ABRF 2003 ESRG Sample

Analysis of a PVDF Blotted Protein with a homogeneous N-terminus

Members of the Committee

John M. Neveu (Chair) - Harvard University
Scott Buckel - Amgen, Inc.
Richard G. Cook - Baylor Col of Med
J. Myron Crawford - Yale Univ.
Nancy Denslow - Univ of Florida
Joseph Fernandez - Rockefeller Univ
Benjamin J. Madden - Mayo Clinic
Laurey Steinke (EB liaison) - University of Nebraska Medical Center

Requested Information

- Sequence information about sample and identification
- Instrument and chemistries used
- Survey of instruments in laboratory
- Submission of actual report given to clients
- pmol per cycle sheet for initial yield (IY) and repetitive yield (RY)

Rationale of Study

- Determine the ability of participating laboratories to sequence an electroblotted sample containing low pmol amounts of protein
- Compare IY and repetitive yield RY
- Determine the ability of laboratories to identify the protein

Choice of Sample

- Sample was a protein from a complex submitted to a committee member's lab for analysis
- Same complex as protein used for ABRF 2002 ESRG
- Protein had a homogeneous N-terminus
- The sequence was available in several databases

Instructions to participants

- Enter single letter code for AA
- Tentative calls should be entered in ( )
- Use a "–" if no identification could be made at a particular position

Sample Preparation

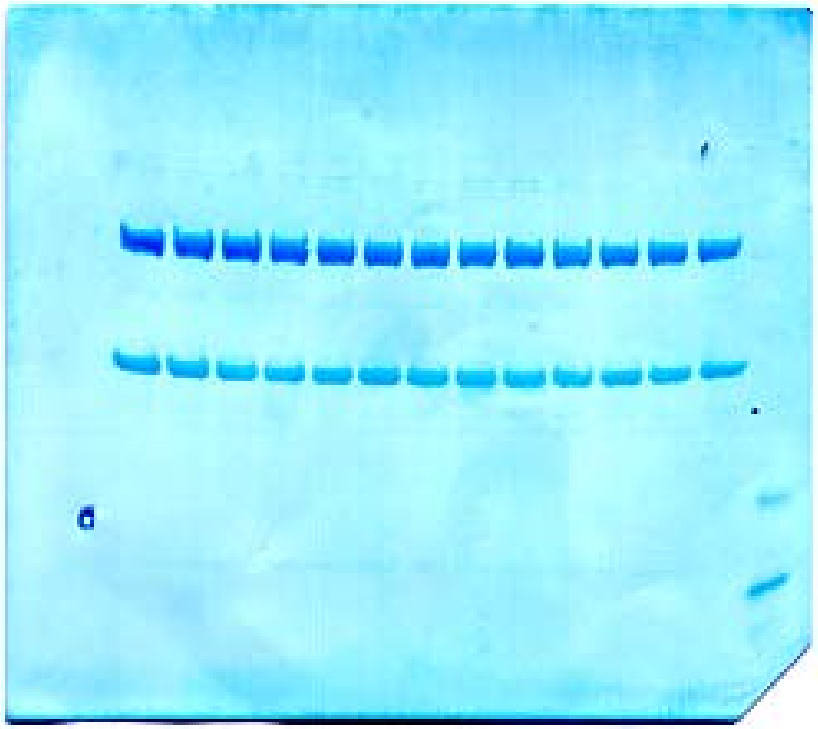
- Sample was electrophoresed on Novex Gels
- Sample was electroblotted to PVDF (Novex 0.22 um) at 250 mAmps for 2 hrs.
- Membranes were stained with 0.05% Coomassie G-250 5 min, 50% MeOH/10% HAc
- Membranes were destained with 40% MeOH/10% HAc
- Membranes were washed in water overnight
- Stored at room temperature for 1 yr

PVDF Membrane

49.4 kDa →

26 kDa →

ABRF-2002



(13.5 pmol)

Sample testing

- Samples were tested by each ESRG member lab to determine sample variability
- By amino acid analysis the bands contained 9-18 pmol
- Samples were then sent to participating laboratories

Data analysis

- Positive Correct Call (PC) – if only 1 correct amino acid listed
- Tentative Correct Call (TC) – if more than 1 positive amino acid was listed including the correct one, or if user indicated the call as tentative
- Positive wrong calls (PW) – if amino acid was wrong
- Tentative wrong call (TW) -- if tentative amino acid was wrong.

Sequence of sample

1 5 10 15 20 25

H M T T R L T R W L T A L D N F E A K M A L L P A

30 35 40 45 50

V R R Y G R L T R A T G L V L E A T G L Q L P L G

Sequence of sample

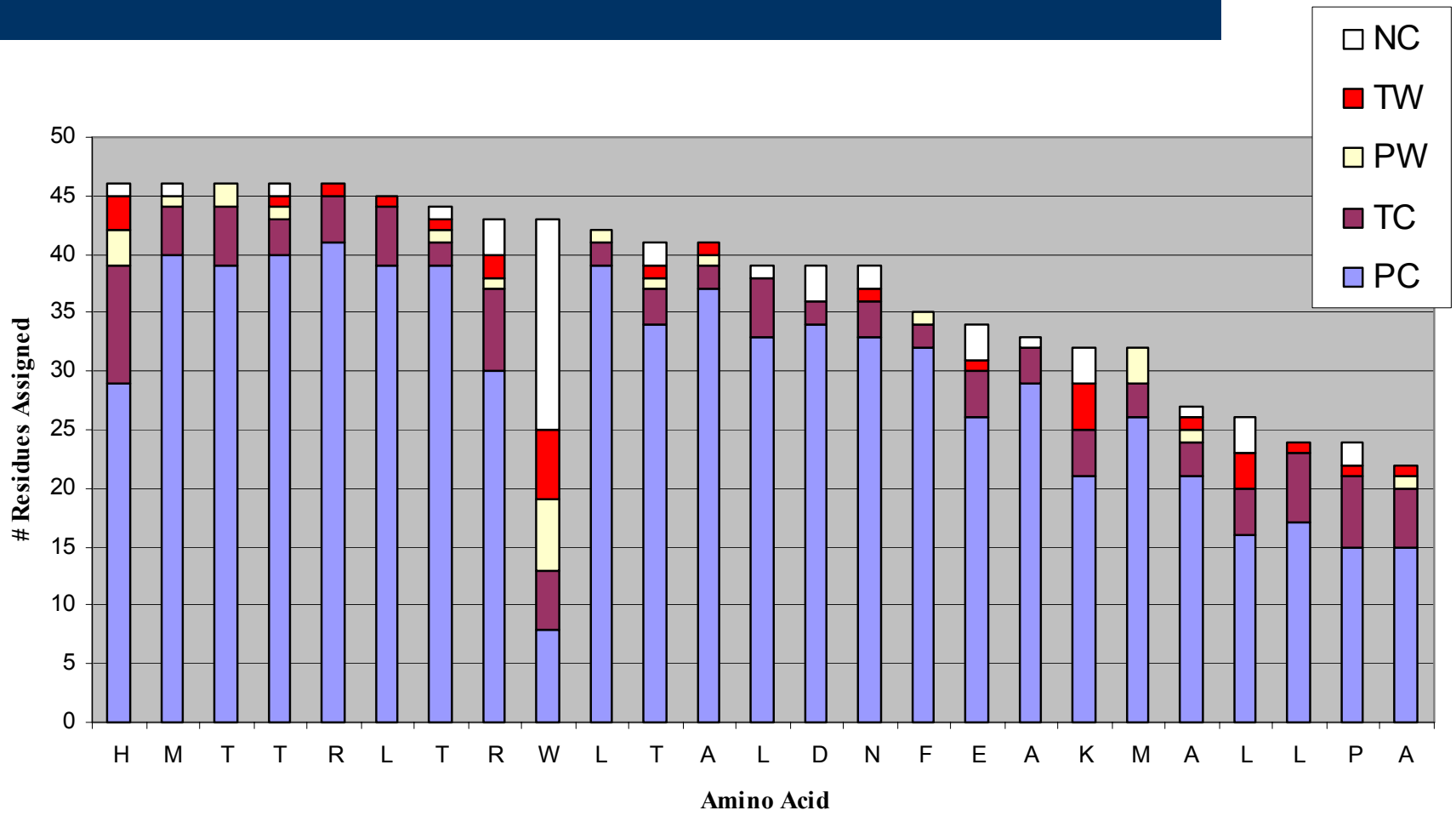
1 5 10 15 20 25

H M T T R L T R W L T A L D N F E A K M A L L P A

30 35 40 45 50

V R R Y G R L T R A T G L V L E A T G L Q L P L G

Sequence Calls For First 25 Residues for ABRF-2003 ESRG



Summary of Sequence Assignments for ABRF-2003ESRG Compared With Other ABRF Edman Sequence Studies of Proteins

Description	Equation	ABRF-2003 ESRG	ABRF-2002 ESRG	ABRF-99SEQprotein
No. responses	R	46	31	45
Avg. no. correct	(PC+TC)/R	19.1	20.3	11.5
Avg. no. positive	(PC+PW)/R	17.3	19.2	10.5
Avg. no. incorrect	(PW+TW)/R	1.2	6.7	0.9
Accuracy of PC calls	PC/(PC+PW)	96.9%	76.3%	98.7%
Accuracy of TC calls	TC/(TC+TW)	77.1%	72.6%	58.1%

PC = positive correct
PW= positive wrong
TC = tentative correct
TW = tentative wrong

Calculations for IY and RY

IY: Subtract pmol amount from previous cycle

$$RY(\%) = \text{log-1}\ \{\underline{\log(\text{pmole B}/\text{pmole A})}\} \times 100$$

(cycle B - cycle A)

Initial Yield and Repetitive Yields

IY MET (2) (pmol)	RY MET (2,20)	RY LEU (6,10,13,22)	RY ALA (12,18,21,25)	Avg # Correct (First 25)
1.96	91.4%	89.8%	91.3%	18

0.24 3.91	87.8 96.2	71.8 102	80.0 116.9

Top Responses – First 25 residues

Minimum of 22 residues – 100% PC (12 labs)

Facility #	# PC	# TC	IY (M2)	W (ID)	RY	Sequencer
41	25	0	3.7	PC	94%	HP G-1005A
29	25	0	2.3	PC	89.80%	ABD Procise-cLC
32	25	0	2.9	PC	-	ABD Procise-HT
17	25	0	-	PC	-	HP G-1005A
38	24	1	1.3	PC	92.50%	ABD Procise-HT
7	24	1	2.7	TC	91.10%	ABD Procise-HT
31	24	0	2.6	-	91.70%	HP G-1005A
20	24	0	2.2	-	91.60%	ABD Procise-cLC
23	24	0	1.4	-	90.80%	ABD Procise-HT
8	23	1	2.6	-	93.80%	ABD Procise- HT
3	23	1	1.9	-	91.30%	ABD Procise-HT
18	22	3	1.2	PC	91.30%	ABD Procise-HT

Responses beyond 25 residues

Facility #	#PC	#TC	%PC	#PW	#TW	Sequencer
41	41	5	100%	0	0	HP G-1005A
17	48	0	98%	1	0	HP G-1005A
7	29	2	100%	0	0	ABD Procise-HT
20	26	4	100%	0	1	ABD Procise cLC
4	25	1	96%	1	1	ABD Procise -HT

Performance by Model of Sequencer

Manufact	Model	n	Avg % PC	Avg % TC	Avg # correct	RY	IY (pmol)
ABD	Procise HT	30	98.8%	82.4%	19.5	90.8%	1.8
ABD	Procise cLC	8	94.7%	75.0%	19.5	88.1%	2.1
ABD	477	3	80.0%	46.0%	7.3	81.8%	2.6
Shimadzu	PPSQ-23A	1	50.0%	NA	5	87.6%	2.3
HP	G-1005A	4	99.1%	50.0%	27	93.1%	3.1

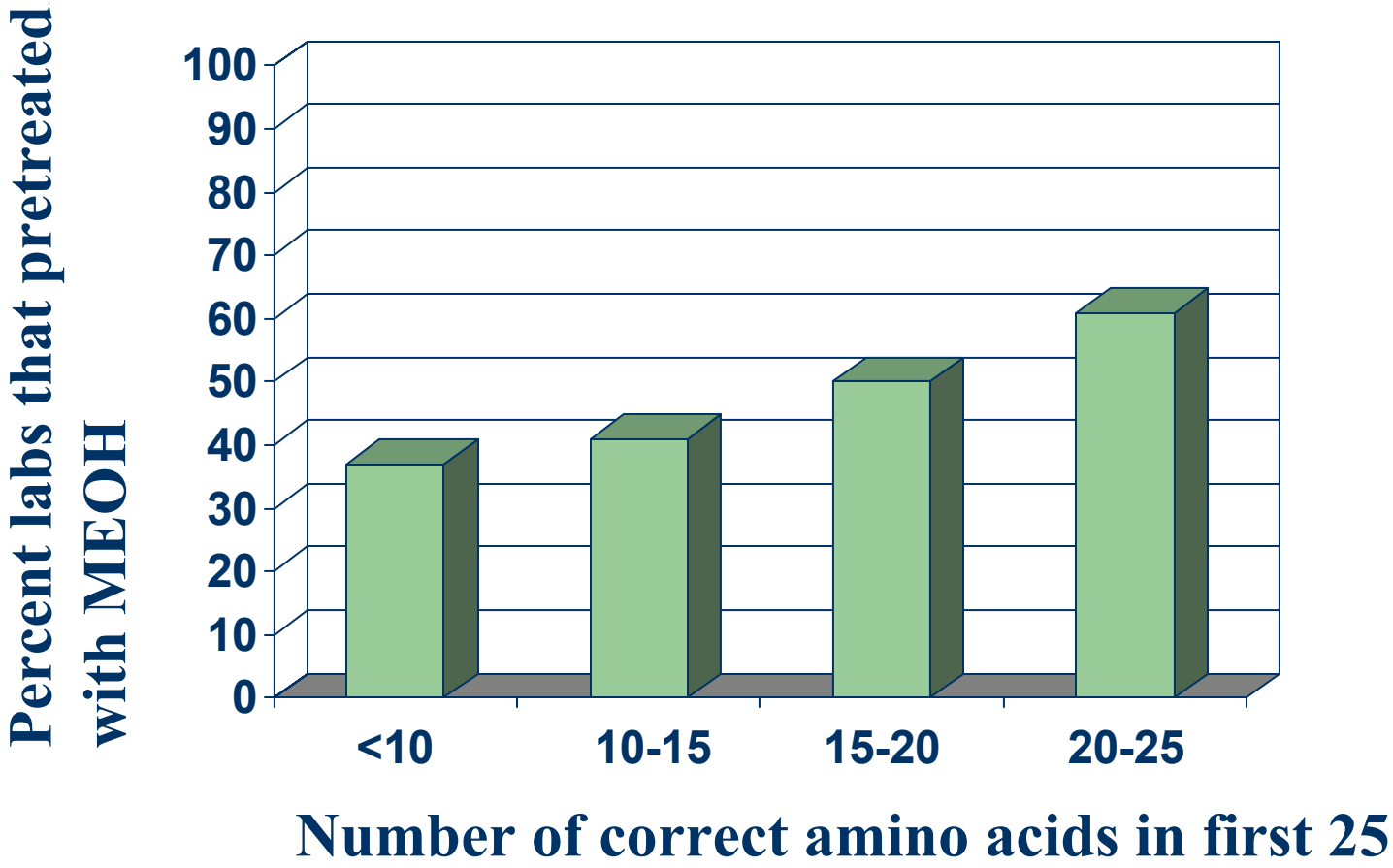
Reagent additives

Additive	ABD (41)	HP(4)	Shimadzu(1)
• DTT in S2B	3	1	-
• DMPTU in SB2	7	1	-
• Acetone in SA3	20	1	1
• Other additives	7	2	1

- Potassium/ sodium phosphate
- TFA, formic acid, acetate
- TCEP
- Trp, Norleu
- TEA
- hexanesulfonate

No observed differences in results

Sample pretreatment -- MeOH (Number of PC calls in first 25 residues)



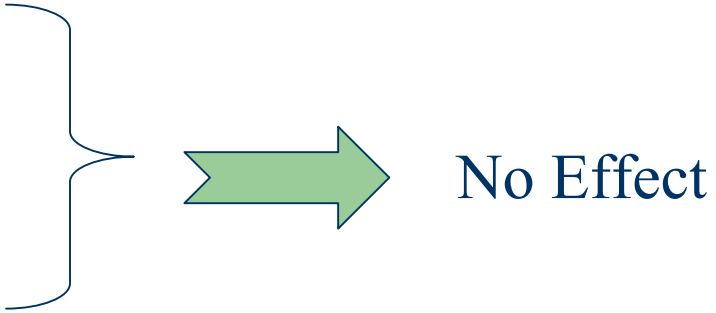
Sample pretreatment and cartridges

Sample Pretreatment

- MeOH wash -- Positive effect
- Polybrene -- no effect
- Water wash – no effect

Cartridges used

- GFF (-filter)
- GFF (+ Polybrene filter)
- PVDF blot cartridge
- SAX column



Database search results

Identity of protein: **Flagellum-specific ATP synthase (*Salmonella typhimurium*)**

		Accuracy of Positive Calls
Correct identification:	30 (65%) →	100%, 91% (1) 94% (1)
Incorrect identification:	1 →	63%
Didn't try:	9	
Ambiguous Identification:	11 →	0-100%
Correct	3 →	89, 98, 100%

Programs used to identify protein

Programs	# labs	Positive ID
Blast(P)	16	12
EMBL	1	1
FASTA (GCG, EMBL)	8	7
FASTA3	4	3
MSEdman	2	2
EMBL (MP search)	2	2
GCG-find patterns	1	1
Protein prospector	1	1
Protein info	1	-

Databases searched

Databases	Correct ID	Not Identified
SP	7	2
SWALL	4	1
nr	11	2
GB	1	-
GPTR	1	1
PIR	2	-

User reports

- Reports ranged from being very formal to informal
- Some were handwritten – others typed memo style
- Some reported sequence only, others reported yield of amino acids, still others gave histograms.

Survey Results

Number of labs that participated		**Top 12**
• Academic core lab – 26	→	8 (31%)
• Academic lab – 4	→	0 (0%)
• Commercial facility – 2	→	2 (100%)
• Sequencing facility for commercial organization – 9	→	1 (11%)
• Not reported – 5	→	1 (20%)

Instrumentation of 12 top labs

- 5 labs have only 1 Sequencer (HT-494)
- 5 labs have at least 2 sequencers (HP, HT-49X, cLC, 477, Shimadzu)
- 1 lab has 3 sequencers (ABD HT-494, (2) 477)
- 1 lab has 4 sequencers (HP, ABD HT-494, cLC, Shimadzu)

Mass Spec capabilities of responding labs

- 32 laboratories (70%) have some form of mass spec capabilities
- 29 laboratories (63%) have MALDI TOF capabilities
- 22 laboratories (48%) have nanospray or LC/MS
- 19 laboratories (41%) have both electrospray and MALDI capabilities

Conclusions

- This was a very successful trial for a blotted sample with an average sequencing yield of 2 pmol
- 22 **(48%)** facilities called 25 or more amino acids,
  20 facilities called 10-24
  4 facilities called less than 10
- 33 **(72%)** facilities had 100% positive accuracy
- Most difficult to call amino acid was W9
  8 PC and 5 TC
- Accuracy of PC 96.9%
  TC 77.1%

Conclusions – con’td

- 30 **(65%)** facilities were able to identify the protein which was Flagellum-specific ATP synthase (*Salmonella typhimurium*) using a variety of search engines and data bases.
- **20%** didn’t try to identify the protein
- **70%** labs now have mass spectrometry as an option

Acknowledgements

- Thanks to Dr. Robert Macnab (Yale University, New Haven, CN) and to Dr. Tohru Minamino (Protonic Nanomachine ERATO, Kyoto, Japan) for the use of their protein complex in this study.

- Thanks to all the participating laboratories for taking the time to analyze the sample and send in their results.