Computational Biology in Proteomics

Zohreh Foladi Dehaghi
Endocrine research center, shariati hospital, Damghan university, Tehran, Iran

* Corresponding Author: **Zohreh Foladi Dehaghi**

Article Info

E-ISSN: 3107-7137
Volume: 02
Issue: 03
Received: 08-03-2026
Accepted: 06-04-2026
Published: 04-05-2026
Page No: 08-12

Abstract
The broad application of proteomics in different biological and medical fields, as well as the diffusion of high-throughput platforms, leads to increasing volumes of available proteomics data. Computational proteomics is the data science concerned with the identification and quantification of proteins from numerous data sources and the biological interpretation of their concentration changes, posttranslational modifications, interactions, and subcellular localizations. Computational proteomics is a highly multidisciplinary endeavor attracting scientists from many fields and incorporates other disciplines like statistics, machine learning, efficient scientific programming, and network and time series analysis.

Keywords: Computational Proteomics, Proteomics Data Analysis, High-Throughput Proteomics, Bioinformatics

Introduction

Proteomics is a fundamental science in which many sciences in the world are directing their efforts. The proteins play a key role in the biological function, and their studies make it possible to understand the mechanisms that occur in many biological events (human or animal diseases, factors that influence plant and bacterial growth). Due to the complexity of the investigation approach that involves various technologies, a high amount of data is produced. In fact, proteomics has known a strong evolution, and now we are in a phase of unparalleled growth that is reflected by the amount of data generated from each experiment. That approach has provided, for the first time, unprecedented opportunities to address the biology of humans, animals, plants, as well as microorganisms at the system level. Bioinformatics applied to proteomics offered the management, data elaboration, and integration of this huge amount of data. It is with this philosophy that this chapter was born. Thus, the role of bioinformatics is fundamental in order to reduce the analysis time and to provide statistically significant results. To process data efficiently, new software packages and algorithms are continuously being developed to improve protein identification, characterization, and quantification in terms of high-throughput and statistical accuracy. However, many limitations exist concerning bioinformatics spectral data elaboration. In particular, for the analysis of plant proteins, extensive data elaboration is necessary due to the lack of structural information in the proteomic and genomic public databases. The main focus of this chapter is to describe in detail the status of bioinformatics applied to proteomic studies. Moreover, the elaboration strategies and algorithms that have been adopted to overcome the well-known limitations of protein analysis without database structural information are described and disclosed. This chapter will shed light on recent developments in bioinformatics and data-mining approaches, and their limitations when applied to proteomic data sets, in order to reinforce the interdependence between proteomic technologies and bioinformatics tools. Proteomic studies involve the identification as well as qualitative and quantitative comparison of proteins expressed under different conditions, together with description of their properties and functions, usually in a large-scale, high-throughput format. The high dimensionality of data generated from these studies will require the development of improved bioinformatics tools and data-mining approaches for efficient and accurate data analysis of various biological systems (for reviews, Li *et al*, 2009; Mathieson & Jensen, 2008; Wright *et al*, 2009). After a rapid move on the wide theme of the genomic and proteomic sciences, in which bioinformatics find their wider applications for the studies of biological systems, the chapter will focus on mass spectrometry, which has become the prominent analytical method for the study of proteins and proteomes in the post-genome era. The high volumes of complex spectra and data generated from such experiments represent new challenges for the field of bioinformatics. The past decade has seen an explosion of informatics tools targeted towards the processing, analysis, storage, and integration of mass spectrometry-based proteomic data. In this chapter, some of the more recent developments in proteome informatics will be discussed. This includes new tools for predicting the properties of proteins and peptides, which can be exploited in experimental proteomic design, and tools for the identification of peptides and proteins from their mass spectra. Similarly, informatics approaches are required for the move towards quantitative proteomics, which are also briefly discussed. Finally, the growing number of proteomic data repositories and emerging data standards developed for the field are highlighted. These tools and technologies point the way towards the next phase of experimental proteomic and informatics challenges that the proteomics community will face. The majority of the chapter is devoted to the description of bioinformatics technologies (hardware, data management, and applications) with particular emphasis on the bioinformatics improvements that have made it possible to obtain significant results in the study of proteomics. Particular attention is focused on the emerging statistical semantics, network learning technologies, and data sharing that is the essential core of system biology data elaboration. Finally, many examples of bioinformatics applied to biological systems are distributed along the different sections of the chapter to lead the reader to completely understand the benefits of bioinformatics applied to system biology (Creston *et al*, 2011). Literature review. One research study by Kontiievskis and their team shows that HIV-1 protease specificity is much more complex than previously anticipated, which cannot be defined based solely on the amino acids at the substrate's scissile bond or by any other single substrate amino acid position. Our results show that the combination of at least three particular amino acids is needed in the substrate for a cleavage event to occur. Only by combining and analyzing massive amounts of HIV proteome data was it possible to discover these novel and general patterns of physicochemical substrate cleavage determinants. Our study is an example of how computational biology methods can advance the understanding of the viral interactomes (kontiievskis and *et al* 2007). The work done by Professor Smedley in 2022 has shown that it is possible to analyze whole genomes across a broad range of rare diseases and find a diagnosis in about 25 percent of people. Now that we have the infrastructure in place for this, the plan is to continue to build on it, incorporating new knowledge from researchers worldwide to try to increase the number of patients that can be helped. Part of this will be the application of new computational approaches to extend our search into the non- coding parts of the genomes, as well as larger structural variants. For example, recently published approaches to predicting intrinsic variants that affect splicing will increase our potential to diagnose more patients and include innovative new artificial intelligence approaches. Having a single healthcare system has made it easier for us to apply WGS diagnostics at scale in the UK. Our findings have already had an impact on our National Health Service (NHS) and will continue to do so with ambitions to sequence around 300,000 genomes outlined in the NHS Long Term Plan. However, we hope that our study will not only transform the UK health system but also be adopted by other healthcare systems to change the lives of rare disease patients worldwide (Ellis *et al* 2022) [3]. The research guided by Dr. Bose and her group demonstrates that for sickle cell disease (SCD), hydroxyurea is an established treatment. Despite this, hydroxyurea remains underutilized. New disease-modifying treatments (DMTs), namely, crizanlizumab, L- glutamine, and voxelotor, have recently been approved. A new JAMA Network Open study delved deeper into understanding the uptake of these DMTs in the context of current treatment strategies (Dr. Bose *et al* 2023) [4].

Project Narrative

There is a method for evaluating computational biology in a proteomics project that is shown in this table.

Table 1: Pros and cons of different computational strategies

Computational Method	Data Type	Database Dependency	Accuracy Level	Protein Mixture Suitability	Modifications Handling	Speed	Key Pros	Key Cons
Database-dependent PMF search	MS	Dependent	Accurate	Simple mixtures	Few modifications	Fast	Fast interpretation, reliable for known/simple samples	Limited modifications, not suitable for complex mixtures
Database-dependent PFF search	MS/MS	Dependent	Accurate	Complex mixtures	Moderate number of modifications	Moderate	Handles complex samples better than PMF	Slower than PMF, still database-limited
Sequence tag + database search	MS/MS	Dependent	Moderate accuracy	Complex mixtures	Many modifications can be included	Fast	Flexible with modifications, relatively fast	Lower accuracy compared to full database searches
De novo sequencing	MS/MS	Independent	Less accurate	Complex mixtures	Few modifications	Fast	No database required, useful for novel proteins	Lower accuracy, challenging interpretation
Spectral library search	MS/MS	Dependent	Highly accurate	Complex mixtures	Few modifications	Fast	Very high accuracy, rapid matching	Limited to existing spectra, less flexible

All Proposal Budgets Must Be Reviewed And Approved By The Office Of Researc and sponsored programs before being submitted to the sponsor.

The following budget outline is a generalized format that will provide a good basis for most proposals. Actual budget categories will vary by sponsoring agency.

1. Personnel: List each position by title and name of employee. Using the annual salary rate and percentage of time to be devoted to the project, compute the salary to be paid to each individual; using established UM fringe benefits formulas shown below, compute benefits associated with each salary. An appropriate escalation rate (e.g., 2-5%) may be used to project salary rates beyond the initial fiscal year. No commitment by the institution and no constraint on the actual rate of increase for an individual is implied by this standard percentage increase procedure.

A. Senior Personnel

- principal investigator/project director
- co-investigators
- faculty/other senior associates

B. Other Personnel

- postdoctoral associates
- other professionals (technician, programmer, etc.)
- graduate students
- undergraduate students
- other [see Note 1]

C. Fringe Benefits

- project salaries for faculty/staff personnel above X, current faculty/staff fringe benefits rate*
- project wages for student personnel above X current student fringe benefits rate*
- *For current rates, see www.research.olemiss.edu/cms/toolbox/current_rates.

2. Travel: List travel expenses of project personnel by purpose and show the basis of computation. For conferences or training projects, travel costs for participants (other than UM employees) should be listed in Participant Support. For the UM travel policy, see www.olemiss.edu/depts/procurement/travel1.html.

- principal investigator/project director, co-investigators, staff
- conference registration fees
- transportation (air and/or ground transportation, parking, etc.)
- personal car mileage
- hotel
- meals

3. Participant Support: For training projects, conferences, workshops, symposia, etc., list number of participants and the unit cost for each type of expense.

- stipends
- travel
- subsistence
- other

4. Supplies: List materials and supplies, directly related to the project, that are expendable or consumed during the course of the project; office supplies are generally not allowed, but covered by F&A. [See Note 1]

- instructional supplies
- laboratory supplies

Equipment: List non-expendable items to be purchased ($5,000 or more per unit).

Expendable items and items costing less than $5,000 per unit should be included either in Supplies or Other. Rented or leased equipment costs should be listed in Contractual.

[See Note 2]

- any nonexpendable item costing $5,000 or more
- any of the following, regardless of cost: weapons, 2-way radio equipment, lawn, maintenance equipment, cellular telephones, chain saws, air compressors, welding, machines, generators, motorized vehicles

If purchase value is $250 and greater: cameras and camera equipment, televisions, computers and computer equipment

Contractual Services ~ List consultants, contracts, or subcontracts, and/or any Other services are to be purchased for the conduct of the project. [See Note 1]

- consultant services (including consultant's compensation rate, estimated days of service, expenses)
- subcontracts (including subcontractor's budget, and F&A costs or fees) [See Note 2]
- lease/rental of equipment
- service or maintenance contracts

Other Direct Costs: List costs not covered in other categories, including tuition charges for graduate assistants to be employed on the project and supported with grant funds.

- publication/documentation/dissemination
- graduate tuition remission* [See Note 2 and Note 4]
- other

*For current rate, see www.research.olemiss.edu/cms/toolbox/current_rates.

Facilities and Administrative Costs ~ Compute and include full F&A costs (indirect costs or overhead) According to UM's federally negotiated rates below, unless limited or prohibited by the sponsor (sponsor documentation required).

Effective through June 30, 2011

- Research - through 06/30/09 44.0% MTDC on campus 26.0% MTDC off campus
- after 07/01/09 43.0% MTDC on campus 26.0% MTDC off campus
- Instruction 50.0% MTDC on campus 26.0% MTDC off campus

- Other Sponsored Activity 28.2% MTDC on campus 16.6% MTDC off campus

[See Note 2] [See Note 2 and Note 3]

Equipment ~ List non-expendable items to be purchased ($5,000 or more per unit).

Expendable items and items costing less than $5,000 per unit should be included either in

Supplies or Other. Rented or leased equipment costs should be listed in Contractual.

[See Note 2]

- any nonexpendable item costing $5,000 or more
- any of the following, regardless of cost: weapons, 2-way radio equipment, lawn

maintenance equipment, cellular telephones, chain saws, air compressors, welding machines, generators, motorized vehicles; if purchase value is $250 and greater: cameras and camera equipment, televisions,

computers and computer equipment

Contractual Services ~ List consultants, contracts, or subcontracts, and/or any

Other services are to be purchased for the conduct of the project. [See Note 1]

- consultant services (including consultant's compensation rate, estimated days of service, expenses)
- subcontracts (including subcontractor's budget, and F&A costs or fees) [See Note 2]
- lease/rental of equipment
- service or maintenance contracts

Other Direct Costs ~ List costs not covered in other categories, including tuition

charges for graduate assistants to be employed on the project and supported with grant funds.

- publication/documentation/dissemination
- graduate tuition remission* [See Note 2 and Note 4]
- other

*For current rate, see

www.research.olemiss.edu/cms/toolbox/current_rates.

Facilities and Administrative Costs ~ Compute and include full F&A costs

(indirect costs or overhead) According to UM's federally negotiated rates below, unless

limited or prohibited by the sponsor (sponsor documentation required).

Effective through June 30, 2011

- Research - through 06/30/09 44.0% MTDC on campus 26.0% MTDC off campus
- after 07/01/09 43.0% MTDC on campus 26.0% MTDC off campus
- Instruction 50.0% MTDC on campus 26.0% MTDC off campus
- Other Sponsored Activity 28.2% MTDC on campus 16.6% MTDC off campus
- [See Note 2] [See Note 2 and Note 3]

5. Equipment: List non-expendable items to be purchased ($5,000 or more per unit).

Expendable items and items costing less than $5,000 per unit should be included either in

Supplies or Other. Rented or leased equipment costs should be listed in Contractual.

[See Note 2]

- any nonexpendable item costing $5,000 or more
- any of the following, regardless of cost: weapons, 2-way radio equipment, lawn

Maintenance equipment, cellular telephones, chain saws, air compressors, welding machines, generators, motorized vehicles; if purchase value is $250 and greater: cameras and camera equipment, televisions,

computers and computer equipment

6. Contractual Services: List consultants, contracts, or subcontracts, and/or any

Other services are to be purchased for the conduct of the project. [See Note 1]

- consultant services (including consultant's compensation rate, estimated days of service, expenses)
- subcontracts (including subcontractor's budget, and F&A costs or fees) [See Note 2]
- lease/rental of equipment
- service or maintenance contracts

7. Other Direct Costs: List costs not covered in other categories, including tuition

charges for graduate assistants to be employed on the project and supported with grant funds.

- publication/documentation/dissemination
- graduate tuition remission* [See Note 2 and Note 4]
- other

*For current rate, see

www.research.olemiss.edu/cms/toolbox/current_rates.

8. Facilities and Administrative Costs: Compute and include full F&A costs (indirect costs or overhead) According to UM's federally negotiated rates below, unless

limited or prohibited by the sponsor (sponsor documentation required).

Effective through June 30, 2011

- Research - through 06/30/09 44.0% MTDC on campus 26.0% MTDC off campus
- after 07/01/09 43.0% MTDC on campus 26.0% MTDC off campus
- Instruction 50.0% MTDC on campus 26.0% MTDC off campus
- Other Sponsored Activity 28.2% MTDC on campus 16.6% MTDC off campus
- [See Note 2] [See Note 2 and Note 3]

Notes

1. According to OMB Circular A-21, Federal Cost Principles for Educational Institutions, items such as

administrative and clerical staff, office supplies, postage, duplication, local telephone costs, and memberships should normally be treated as indirect costs. Direct charging of these items may be appropriate where a major project or activity explicitly requires such costs, and they can specifically be identified with the project or activity. In these cases, the costs should be sufficiently justified. These principles apply only to federal funding; non-federal agencies may allow such costs to be routinely charged as direct costs. See http://www.olemiss.edu/depts/contracts/direct%20cost%20charging%20policies.html.

2. Modified total direct costs (MTDC) consist of all salaries and wages, fringe benefits, materials and supplies, services, travel, and subgrants and subcontracts up to $25,000 of each subgrant or subcontract (regardless of the period covered by the subgrant or subcontract). Modified total direct costs shall exclude equipment, capital expenditures, charges for patient care and tuition remission, rental costs of off-site facilities, scholarships and fellowships, as well as the portion of each subgrant and subcontract in excess of $25,000. In addition, recharge center rates (e.g., Supercomputer, Teleproductions, Field Station, etc.) may have to be excluded; check with ORSP. See http://www.olemiss.edu/depts/contracts/fac%26admin.html.
3. Off-Campus definition applies to all activities performed in facilities not owned by the institution and to which rent is directly allocated to the project. Only one F&A cost rate will be used on a project budget. If more than 50% of a project is performed off-campus, the off-campus rate will apply to the entire project. See http://www.olemiss.edu/depts/contracts/fac%26admin.html.
4. Tuition remission for graduate assistants is REQUIRED on grants and contracts unless not allowed by the agency or program. Any variations must be approved by the Vice Chancellor for Research and Sponsored Programs. See http://www.research.olemiss.edu/cms/orsp/policies/tuition_charges.

Timeframe

This work can be done in Jolley 2024.

References

1. Cristoni S, Mazzuca M. Bioinformatics applied to proteomics. 2014;3.
2. Kontiievskis A, Wikberg JES, Komorowski J. Computational proteomics analysis of HIV-1 protease interactome. 2010;7.
3. Ellis D, *et al*. Computational biology in rare disease research. 2022.
4. Bose P, *et al*. Sickle cell disease patients and the use of disease-modifying treatments. 2023;23.

How to Cite This Article

Dehaghi ZF. Computational biology in proteomics. International Journal of Biological and Biomedical Research. 2026;2(3):8-12.

Creative Commons (CC) License