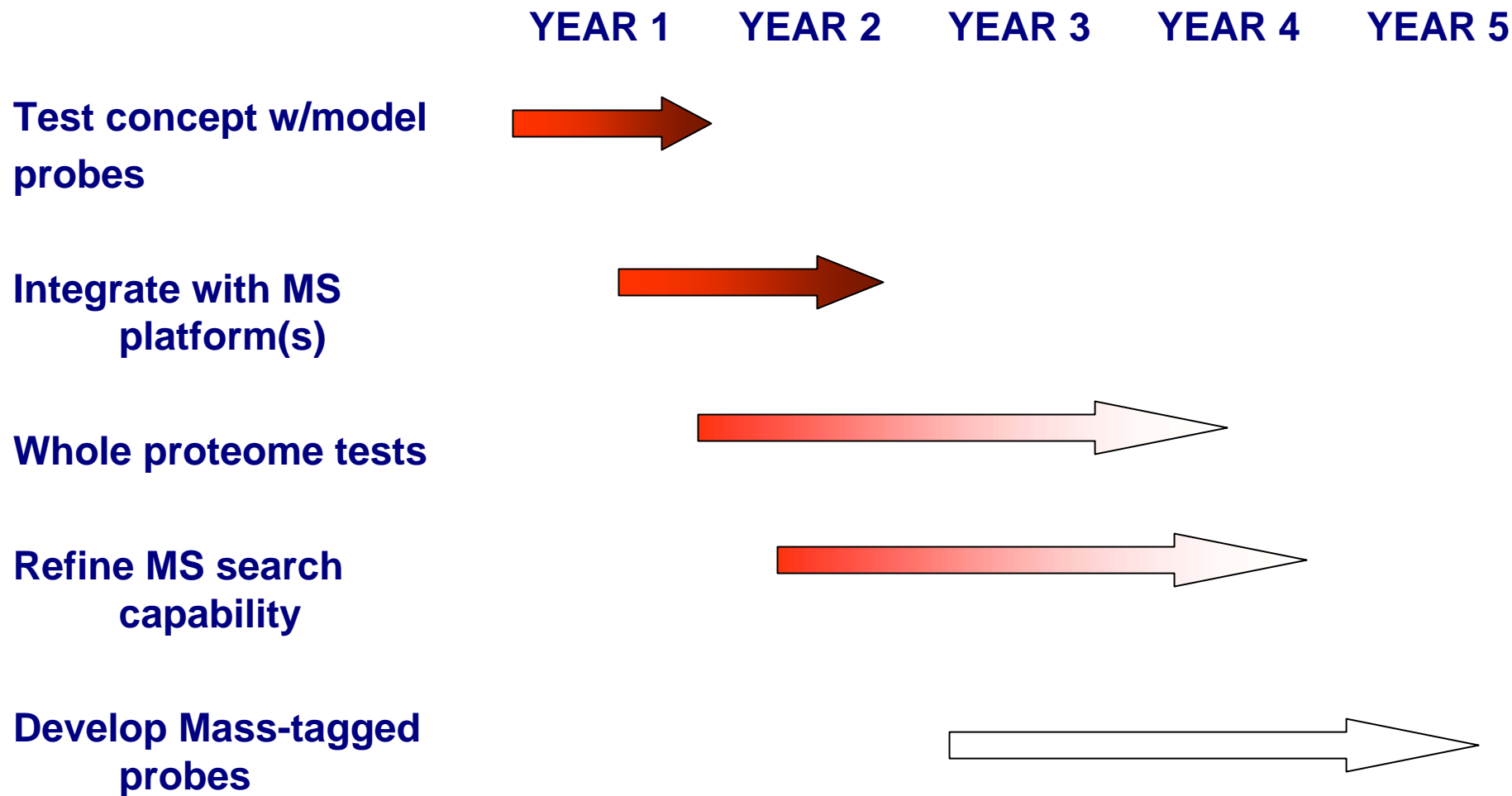
Chemistry / sample prep works: block Lys-side chains, allowing specific N-terminal modification

Use two systems of N-terminal modification for detection of a proteolytic signature:

	+	-
NHS-SS-Biotin	proteome simplifi cation	NO quantitation
iTRAQ	quantitation	NO proteome simplification

TRP 3: Product Terminal Isotope Coding

Guy Salvesen, The Burnham Institute



Acknowledgements

Guy Salvesen

John Timmer

Eric Wildfang

Jeff Smith

Wenhong Zhu

Alexey Eroshkin

Yoshinobu Igarashi

Andrei Osterman

Adam Godzik

Andy Tao

Jingwei Li

Melanie Richter

Matt Bogyo

Steven Verhelst

Yie-Hwa Chang

Salvatore Sechi

Alexey Nesvizhskii