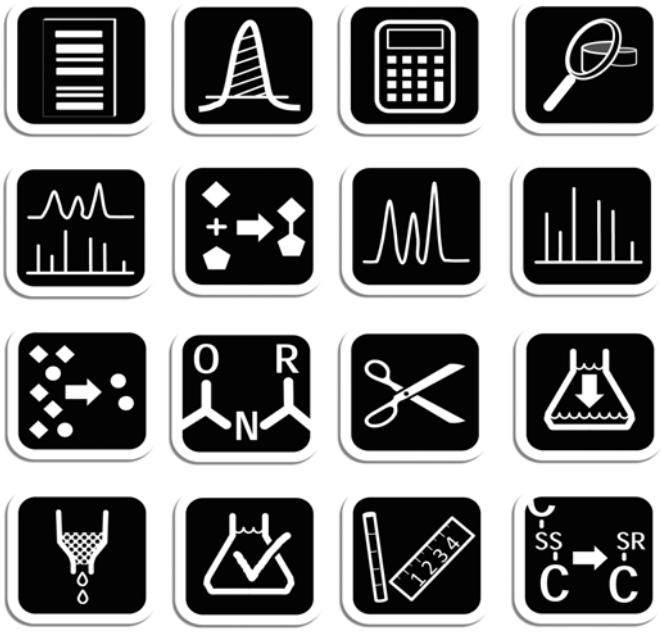
An overview of the Proteomics Resource Center at The Rockefeller University

September 2011 to July 2013



By Henrik Molina. July 2013

An overview of the Proteomics Resource Center at The Rockefeller University.

Contents.

Overview.

The Proteomics Resource Center (PRC) at The Rockefeller University offers analysis and synthesis of biomolecules for The Rockefeller University community on a subsidized, cost-recovery basis. The Center is also available to researchers from other academic institutions.

Though the main focus is identification of proteins and peptides by bottom-up tandem mass spectrometry analysis of other biomolecular compounds - for example metabolites and drugs - are within the scope of the Proteomics Resource Center. The peptide synthesis laboratory synthesizes single peptides with or without modifications and/or stable isotope labeled residues as well as peptides for arrays.

The Center currently has a staff of 5 (see **Table 1**) and houses several mass spectrometers, liquid chromatographic setups and peptide synthesizers (**Table 2**, following page). Selected instruments housed in the Center are available for walk-up use, marked [U] in **Table 2**.

Table 1. Staff.

Name	Expertise	Title
	Peptide synthesis	Res. Support Specialist THIS POSITION HAS BEEN ELIMINATED AS OF JULY 1 2013.
	Proteomics	Associate Director
	Proteomics	Director
	Admin./Peptide synthesis	Admin. Assistant
	Proteomics	Res. Support Specialist
	Peptide synthesis	Res. Support Specialist
	Proteomics	Res. Support Specialist

Table 2. Instrumentation overview.

Equipment	Vendor	Model	Type	Flow	Received	Age¤	Units
MS*	Thermo	Q-Exactive	Orbitrap	Nano	Feb2012	1	1
MS*	Thermo	Vantage	Triple Quad.	Low micro	July2008	5	1
MS*	Thermo	LTQ	Linear Ion Trap	High micro	June2006	7	1
MS*	Thermo	Orbitrap XL+ETD	Hybride ion trap/Orbitrap	Nano	Mar2008	5	1
MS*[U]	AB Sciex	Voyager STR	MALDI	-	Feb2004	10+	1
MS	AB Sciex	QSTAR Elite	Quad. ToF	NanoES	Apr2004	8	1
LC*	Dionex	HPLC	Semi-prep HPLC	mL	Apr2008	5	1
LC*	Waters	HPLC	Prep. HPLC w/auto sampler & Fraction collector HPLC	mL	July2008	5	1
LC*	Waters	UPLC	Analytical w/auto sampler	mL	Apr2007	15+	1
LC[U]	Agilent	HPLC		Micro	?	15+	1
LC*[U]	GE Health.	AKTA Purifier	FPLC	mL	Donated in 2008	10+	1
Scanner*[U]	GE Health.	Typhoon 9400	Scanner	-	Sep2003	9	1
Off Gel[U]	Agilent	G3100A	Off-gel fractionator	-	June2010	3	1
Peptide Syn.*	Protein Tech.	Rainin Symphony	12 port Synthesizer	-	Nov2001, Nov2003	10+	2
Peptide Syn. *	CEM Corp.	Liberty pep. syn.	Microwave Peptide Synthesizer	-	Mar2006	6	1
Peptide Syn.	Intavis LLC	MultiPep system	96 well Peptide Synthesizer	-	Oct2004, Aug2007	6+	2
Lyophilizers	-	-		-	1989,1993, 2005	10+	3

*: On service contract. [U]: User instrument. ¤ : Approx. age in years in January 2013.

Proteomics Laboratory Information Management System (pLIMS).

With many new and different tools used by the Center (**Table 3**, page 10) strong interaction with the users has become increasingly important and we have implemented that the staff interviews the users and keeps track of information related to sample matrix and objective. For example: samples of lower amounts should be analyzed before samples of higher amounts to minimize potential carry-over in-between samples, detergent(s) in samples can make in-solution based strategies impossible, and if a sample contains primary amines desalting is required prior to chemical labeling.

For a proteomics operation with great sample variability, structure is needed. Together with the RU Data Center a web based database structure has been designed. The system is named pLIMS, where 'p' is for proteomics and LIMS is short for Laboratory Information Management System. Though our pLIMS is not designed to track all steps in the experimental process I still use this term to describe the system that allows us to log user and selected sample relevant information for all projects:

- User identifying information is added to the system (affiliation, name, laboratory, address) and a PRC scientist is assigned to the project.
- The sample matrix is defined (gel, in-solution, solid, FASP)
- The user is interviewed re. experimental details exemplified by: Target taxonomy?, Potential contaminating taxonomies?, Is reduction & alkylation of cysteine's required?, Expected complexity?, Sample composition? and Amounts?
- Type of analysis: simple LC-MS(/MS), MALDI, MRM, quantitation etc.
- Additional information, for example a gel picture, can be uploaded to the pLIMS.

The pLIMS also allows for creating sample labels which is particular useful when dealing with an average of 300+ samples per month.

Because all user information is logged in the pLIMS and because the pLIMS allows us to document the number of different tools used for an analysis/experiment, it is possible to use the system for billing. This saves time and structures the billing process. Shown in **Figure 1** is the pLIMS 'home page' where projects are listed and marked according to process status. Examples of a pLIMS submission report and an invoice is available as **Appendix 2** and **Appendix 3**, respectively.

Should this system be successful, a module related to actual sample can be added.

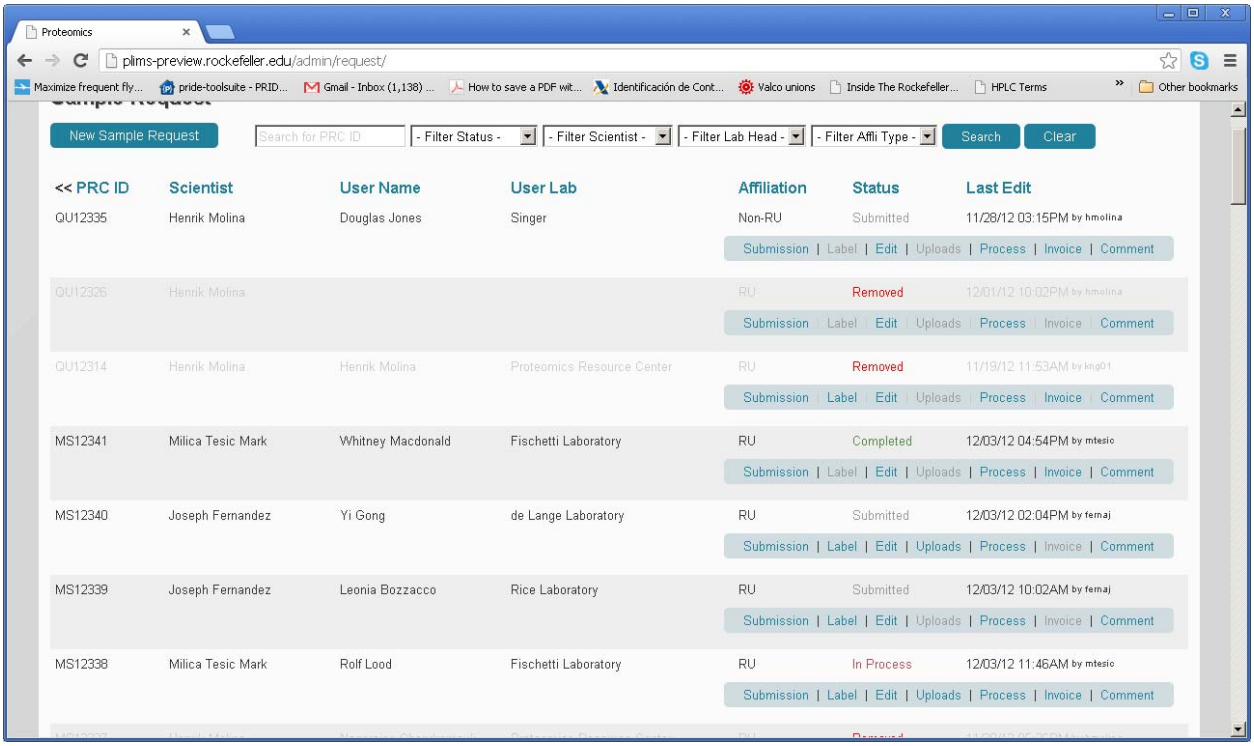


Figure 1. Screen shot from pLIMS. 'Home screen' with projects listed.

An important feature of the pLIMS system is that it allows to better estimate usage. In the past only the number of projects were recorded and because projects ranges from one to more than 100 samples this measure is not useful when trying to actual usage of the Center.

Data structure and website.

The website of the Proteomics Resource Center has received a major revamping (http://inside.rockefeller.edu/proteomics/). During this process the site was moved to the RU Data Centers Do-It-Yourself platform. This allows us to quickly update the PRC website which is useful to let users know changes in the availability of user instruments.

Needless to say, the Center creates large amounts of mass spec data files that needs to be searched and stored. The previous data infrastructure was old and a new structure based on ProteomeDiscoverer

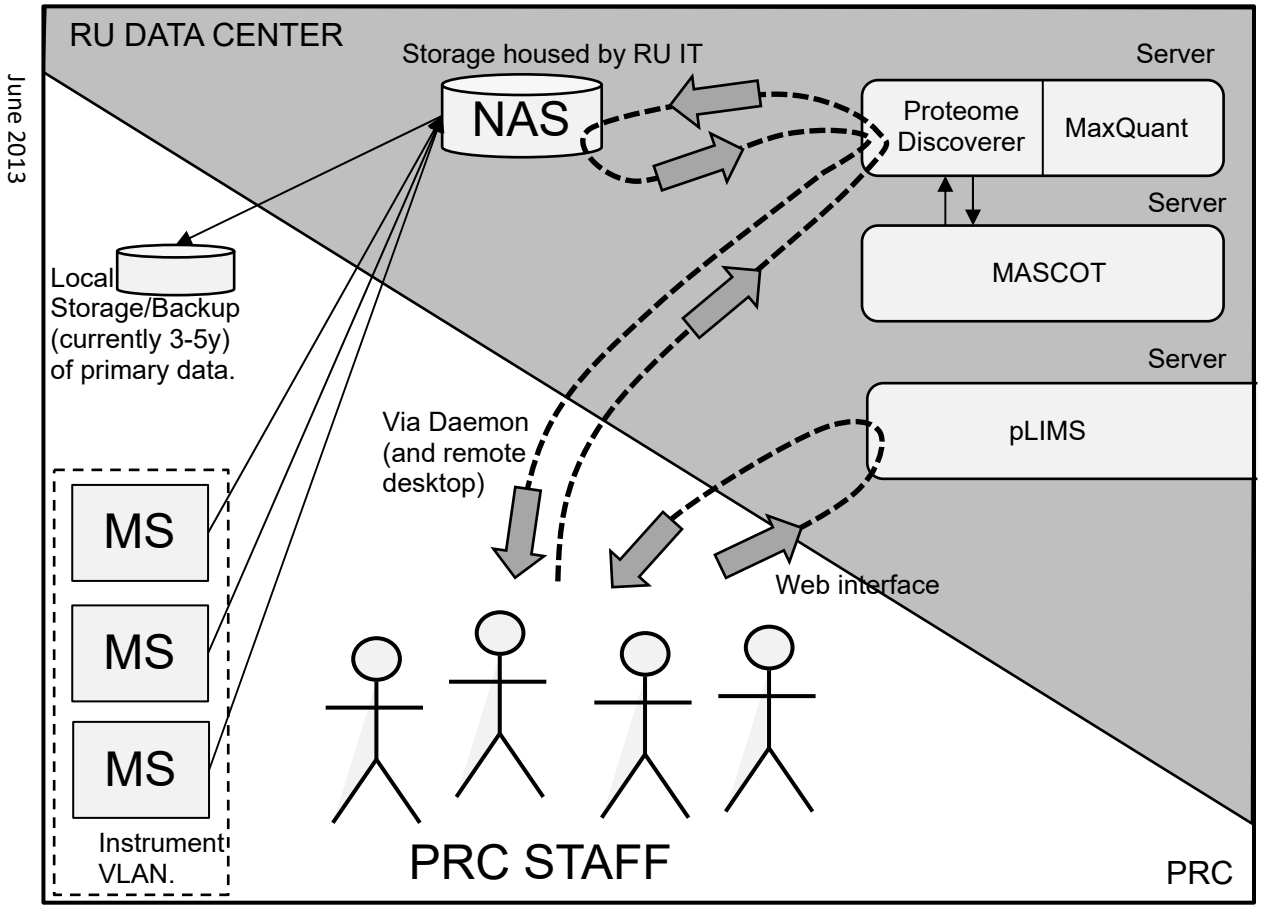


Figure 2. Schematics of data infrastructure.

(Thermo) and Macot (MatrixScience) has been designed, purchased and implemented. The ProteomeDiscoverer software (MS/MS extraction, quantitation and sorting) is expensive ($15k) and in an effort to save cost I have installed this platform on a server bought by the PRC but housed by the RU Data Center. This makes it possible to operate with only one license – but capacity is at its limit.

In May 2013 a simple RAID5 system was purchased and installed and is now functioning as a storage backup of primary data stored on the RU Data Center.

The PRC data structure is illustrated in **Figure 2** (previous page).

Instrumentation.

One new mass spectrometer (Q-Exactive) was added to the Center in early 2012. So far this instrument has performed well though this more sensitive platform requires more attention. The Q-Exactive is mainly used for very complex samples and for PTM type analysis.

A request from a user to use HPLC resulted in establishing this service using an older Waters system (with auto sampler but not fraction collector). The user lost interest but the Center can now also offers an HPLC as a user instrument. This system also serves as back-up for peptide purification.

The user MALDI mass spectrometer has for years been on a full vendor based service contract ($30k). The cost recovery was approximately half of the service contract cost and a large part of the MALDI users were non-RU. Last year the full service contract was downgraded to a pay-per-visit contract. This still proved expensive (compare to cost recovery). This summer we signed a contract with a 3rd party service provider that offers full service and one extra PM for ~$17k/year which is close to cost recovery for this instrument. So far this arrangement has worked well. Via the 3rd party vendor the operating system of the computer controlling the MALDI was upgraded ($6k) and is now running Win XP Professional which is a major improvement in stability.

A list of The Proteomics Resource Center instrumentation is presented in **Table 2**.

Typical and not so typical projects.

With 500+ projects being processed in the Center this year the Center has covered much ground. Though most of the projects are comparisons of two or more samples some project stands out. Examples are:

- Identification of regulated o-GlcNAc modified proteins using label free, dimethyl and TMT based quantitation.
- Pulse-chase type studies in SILAC mice focused at specific histones .
- Measuring mutations effect on the stoichiometric of a 5-protein-complex.
- Synthesis of a serine phosphorylated and bitonylated 33-mer peptide.

- Epitope mapping project where 330 peptides were made to develop a conjugate vaccine against group A *Streptococcus* targeting M protein.

Concerns.

With 300+ proteomics projects in one year and only three staff it is very demanding and challenging to be on-top and fully understand all projects. Though we are striving to have good communication with the users mistakes and misunderstanding are bound to happen which is a concern to me. I do think that this problem will decrease over time because of gained experience. In the proteomics group we meet weekly meeting with the purpose of discussing specific samples and share our experience.

A more technical concern is related to actual sample acquisition and analysis. Some samples deserve more attention than available resources allow. Also, with the large range of samples being analyzed by the same instrumentation, carry-over is happening though we are trying our very best to minimize this by prioritize samples based on amounts.

Though I mention that the above is concerning to me, I am sure that this is a valid problem for nearly all proteomics laboratories. One of my goals for the coming year is address carry-over by testing the use of using multiple columns and extra washing cycles in-between samples.

Usage and Finance.

The Center has since 2004 been subsidized with between 60% and 80% per year. Such subsidy rates might seem high but is related to that proteomics requires many staff hours - different from e.g. microscopy where a very large fraction of the experiments are carried out by the users them self. 70% of the cost associated with the Proteomics Resource Center is salaries. Service contracts are also a heavy burden which currently is ~25% of total expenses. Total spending's and recoveries are shown for FY07 to FY13 in **Figure 3**. Subsidy for FY13 was 60.7%.

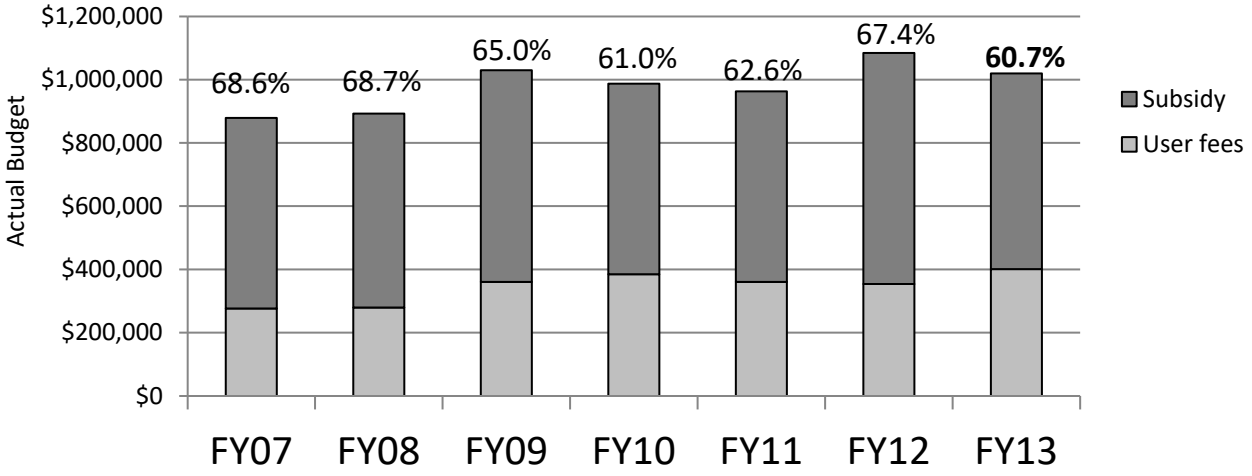


Figure 3. Actual budget and subsidy for Fiscal Year (FY) 2007-2013. Average subsidy of 64.9% for FY07-13.

The number of laboratories using the Center per quarter has been constant. Close to half of the laboratories at The Rockefeller University use the Center in a period of 3 months - an increase compared to the past (**Figure 4**). In the Financial year 2012 a total of 58 RU labs used services provided by the Proteomics Resource Center. For FY13 this number was 47.

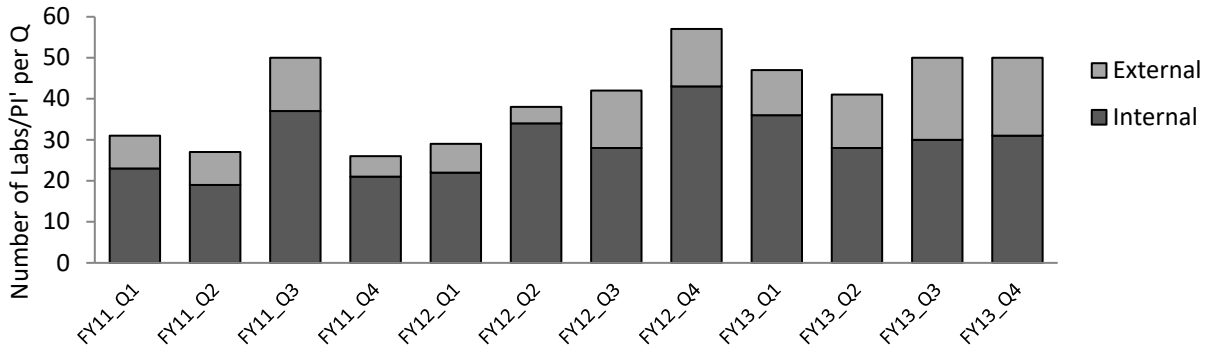


Figure 4. PRC usage for the period August 2010 to July 2013. The numbers represent the number of individual internal (dark grey) and external (light grey) laboratories/PI's and is shown.

In **Figure 5** (following page) are shown historical cost recovery numbers for the Center as whole (top panel) and separated into the peptide (mid panel) and proteomics (lower panel) group. Internal and external user fees are summed. During the past year we have seen a steady increase of proteomics projects. Approximately 350 proteomics/small molecule projects where processed in the calendar year 2012. The average number of samples in a project was 10 with a median of 6. Moving more towards in-solution samples (before September 2011 90% of all samples were protein identification from gel bands) I am expecting a decrease in number of samples per project. In FY13 Q4 the average number of samples in a project was 8 (with a median of 4) and I foresee a shift towards more computational work and less time spend at the bench, which can impact recovery because analysis is only partly charged (see also page 11).

Cost recovery related to peptide synthesis has declined and peptide projects (~200 in 2012) are down approximately 50% compared to some of the best years. The decline is likely related to the closure of the Steinman Lab following Ralph's death. Because of this decline one peptide position was eliminated by July 2013.

Tools and new user fee structure.

During the past 18 months several new proteomics tools has been implemented. Because gel-bands were the preferred sample matrix for the Center in the past, new tools and concepts required a new

user fee structure. New implementations include dimethyl labeling, tandem mass tags together with micro fractionation strategies. **Table 3** (page 10) summarizes tools with respective fees, offered by the

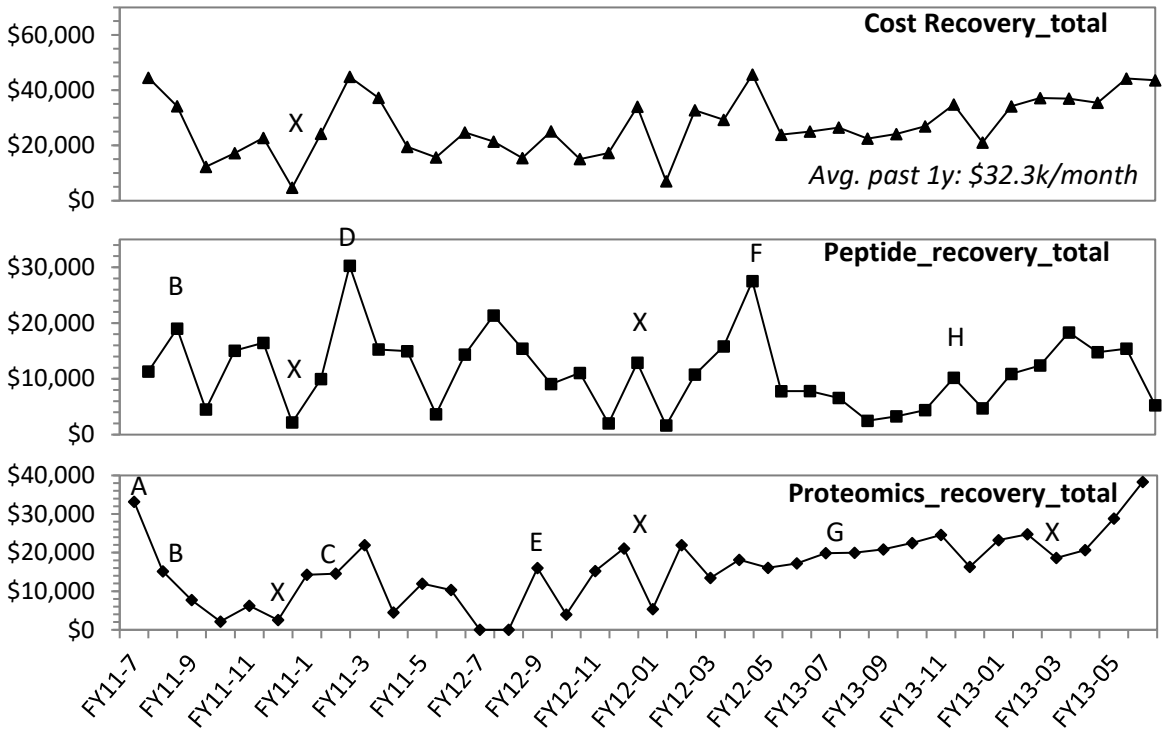


Figure 5. PRC Cost Recovery and usage for the period August 2010 to July 2013. Some extremes explained: **A:** Haiteng (STARR grant) purchases $21k worth of MS service - **B:** 2x $10k worth of peptides and MS analysis to 'Luo@MSKCC' (STARR grant) at internal fee rate. **C:** Haiteng departs. - **D:** $19k peptide order from OCME / NYC (external) - **E**: Henrik arrives. - **F:** $7.5k peptide order from Zhao/Rumah/ Vartanian/CUMC . - **G:** New user fee structure - **H:** One peptide library worth $8k. - **X:** Holiday breaks.

Center. FY13 and suggested FY14 fees are shown. The user fees have been made modular which makes it possible to calculate fair fees for experiments. For example, in the past it was difficult to charge for replicated LC-MS/MS analysis – and technical replicates are often very useful – but in the modular system this is possible. The FY12 and FY13 prices was designed with the previous gel-band price in mind meaning that one gel band before and after the fee structure change has been kept constant. For FY14 I have suggested an internal fee increase of ~11% for the three most used services.

Even with the suggested increase the internal fees are still competitive according to a recent survey (**Appendix 1**). For the survey I compared 11 academic proteomics centers by checking fees listed on the respective centers website. It must be noted that because different centers have different fee structure/model it can be very difficult to make an 'apples-to-apples'. For the comparison I investigated internal and external pricing for the analysis of a simple protein gel band and for the analysis of a complex sample – our most common services.

The peptide user fee structure is unchanged except for the usual 3% increase.

A special part of the new user fee structure is a Do-It-Yourself (DIY) option. As mentioned previously labor is by far the largest component of the budget – which is why sample processing that requires staff (digestion, chemical modification, micro fractionation etc.) is set higher and LC-MS/MS, that require less human interference, is set lower. The DIY option relies on training of users with respect to e.g. digestion and only charge for LC-MS/MS and chemicals. In short: the DIY option makes sense for larger project where costs are lowered and in addition benefits with having users trained in pre-MS sample handling.

Table 3. Tools and Modular User Fees from FY13 and for FY14 (in bold).

Process	**Internal User Fee (FY13)** ->FY14 price	**External User Fee* (FY13)** ->FY14 price
Digestion (trypsin, Lys-C, Chymotrypsin)	$60-**>$65**	$90-**>$110**
Digestion, large (up to 400ug of protein)	$80-**>$90**	$120-**>$140**
Filter Assisted Sample Preparation		
LC-MS(/MS) per hour (increments of 1h)	$35-**>$40**	$70-**>$90**
QSTAR/MALDI full service per sample	$50#-**>$55**	$75-**>$90**
µ-Solid Phase Extraction (SPE)	$20-**>$25**	$30-**>$40**
miligram amount Solid Phase Extraction (SPE)	$30-**>$35**	$45-**>$55**
µ-Fractionation	$250-**>$260**	$375-**>$400**
Dimethyl (up to 3 states)	$120-**>$120**	$180-**>$190**
TMT per label (up to 6 states)	$120-**>$130**	$180-**>$200**
Phospho peptide enrichment (IMAC)	$80-**>$100**	$120-**>$150**
Phospho peptide enrichment (TiO)	$80-**>$100**	$120-**>$150**
Reduction & Alkylation	$20-**>$22**	$30-**>$40**
Silver destaining	$20-**>$22**	$30-**>$40**
Data base search only (per search)***	$20-**>$25**	$40-**>$50**
Custom, e.g. focused analysis or setting up small molecule analysis.	Min. **$100**, but based on time used	Min. **$200,** but based on time used
Chemicals etc. purchased via PRC.	At cost	-
Large/custom data analysis**	Collaboration	-

*) For **Non-RU/external** a start-up fee of $100 is applied per project.
#) Startup fee of $50 is applied for internal users.
**) Some data analyses are very simple (e.g. 'identification only') while others can be time consuming. While standard data analysis is included for all protein identifications, custom and lengthy analyses are not. The latter is typically analyzed on a collaborative basis.
***) Standard data analysis is included in *full service* (digestion+LC-MS/MS) protein identification.

The PRC staff provides a significant amount of intellectual contribution to a number of projects. While I have not formally evaluated this, I estimate that 20% of the proteomics samples processed by the Center take up 80% of the staff time. The 20% includes samples that require very special experimental design and planning, analysis different from simple lists of identified proteins or for example obtaining and creating publication quality evidence for post-translational modified residues. Based on the trend I have observed the last 16 months I sense that the above mentioned 20% will only increase….and that this is a good thing. Assessing user fees for these intellectual contributions is difficult and deemed counterproductive, i.e., it would discourage researchers from taking advantage of the most important resource the PRC has to offer, namely the professional staff's expertise. The University is aware of this and that it is a primary reason that the subsidy level for the PRC is as high as it is.

It is important to point out that these intellectual contributions should be considered in light of widely accepted guidelines on co-authorship. Recognition of the Resource Center's contributions, through co-authorship when it is warranted, and/or in publication acknowledgements does not only meet these widely accepted guidelines but enhances the perception of the Center which can be beneficial to future grant applications that reference the Center and to future development proposals seeking funds for the Centers.

APPENDIX 1. Fee comparison.

Price/Fee analysis for two common types of services. Analysis was conducted between July 9 and July 10, 2013

					Simple gel band		In-solution, more complex	
#	URL	Name	Institution	Location	Internal	External	Internal	External
1	https://taplin.med.harvard.edu/prices	Taplin¤	Harvard	Boston, MA	$60/$140	$80/$160	$300	$320
2	http://www.molbio1.princeton.edu/synsec	Princeton Proteomics & Mass Spectrometry Core Facility	Princeton University	*Princeton, NJ*	$95-$275	?	?	?
3	http://ocs.med.nyu.edu/proteomics	NYU	NYC	New York, NY	$116	$175	$127	$191
4	**http://proteomics.rockefeller.edu/prot**	**Proteomics Resource Center**	**The Rockefeller University**	**New York, NY**	**$105**	**$200**	**$232**	**$370**
5	http://www.genome.duke.edu/cores/prote	Duke Proteomics Core Facility	Duke University	Durham, NC	$110	$173	$280	$440
6	http://www.brc.cornell.edu/brcinfo/?p=pri	Biotechnology Resource Center	Cornell University	Itacha, NY	$150	$280	$280	$400
7	http://www.osa.sunysb.edu/Proteomics/fi	Proteomics Center	SUNY Stony Brook	Stonybrooke	$150	?	?	?
8	http://medicine.yale.edu/keck/proteomics	Keck¤	Yale	New Haven, CT	$165	$198	$350	$420
9	http://www.dfhcc.harvard.edu/core-faciliti	Cancer Proteomics¤	Dana-Faber Harvard Cancer Center	Boston, MA	$165	$248	?	?
10	http://proteomics.ucdavis.edu/protein-ide	Proteomics Core Facility¤	UC Davis	Davis, CA	$169	$259	169?	259?
11	http://njms.rutgers.edu/proweb/	Center for Advanced Proteomics Research	Rutgers, New Jersey Medical School	Newark, NJ	$200	?	$275	
12	http://mass-spec.stanford.edu/Rates.htm	Standford University Mass Spectrometry¤	Standford University	Standford, CA	$216	$288	$288	$384
13	http://www.mskcc.org/research/microche	MSKCC	MSKCC	New York, NY	?	?	?	?

This analysis does not take start-up fees or minimum sample numbers into account. Not all the centers have fee structures that is easy to decipher and numbers are best assessment.
¤) Indicates that the Center might allow For-Profit users.
#12-13) Note: It has not been possbile to find public avaiable pricing informastion from

APPENDIX 2. Sample submission report example.

THE PROTEOMICS RESOURCE CENTER | p +1 212 327 8446 f +1 212 327 8620
THE ROCKEFELLER UNIVERSITY| 1230 YORK AVE, NEW YORK, NY 10065, USA

MS12343

SUBMIT DATE - Dec 4, 2012
PRC SCIENTIST - Henrik Molina

NUMBER OF SAMPLES - 6

AFFLIATION - The Rockefeller University
USER NAME - Natalia Frias-Staheli
USER EMAIL - nfrias@rockefeller.edu
USER PHONE - 212-327-7054

TAXONOMY - Homo sapiens (human)

CHEMICAL MODIFICATIONS - no

SAMPLE TYPE - Gel band
STAINING - Silver; Coomasie blue;
EXPECTED COMPLEXITY - Estimated complexity: <50;

REQUEST TYPE - Proteomics
ANALYSIS TYPE - ID by LC-MS/MS;
ID by LC-MS/MS comment - Amounts are high for sample 2. For this comparison (incl. the control) only a fraction (e.g. 1/5) should be injected. 1/3 for the silver stained samples.

DATA ANALYSIS - Each LC-MS/MS analysis searched independently;
SAMPLE NAMES - band_4; band_2A; band_2B; band_3; band_1A; band_1B

NAME OF FILES UPLOADED - Natalia_gel_pic.pdf

APPENDIX 3. Invoice example.

12/11/12 Proteomics

THE PROTEOMICS RESOURCE CENTER | p +1 212 327 8446 f +1 212 327 8620
THE ROCKEFELLER UNIVERSITY| 1230 YORK AVE, NEW YORK, NY 10065, USA

INVOICE: MS12336 **SUBMIT DATE** - Nov 28, 2012 **COMPLETED** - Dec 1, 2012 **INVOICED** - Dec 6, 2012

The Rockefeller University
Attn: Accounts Payable, Box 259
1230 York Avenue
New York City, NY 10065
U.S.A.

USER NAME - Beatrix Ueberheide
USER AFFILIATION - New York University Medical Center
USER EMAIL - Beatrix.Ueberheide@nyumc.org
USER PHONE - 212-263-2546
USER LAB - Ueberheide

NUMBER OF SAMPLES - 2 **Total cost:** $ 360

Payment Type Information - Grant 1010700-01-OSU1013022-64110-500000

Accepted payment types includes: RU Lab Account, Credit Card, Purchase Order (PO), and Check.
Please make check payable to: The Rockefeller University and mail to: The Rockefeller University / Proteomics Resource Center, Box 105 / 1230 York Avenue / New York, N.Y. 10065-6399 The Rockefeller University is a not-for-profit insitution. Tax Exempt No. 131624158

Process	Comments	Price/Unit	Actual #	# Units Charged
Administration/Start-up Fee		$ 150	0	1
LC-MS(/MS) per hour (increments of 1h)		$ 70	0	3
				Total: $ 360

plims-preview.rockefeller.edu/admin/request/invoice/?id=336 1/1