



Edman Sample Preparation Part 2

- Bill Henzel Genentech Inc.



Determining a Protein is Blocked

- Estimate the amount of protein applied to the protein sequencer. Compare the staining intensity of the band of interest to a standard.
- Does the amino acid background increase with time?
- Is the protein sequencer working?
- Can other bands on the electroblott with similar staining intensity be sequenced?




Common Blocking Groups

- Pyroglutamic Acid
- Acetyl Ser, Ala, Met, Thr, Gly



Identification of Blocked Proteins

- Pyroglutamate aminopeptidase deblocking (pyrococcus furiosus)
 - 1hr at 90 °C
- Rapid enzymatic cleavage (5-30 min.) and Edman degradation.
- Chemical cleavage methods:
- CNBr Met/X
- Dilute Acid Asp/Pro
- Hydroxylamine Asn/Gly
- Iodosobenzoic acid, DMSO/HCl Trp/X



Deblocking Protein Containing N-Terminal Pyroglutamic Acid

- Block PVDF membrane with 1% PVP-360 at 25° C on a shaker for 20 Minutes.
- Cleave with 1 mU of *Pyrococcus furiosus* pyroglutamate aminopeptidase in 50 mM sodium phosphate at 90° C for 1 hr.

Sequence Sorting of A Simple Protein Mixture from a Single Band



Notes: Major sequence matches *E.coli* ribosomal protein; *p1.BR3_ecd* sequence also present.

Sequence Data Results

Cycle/Run	1/108	2/109	3/110	4/111	5/112	6/113	7/114	8/115	9/116	10/117	11/118	12/119
Seq 1	S 0.232	L 0.540	S 0.729	T 0.484	E 0.512	A 0.232	T 0.447	A 0.263	K 0.191	I 0.503	V 0.472	S 0.282
Seq 2	G 0.561	S 0.344	S 0.729	H 0.320	H 0.435	H 0.425	H 0.416	H 0.372	H 0.294	S 0.302	S 0.256	G 0.282

Sequences Found

- Correct Sequence
 Unprocessed Sequence
 Internal Cleavage
 Wrong Protein
 Ragged N-term
 Contaminant Protein

RS15_ECOLI <RS15_ECOLI> 30a ribosomal protein s15 - escherichia coli

<0/site: Infr_smt

<65/site: Conflict MISSING (in Ref. 3).

<98 residues

SLSTEATAKIVS EFGRDANDTGSTEVQVALLTAQINHLQGHFAZHKKDHSRRGLLRMVSQRRKLLDYLKRKIDVARYTQLIERLGLRR

p1.BR3_ecd

MQS **S** **S** **S** **H** **H** **H** **H** **H** **S** **S** **G** **L** **V** **P** **R** **G** **S** **H** **M**

RRGPRSLRGRDAPAPTFCVFABCFDLLVRHCVACGLLRTPRPKPAGASSPAPRTALQFQESVGAGAGE

Sequence Sorting of a Complex Mixture

Notes: Multiple clips found; ragged N-terminus.

Sequence Data Results

Cycle/Run	1/15	2/16	3/17	4/18	5/19	6/20	7/21	8/22	9/23	10/24
Seq 1	R	R	L	R	V	R	S	Q	R	S
	2.812	1.859	2.003	1.623	0.866	1.673	0.285	1.755	1.895	0.215
Seq 2	S	Q	R	S	R	D	S	S	V	P
	0.514	2.005	1.593	0.227	1.630	1.232	0.285	0.215	-	0.645
Seq 3	A	R	R	L	R	V	R	S	Q	R
	0.849	1.859	1.593	2.311	1.630	0.939	1.698	0.215	0.120	0.895
Seq 4	S	R	D	S	S	V	P	T	Q	C
	0.514	1.859	1.698	0.227	0.168	0.839	0.644	1.793	0.120	0.216

Sequences Found

- Correct Sequence
 Unprocessed Sequence
 Internal Cleavage
 Wrong Protein
 Ragged N-term
 Contaminant Protein

1	RRLRVRSQRS	2.68 pmol
2	SQRSRDSSVP	1.03 pmol
3	ARRLRVRSQR	0.79 pmol
4	SRDSSVPTQC	0.72 pmol

p1.pRK5mBR3.ECD.mIgG

< Tue Jul 24 15:47:44 2001

< /home/ruby/va/Molbio/sam/sequences/tntfrag/like/consdomain/ss.pRK5mBR3.ECD.mIgG (5568 bases)

<314 residues, 0 stop; molecular weight: 34578.73

MSALLILALVGAAVASTG**ARRLRVRSQRSRDSSVPTQC**NQTECFDPLVRNCVSCLEFHTPDGTGHTSSLEPGTALQPQBSQ

VTGGKKIVPRDOGCKPCICTVPEVSSVFIFPPKPKDVLITITLTPKVTCCVVVDISKDQPEVQFSMPVDDVEVHTAQQQPRE

EQFNSTFRSVSELPIMHQDLNGKEPKCRVNSAAFPAPIEKTIKTKWGRPHAPQVYTIPPPKEQMAKDKVSLTCHITDFP

PEDITVENQMGQPAENYKNTQPIQNTNGSYFVSKLNVQKSNWEAGNTFTCSVLHGQLNHRTEKSLGHSFGK



CNBr Cleavage of Blocked Proteins

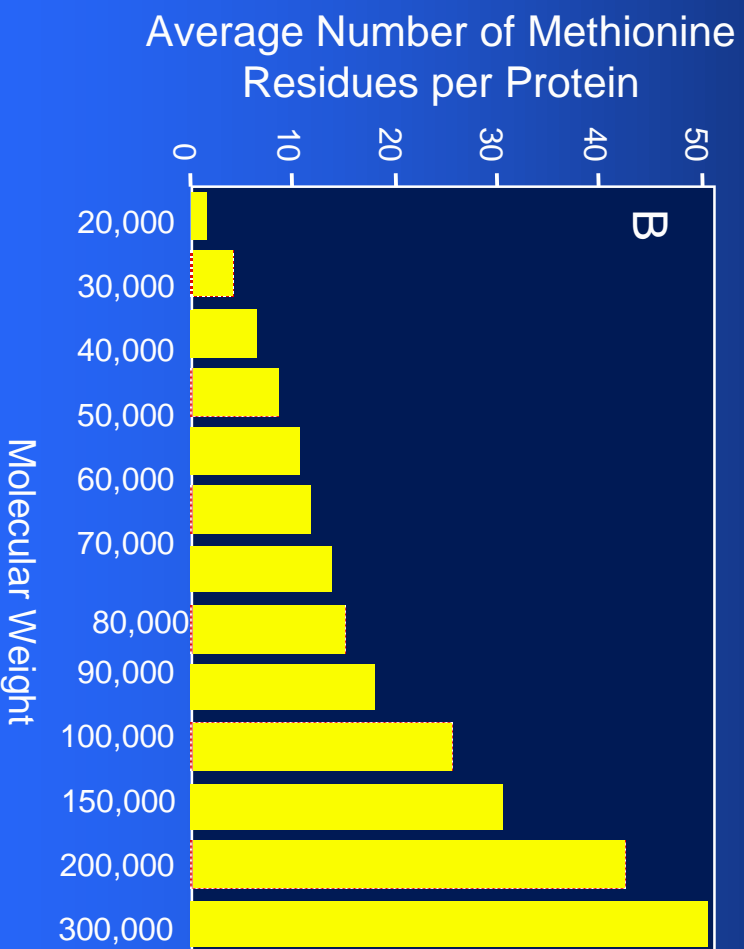
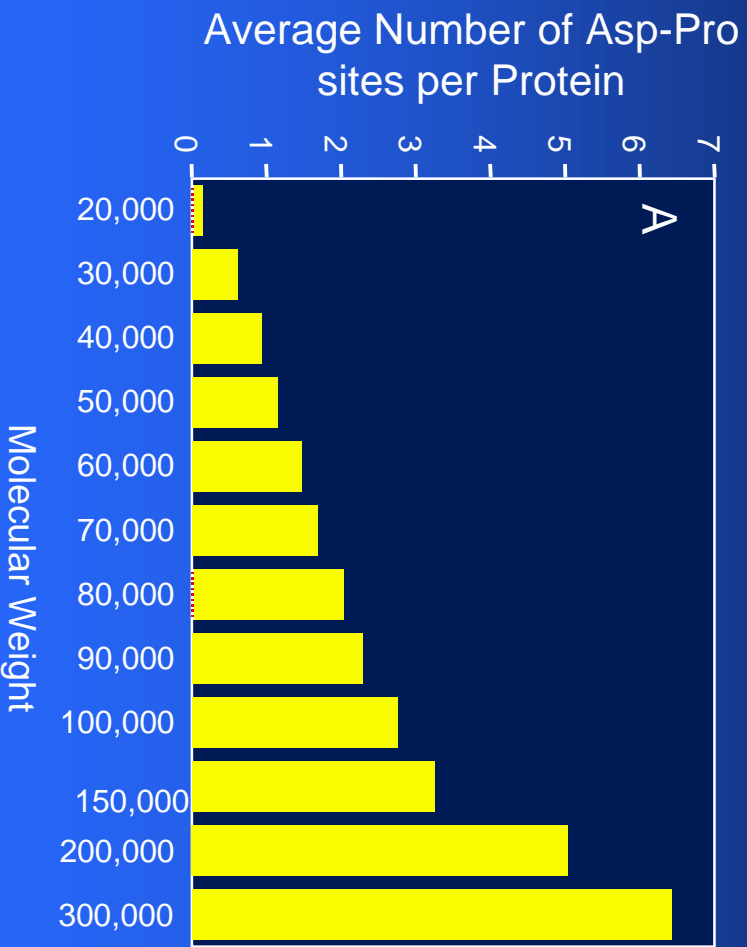
- Wet PVDF with 1 μ l of methanol.
- Add 20-30 μ l of 0.1 N HCl 45°C 3 hr.
- Speed-Vac to completely dry PVDF membrane.
- Journal of Biochem. Biophys. Methods (1980) 3, 11-30.

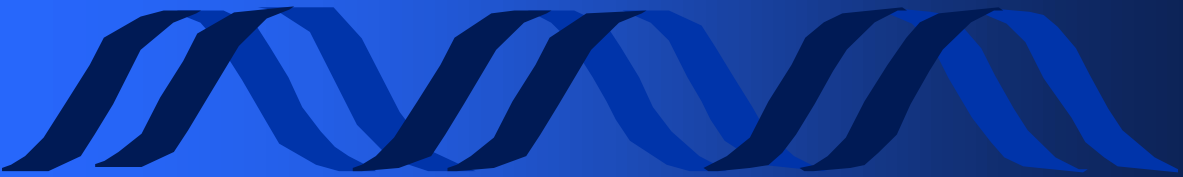


Rapid Method for Generating Internal Sequence of N-terminal Blocked Proteins

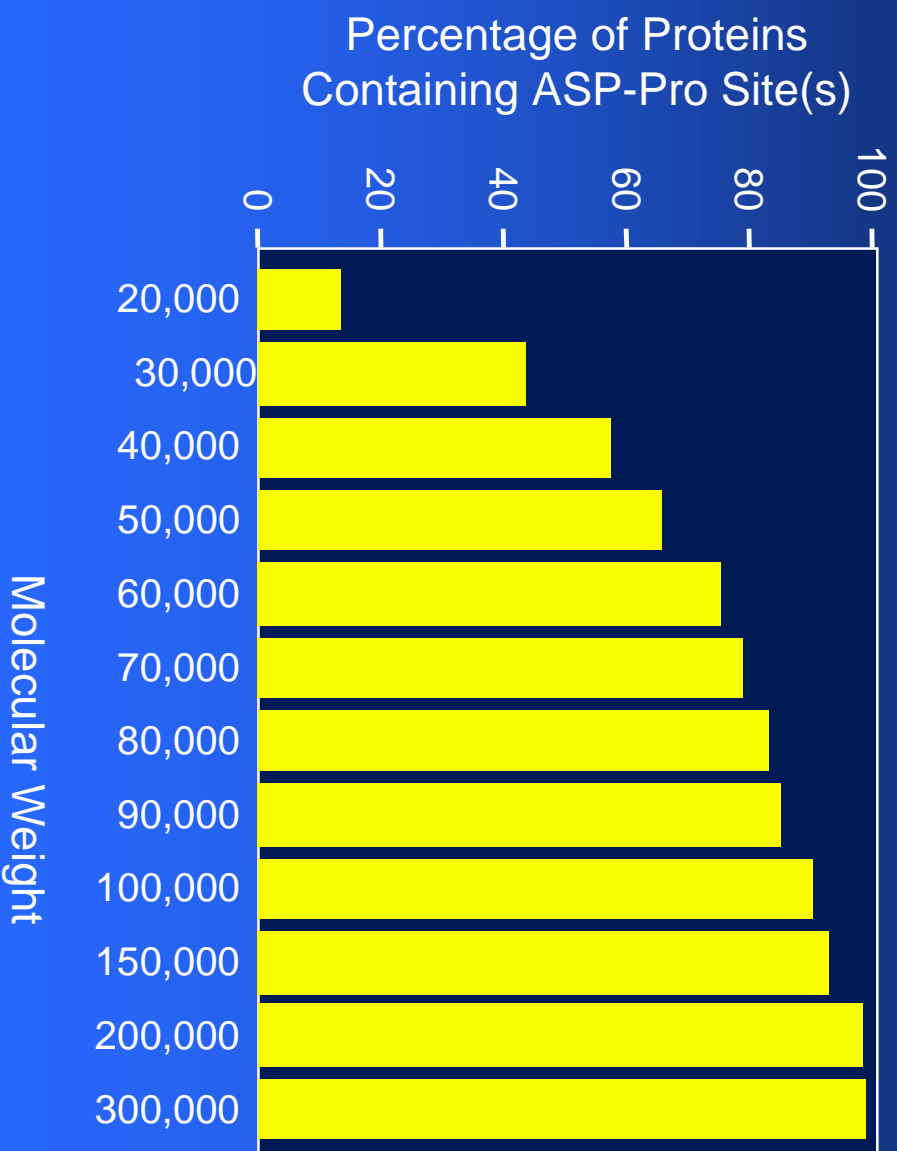
- Asp-Pro Cleavage
- 10% acetic acid, 7M guanidine HCl pH2.5 90°C
1 hr +OPA (Blocking all AA's except Pro)
- Anal. Chem. 72,5431-4536 (2000)

Comparison of Asp/Pro and Met Cleavage Sites





Percentage of Proteins Containing Asp/Pro Sites



Sequence Analysis of a Asp/Pro Cleavage Reaction of Human Acetylglucosaminyltr Poly-His



Human Acetylglucosaminyltr Poly-His

><subunit 1 of 1, 424 aa, 1 stop

><MW: 48170, pI: 7.45, NX(S/T): 2

MSYAIRCAFYQLLLAALMLVAMLQLLYLSLLSGLHGQEBQDQYFEFFPPSPRSVDQVKAQLRTALASGGVLDASGDYRVY
RGLLKTTMD **PNDVILATHA** SVDNLLHLSGLLERWEGPLSVSVFAATKEEAQLATVLAAYALSSHCPDMRARVAMHLVCPSR

YEAAVPD **PREPGEFALL** RSCQEVFDKLARVAQPGINYALGTNVSYPNNLLRNLRBEGANYALVIDVDMVPSEGLWRGLRE

MLDQSNQWGGTALVVPAPFEIRRARRRMPMNKNELVQLYQVGEVRPFYYGLCTPCQAPTNYSRWVNLPEESLLRPAYVVPWQ
DWPEFFYVAGG KVPTFDERFRQYGFNRSQACELHVAGFDPEVLNEGFLVHRGPFKEALKPHQPQKEAENQHNKILYRQFKQ

ELKAKYPNSPRRCAAHHHHHHHH

Sequence Data Results

Cycle/Run	1/3	2/4	3/5	4/6	5/7	6/8	7/9	8/10	9/11	10/12
Seq 1	P 0.721	R 0.310	E 0.878	P 1.197	G 0.183	E 0.037	F 0.306	A 0.703	L 0.659	L 0.256
Seq 2	P 0.721	N 0.100	D 0.514	V 0.495	I 0.167	L 0.428	A 0.063	T 0.122	-	A 0.195
Seq 3	P 0.721	W 0.706	E 0.878	P 1.197	F 0.012	Y 0.171	V 0.558	A 0.703	G 1.334	G 0.381

N-terminally Blocked Proteins Identified by Database Searching from Asp-Pro Cleavage Sequence Data

Accession #	Protein Name	Molecular weight (Da)	No. sequences found/expected	Starting Residue No.	Sequence Found
SY16_HUMAN ^a	SY_16 Human small inducible cytokine a16 precursor	11,184	1/1	91	DPNLPLLPTRN
LITB_HUMAN ^a	lithostathine 1-beta precursor human	16,414	2/2	54 105	DPETWVDADLYCQN DPKKNRRWHWSSGS
NM_001330_1 ^a	Cardiotrophin 1 human	21,227	2/2	11 55	DPQTDSSVSLL DPFGLPSFSPP
CAH2_BOVIN ^a	Bovine carbonic anhydrase	28,508	2/2	41 179	DPALXPLALVYG DPGSLLPNVLDY
JW0047 ^b	Class I cytokinase receptor precursor human	40,431	2/2	144 361	DPLEATVHWAP DPLEXLNWVRL
RAGE_MOUSE ^a	Advanced glycosylation end product specific receptor precursor mouse	40,448	1/1	128	DPASELTASVPN
HGF_HUMAN ^a	Hepatocyte growth factor precursor human, alpha chain	57,157	3/3	55 276 370	DPALKIKTK DPHTRWEYC DPNIRVGYC
AF_156884_1 ^a	Receptor interacting serine threonine kinase 3 human	56,901	3/3	88 152 503	DPKPALVTK DPELHVKLA DPEAWSRPQ
AF265242_1 ^a	Type-1 T cell cytokine receptor human	69,460	2/3	144 361	DPLEATVHWAPP DPLEKLNWVRLP
RIP_HUMAN ^a	Serine-threonine protein kinase rip human	75,959	3/3	419 481 589	DPFAQQRPYEN DPGTAGPRVWY DPIRENLGKHW
AF316597_1 ^a	Paracaspase human	92,272	4/4	7 563 641 754	DPLQALPPSAAPT DPIQGTEYSAESL DPKDANKGTPEET DPFHGVYHSHPGN
PHS2_RABIT ^a	Glycogen phosphorylase rabbit	97,158	3/3	80 322 635	DPXRIYYLSLEF DPVRTNFDAFPD DPVVGDRLRVIF

^a Swissprot Database

^b Genebank Translated Database

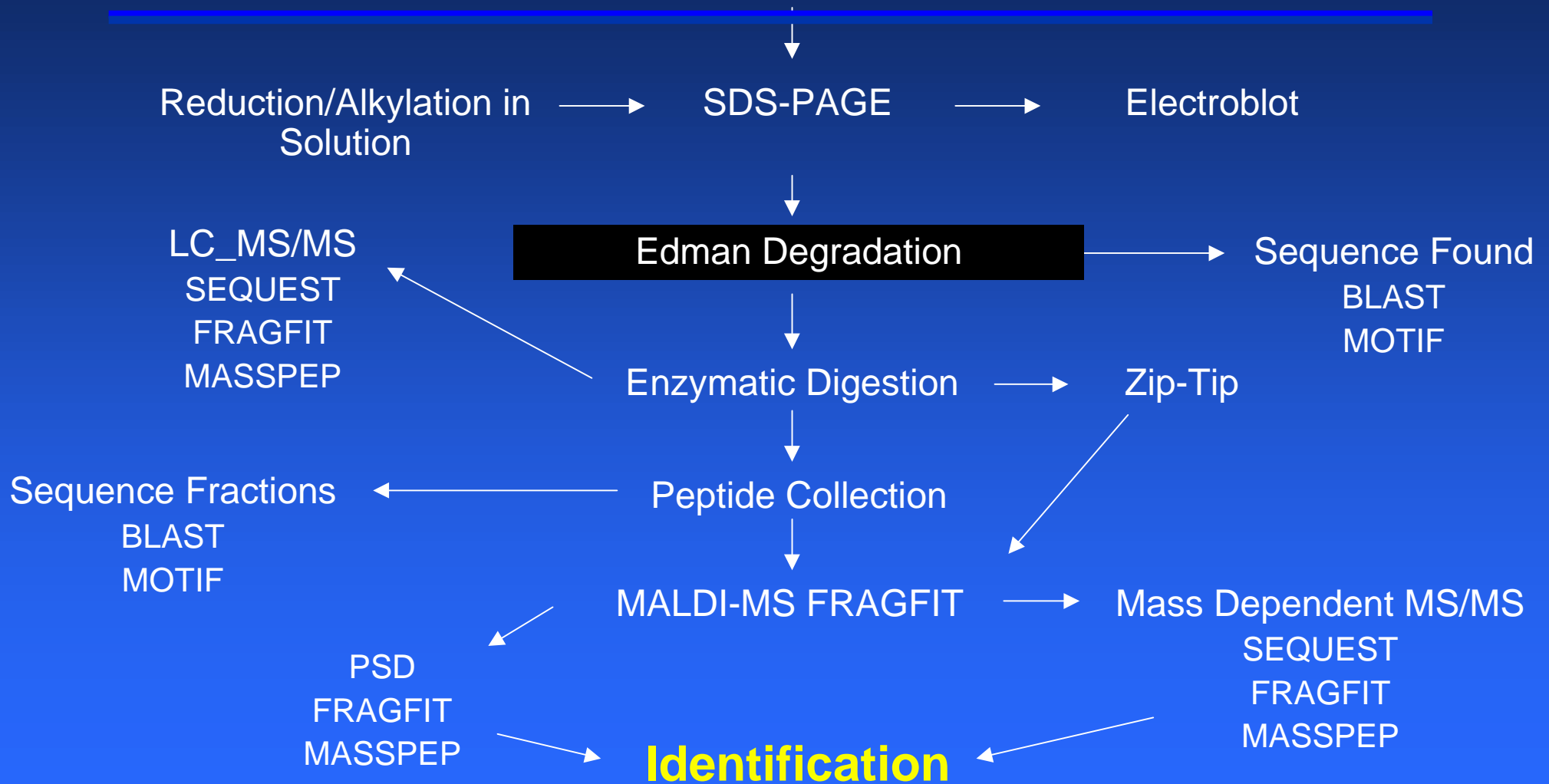


Rapid Enzymatic Cleavage of Blocked Proteins

- Wet PVDF with 1-2 μ l of Methanol.
- Block with 0.5 ml of 1% PVP-360.
- Add 0.2 μ g of Trypsin, Lys-N, Asp-N, Glu-C, Lys-C or Arg-C.
- Incubate 5-20 minutes at 37° C.
- Remove buffer and wash with water.
- Sequence peptides mixtures remaining on PVDF membrane.

Summary of Protein Identification Methods

Unknown Protein or Mixture of Proteins





Acknowledgements

Victoria Pham

Wendy Sandoval

Suzy Wong



Protein identification based on similarity to amino acid composition

- Each amino acid has an assigned weight, which were derived from a statistical analysis of the difference of the theoretical and experimental mole percentage values of 25 proteins.
- The molecular weight is also used as a weighting factor.
- A number of websites have similar programs some of which are listed on the ABRF website.



AminoSearch Search Results

Protein Name	Rank of match
xTGFb.gIHi.V5L.RK.theo	1
FGF19	1
HGH	1
HLFA1	1
Indian Hegehog	1
Myoglobin	1
MAMG	1
Complement Hlog	1
Chemokine beta 8	1
FKBP	1
VEGF 165 KDR	3

Distribution of Amino Acids in the N-terminal Region of Human Secreted Proteins

