

ABRF-PRG04: Differentiation of Protein Isoforms

D. P. Arnott¹, M. A. Gawinowicz², J. A. Kowalak³, W. S. Lane⁴, T. A. Neubert⁵, K. D. Speicher⁶, C. W. Turck⁷, and K. A. West⁸

¹Genentech, Inc., South San Francisco, CA, ²Columbia University, New York, NY, ³NIH, Bethesda, MD, ⁴Harvard University, Cambridge, MA, ⁵New York University School of Medicine, New York, NY, ⁶The Wistar Institute, Philadelphia, PA, ⁷Max Planck Institute of Psychiatry, Munich, Germany, ⁸Cleveland Clinic Foundation, Cleveland, OH

MISSION

The mission of the Proteomics Research Group (PRG) is to assist ABRF members in evaluating their capabilities to identify "unknown" proteins in order to establish realistic expectations for this technology.

ABSTRACT

Accurate protein identification sometimes requires careful discrimination between closely related proteins. Protein isoforms and/or protein homologues may differ by as little as a single amino acid substitution or posttranslational modification, yet these differences can have profound effects on the structure and function of the proteins. The ABRF-PRG04 sample was designed to study a simple mixture of 3 closely related proteins, each at 3 picomoles. The Proteomics Research Group (PRG) sent the sample to interested laboratories in the form of intact proteins, and participating laboratories were asked to identify the proteins and report their results. The primary goal of the PRG04 Study is to give participating laboratories a chance to evaluate their capabilities and practices with regards to sample fractionation (1D- or 2D PAGE, HPLC, or none), protein digestion methods (in-solution, in-gel, enzyme choice), and approaches to protein identification (instrumentation, use of software and/or manual techniques to facilitate interpretation), as well as determination of amino acid or posttranslational modifications. Compilation of submitted data will allow a comparison of the strategies used and aid in optimization of these techniques.

INTRODUCTION

A common quest in proteomics core facilities, aside from the usual identification of proteins, is to locate differences within protein(s) of interest. These may simply be cross-species differences or modifications associated with a vital role, for example in protein function or activity. The investigator may have only minimal information regarding the type or location of these differences. These considerations make the ability to locate such variations very important. To make this determination, a core facility must investigate the protein in depth, going beyond the identification of the protein. This study will help in evaluating a core facility's ability to identify closely related proteins and determine where the differences exist between them. Therefore, the primary goals of this study are to give each laboratory a chance to evaluate its capabilities and practices with regards to:

- Protein Digestion methods (solution-based, in-gel, choice of enzymes)
- Protein identification methods
- Methods for the determination of amino acid differences between protein isoforms
- Amino acid sequence coverage of the identified proteins

as well as to obtain data that would allow a comparison of the strategies used and aid in optimization of these techniques.

With this in mind, we provided laboratories that requested samples (106 labs) with a mixture of three intact proteins: two bovine carbonic anhydrase isoforms differing by a single AA substitution plus a human carbonic anhydrase. The identity of the respondents was not known to members of the PRG. Since the sequences of these proteins are in the public databases, we required that the participants supply proof of the differences ascertained by MS/MS or some other means they may have used. This study relates to those previous PRG studies in that proteins in a mixture will be identified and AA modifications/substitutions will be analyzed. However, in this study, intact proteins, rather than a predigested mixture, were examined. Furthermore, the analysis involved searching for differences between 3 very similar proteins, rather than modifications within the same protein. Also, because the participants did their own digests, this study addressed, to some extent, the issue of front end sample preparation. We asked that participants return proof of the differences they have determined, along with the completed questionnaire.

METHODS

Three carbonic anhydrase proteins (C6165, C2522, and C3640) were purchased from Sigma-Aldrich CO (St Louis, MO). Small amounts of each protein were weighed out on a Cahn microbalance and an appropriate amount of 1% acetic acid was added to give a final concentration of 1 µg/µL (stock solutions). Three microliters of each protein solution was placed into an amino acid analysis tube, each sample was prepared in triplicate, with 3 µL of 1% acetic acid as a control. The samples were lyophilized in a vacuum centrifuge, sealed by wrapping in parafilm and analyzed for amino acid content, which determined the protein content of the solid proteins to be approximately 60%.

Fresh stock solutions (1 µg/µL) of each of the three proteins were prepared in 1% acetic acid as were fresh working solutions at a concentration of 3 picomoles/µL. Samples were prepared by aliquoting 1 µL of each protein solution into 0.5 µL Eppendorf tubes containing 10 µL of 1% acetic acid. All samples were lyophilized in a vacuum centrifuge, and sealed by wrapping in parafilm. Lyophilized samples were then sent to 85 requesting laboratories in the USA and 21 labs in Europe and Israel for protein identification.

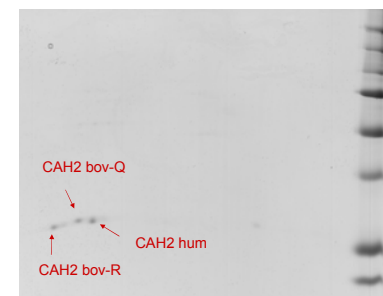


Figure 1. 2-D PAGE of 3 pmol carbonic anhydrase protein mixture. The first dimension pH gradient was 3-10 nonlinear, the second dimension SDS-PAGE gel contained 12% acrylamide. Proteins were visualized by Colloidal Coomassie Blue staining.

Basic

Acidic

Table 1. Sample Prep Information

ID	Solvent	Volume	Separation Method	Gel or Column Type	Staining Method	Protein 1			Protein 2			Protein 3		
						Protein 1 ID	Confidence	Total %	Protein 2 ID	Confidence	Total %	Protein 3 ID	Confidence	Total %
715	SDS SB pH 8.5	10	SDS-PAGE	10% Tris-Gly Standard	CBB	CAH2bov-Q	P	79	CAH2bov-R	P	79	CAH2hum	P	77
98186	SDS SB HNE	20	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	66	CAH2bov-R	P	65	CAH2bov-Q	P	66
20702	1% FA	20	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	63	CAH2bov-R	P	59	CAH2bov-Q	P	60
27974	water	20	SDS-PAGE	12% Bis-Tris NuPAGE	SYPRO Orange	CAH2bov-Q	P	57	CAH2bov-R	P	58	CAH2hum	P	47
82013	water	80	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	58	CAH2bov-R	P	59	CAH2hum	P	51
4343	25 mM ABC	20	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	52	CAH2bov-R	P	52	CAH2hum	P	52
29103	100 mM ABC	100	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	43	CAH2bov-R	P	40	CAH2bov-Q	P	39
11113	50 mM ABC, pH 8	15	SDS-PAGE (33% of sample)	12% Tris-Gly Mini	MS Friendly Silver	CAH2bov-Q	P	22	CAH2bov-R	P	20	CAH2bov-Q	P	20
11010*	water	10	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-R	P	49	CAH2bov-R	P	49	CAH2horse	T	31
1096352	0.1% TFA/0.1% ACN (8/1)	6	RP-HPLC	50% MeCN	Colloidal CBB	CAH2bov-Q	P	42	CAH2hum	P	45.9	CAH2bov-Q	P	42
21952	0.1% FA, 0.25% N-ethylmaleimide	15	2DE (50% of sample), RP-HPLC (33% of sample)	C8 Vydac	Colloidal CBB	CAH2bov-Q	P	79	CAH2hum	P	69	CAH2bov-Q	P	69
25519	water	30	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-R	P	73	CAH2hum	P	64	CAH2hum	P	64
11111	8 M Urea, 0.2 M Tris-HCl	20	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-R	P	63	CAH2hum	P	64	CAH2hum	P	64
90894	8 M Urea in 0.4 M ABC	10	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	62	CAH2bov-R	P	66	CAH2bov-Q	P	66
23026	6% ACN	6.3	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	67	CAH2hum	P	50	CAH2hum	P	50
94591	water	10	SDS-PAGE	12% Tris-Gly Mini	CBB	CAH2bov-Q	P	56	CAH2bov-R	P	55	CAH2bov-Q	P	55
23312	ABC	20	SDS-PAGE	12% Tris-Gly Mini	CBB	CAH2bov-Q	P	55	CAH2bov-R	P	55	CAH2bov-Q	P	55
2115	50 mM ABC	30	SDS-PAGE	12% Tris-Gly Mini	CBB	CAH2bov-Q	P	70	CAH2bov-R	P	66	CAH2bov-Q	P	66
24389	ABC	5	SDS-PAGE	12% MES Mini	SYPRO	CAH2bov-Q	P	46	CAH2bov-R	P	41	CAH2bov-Q	P	41
32559	ABC	20	SDS-PAGE	12% MES Mini	SYPRO	CAH2bov-Q	P	60	CAH2bov-R	P	50	CAH2bov-Q	P	50
11787	25 mM ABC, 10% ACN	50	SDS-PAGE	12% MES Mini	SYPRO	CAH2bov-Q	P	54	CAH2bov-R	P	44	CAH2bov-Q	P	44
73108	100 mM ABC	100	SDS-PAGE	10% Tris-Gly Mini	Silver	CAH2bov-Q	P	38	CAH2bov-R	P	30.4	CAH2bov-Q	P	30.4
89743	ABC	10	SDS-PAGE	10% Tris-Gly Mini	Silver	CAH2bov-R	P	11	CAH2bov-R	P	11	CAH2bov-R	P	11
10567	Water	25	RP-HPLC	4-12% NuPAGE	Silver	CAH2bov-Q	P	11	CAH2hum	T	7	CAH2bov-Q	P	7
24770	8 M Urea	150	RP-HPLC	C18, 75 µm x 15 cm, Pico20 µm	BSA	CAH2bov-Q	P	46	CAH2hum	T	20	CAH2bov-Q	P	20
92711	8 M Urea	6	RP-HPLC	C18, 75 µm x 15 cm, Pico20 µm	BSA	CAH2bov-Q	P	54	CAH2bov-R	P	46	CAH2hum	P	46
11735	5% ACN, 100 mM ABC	19	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	42	CAH2bov-R	P	77	CAH2hum	P	63
48583	50 mM ABC	10	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	66	CAH2bov-R	P	59	CAH2hum	T	16
89186	Tris-HCl, 8 M Urea	5	SDS-PAGE	10% Tris-Gly Mini	Zinc	CAH2bov-Q	P	47.9	CAH2bov-R	P	48	CAH2sheep	P	32.4
87050	SDS SB	60	SDS-PAGE	5-20% Tris-Gly Mini	SYPRO Ruby	CAH2bov-Q	P	62.5	CAH2bov-R	P	22.7	CAH2sheep	P	15.8
31815	50 mM ABC	20	SDS-PAGE	10% Tris-Gly Mini	CBB	CAH2bov-Q	P	38	CAH2sheep	P	32	CAH2bov-R	P	24
21068	0.1% TFA, 0.1% ACN	20	SDS-PAGE	10% Tris-Gly Mini	CBB	CAH2bov-Q	P	72	CAH2bov-R	P	22	CAH2sheep	P	17
640521	2% ACN, 0.1% FA	50	SDS-PAGE	10% Tris-Gly Mini	CBB	CAH2sheep	P	50	CAH2hum	P	40	CAH2bov-R	P	67
80053	50 mM ABC	10	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								
89117	40% ACN	80	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								
89117	5% FA	10	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								
106369	50 mM ABC	20	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								
13791	50 mM ABC	20	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								
11747	SDS SB	100	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								
13053	30% ACN, 0.1% TFA	20	SDS-PAGE	12% Tris-Gly Mini	CBB	no results								
11596	3% TFA	30	SDS-PAGE	12% Tris-Gly Mini	CBB	no results								
11596	25 mM ABC	10	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								
11128	25 mM ABC	10	SDS-PAGE	10% Tris-Gly Mini	CBB	no results								

ABC: ammonium bicarbonate
FA: formic acid
Glu-C: endoprotease Glu-C
SDS SB HNE: SDS sample buffer plus beta mercaptoethanol
TFA: trifluoroacetic acid

Sample Preparation Summary:

- 106 samples were sent to requesting laboratories, and 42 labs returned data
- 11/42 labs separated the proteins by 1-D SDS-PAGE (see corresponding protein ID results in Figure 2), 1 lab used 2-D SDS-PAGE
 - 5/11 labs used Coomassie Brilliant Blue to stain the gels
 - 1/11 labs used zinc staining
 - 3/11 labs used SYPRO orange or ruby staining
 - 3/11 labs used silver staining
- 2/42 labs separated the intact proteins by reverse phase HPLC
- 2/11 labs did not see any proteins by Coomassie staining of 1-D SDS-PAGE gels

Figure 2. Effect of Gel Stain on Protein Sequence Coverage

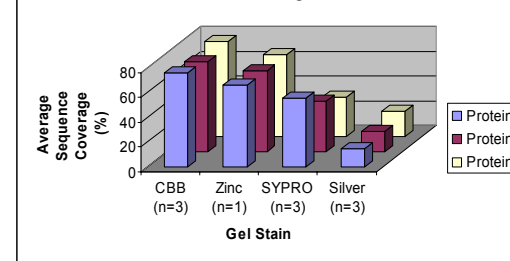


Table 2. Mass Spectrometry Results

ID	Protein(s) Dipeptides	Enzymes used	Peptide Masses Measured Without PSD or ms/ms			Peptide Masses Measured with PSD or ms/ms			Protein 1			Protein 2			Protein 3																	
			#Peptides Matched	%Sequence Coverage	#Peptides Matched	%Sequence Coverage	Protein 1 ID	Confidence	Total %	Protein 2 ID	Confidence	Total %	Protein 3 ID	Confidence	Total %																	
715	Y	T, Ch	19	79	19	79	21	77	18	78	18	78	16	67	CAH2bov-Q	P	79	Y	Y	Acetyl	CAH2bov-Q	P	77	Y	Y	Acetyl	CAH2hum	P	77	Y	Y	Acetyl
98186	Y	T, Ch	14	66	13	65	13	66	CAH2bov-Q	P	65	Y	Y	Acetyl	CAH2bov-Q	P	65	Y	Y	Acetyl	CAH2bov-Q	P	66	Y	Y	Acetyl	CAH2bov-Q	P	66	Y	Y	Acetyl
20702	Y	T	15	63	15	63.2	15	59.6	CAH2hum	P	63	Y	Y	Acetyl	CAH2bov-R	P	59	Y	Y	Acetyl	CAH2bov-Q	P	60	Y	Y	Acetyl	CAH2hum	P	60	Y	Y	Acetyl
27974	Y	T, Glu-C, Ch	12	57	12	58	12	43	2	11	2	12	2	20	CAH2bov-R	P	57	Y	Y	Acetyl	CAH2bov-Q	P	58	Y	Y	Acetyl	CAH2hum	P	47	Y	Y	Acetyl
82013	Y	T, Glu-C, Ch	11	58	11	58	12	51	6	28	6	28	6	28	CAH2bov-R	P	58	Y	Y	Acetyl	CAH2bov-Q	P	58	Y	Y	Acetyl	CAH2hum	P	51	Y	Y	Acetyl
4343	Y	T	14	52	14	53	13	52	CAH2bov-Q	P	52	Y	Y	Acetyl	CAH2bov-Q	P	52	Y	Y	Acetyl	CAH2bov-Q	P	52	Y	Y	Acetyl	CAH2bov-Q	P	52	Y	Y	Acetyl
29103	Y	T	7	35	10	40	6	32	3	9	1	8	CAH2bov-R	P	43	Y	Y	Acetyl	CAH2bov-Q	P	40	Y	Y	Acetyl	CAH2bov-Q	P	39	Y	Y	Acetyl		
11113	Y	T	4	22	3	20	3	20	CAH2bov-Q	P	22	Y	Y	Acetyl	CAH2bov-Q	P	20	Y	Y	Acetyl	CAH2bov-Q	P	20	Y	Y	Acetyl	CAH2bov-Q	P	20	Y	Y	Acetyl
11010*	Y	T	8	30	7	34	0	0	6	28	5	20	4	15	CAH2bov-R	P	55	Y	Y	Acetyl	CAH2bov-R	P	49	Y	Y	Acetyl	CAH2horse	T	31	Y	Y	Acetyl
1096352	Y	Lys-C	3	13.9	0	0	0	0	18	42	13	45.9	18	42	CAH2bov-Q	P	42	Y	Y	Acetyl	CAH2hum	P	45.9	Y	Y	Acetyl	CAH2bov-Q	P	42	Y	Y	Acetyl
21952	Y	T	18	70	14	58	17	63	5	25	5	25	CAH2bov-Q	P	73	Y	Y	Acetyl	CAH2hum	P	69	Y	Y	Acetyl	CAH2bov-Q	P	69	Y	Y	Acetyl		
25519	Y	T	12	59	13	56	1	1	15	73	16	64	CAH2bov-R	P																		