

The Trip Adviser guide to the protein science world: a proposal to improve the awareness concerning the quality of recombinant proteins

Mario Lebediker^{1,†}
Email: mario.l@mail.huji.ac.il

Tsafi Danieli^{1,†}
Email: tsafi.danieli@mail.huji.ac.il

Ario de Marco^{2*}
* Corresponding author
Email: ario.demarco@ung.si

¹ Protein Expression and Purification Facilities, Wolfson Centre for Applied Structural Biology, Givat-Ram Campus - The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

² Department of Biomedical Sciences and Engineering, University of Nova Gorica, Glavni Trg 9, SI-5261 Vipava, Slovenia

[†] Equal contributors.

Abstract

In many research articles, where protein purification is required for various assays, (protein-protein interactions, activity assays, etc.), we always have access to the final results, but seldom have access to the raw data required for an accurate evaluation of the protein quality. This data is extremely important on one hand to critically evaluate the quality of the proteins used in the described research and, on the other hand, to allow other laboratories to safely use the described procedure in a reproducible manner. We hereby propose to include a standardized methodology that can easily be incorporated in research papers. Moreover, this methodology can be utilized as a “quality control” ladder, where the more information given, will lead to a higher ranking of the article. This “quality control” stamp will allow researchers retrieving relevant and useful materials and methods in the field of protein research.

Keywords

Data annotation, Data reproducibility, Protein quality

Commentary

At least 50% of the published studies –even in the most prestigious journals- cannot be reproduced [1-8]. Designing, writing, reviewing, publishing, and referring to data obtained under misleading experimental conditions is clearly an expensive and unproductive procedure

for all the actors involved in the scientific system. Moreover, the accumulated errors are amplified by each secondary publication that was based on non-reproducible data.

Minimal Information (MI) checklists have been proposed for standardization of experimental description but a general drawback of these platforms is that they have been primarily conceived for simplifying the bioinformatics (re)use of experimental data. This effort is meaningful because metadata analysis of standardized datasets represents a valuable source of information and maximizes the usage of already existing results. [9-11]. Nevertheless, annotation following mandatory guidelines is often cumbersome and conflicting MI checklists have been proposed, despite the simplification efforts made by the community [12-14]. Moreover, only a few cases of MI checklists focus on methodologies for recombinant protein production and quality evaluation [15].

Protein production at lab scale is a straight-forward procedure. Nevertheless, each step implies making choices, providing controls, and dealing with the evident as well as the unappreciated pitfalls of the technology, such as changes in protein expression, physical and chemical alterations in protein structure, aggregation, and proteolysis. Since protein production is very often not the aim of most research projects, but simply the way to obtain intermediate reagents to start a research project, poor protein quality will undermine the robustness of complex multidisciplinary efforts. At the same time, general (cell) biologists are less aware of protein quality than specialists such as crystallographers, enzymologists, or protein chemists and biotechnologists. Therefore, we wish to propose a methodology for improving the qualitative evaluation of their proteins to researchers who are not “protein production specialists”.

Based on many years of experience in the protein production field, we would like to propose a practice that should simplify the assessment of the experimental set based on a flowchart for initial evaluation of experimental steps in protein production together with the corresponding data to append as Additional file 1 according to the guidelines of established initiatives such as Biosharing/MIBBI Foundry.

The protein production flowchart (Additional file 1) should help following the design of the protein production protocol outlining the critical points and to standardize and reproduce the results in other laboratories [16,17]. We suggest editors and reviewers to encourage (not compel) researchers to fill as many as possible of the listed requests (following the already available standards) to acquire the necessary information for the reliable evaluation of the proposed work. Clearly, the set of relevant data will change according to the final use of the protein and, therefore, there is no reason for mandatory universal guidelines (Table 1).

Table 1 The most basic requirements for evaluating protein quality

In-deep protein biophysical characterization needs specific expertise and specialized equipment, but any biology lab should be able to assess the produced proteins using to at least two complementary techniques:

1. PAGE-SDS provides multiple information regarding the quality of the protein such as the presence of degradation products as well as the absence of protein contamination.
 2. Analytical size exclusion chromatography (SEC) [18] provides information regarding the correct oligomeric structure of the protein and the absence of soluble aggregates that can cause non-specific results in downstream experiments.
-

Some editors might even consider attaching a special section of comments to the electronic version of the paper, allowing peers to grade the quality of the described protein production procedure, similar to sites such as TripAdvisor and others.

Competing interests

The authors declare no competing interests and equally contributed to design the flowchart and to write the text.

References

1. Bauerlein M, Gad-el-Hak M, Grody W, McKelvey B, Trimble SV: **We must stop the avalanche of low-quality research.** *Chron Higher Educ* 2010, **57**: June 13. <http://chronicle.com/article/We-Must-Stop-the-Avalanche-of/65890/>.
2. Chafen JJ, Newberry SJ, Riedl MA, Bravata DM, Maglione M, Suttorp MJ, Sundaram V, Paige NM, Towfigh A, Hulley BJ, Shekelle PG: **Diagnosing and managing common food allergies: a systematic review.** *JAMA* 2010, **303**:1848–1856.
3. de Marco A: **Reagent validation: an underestimated issue in lab praxis.** *J Mol Recognit* 2011, **24**:136.
4. Rich RL, Myszka DG: **Grading the commercial optical biosensor literature – Class of 2008: “The Mighty Binders”.** *J Mol Recognit* 2010, **23**:1–64.
5. Rich RL, Myszka DG: **Survey of the 2009 commercial optical biosensor literature.** *J Mol Recognit* 2011, **24**:892–914.
6. Garwood J: **Academic truth or biotech bullshit?** *LabTimes* 2012, **3**:40–44.
7. Prinz F, Schlange T, Asadullah K: **Believe or not: how much can we rely on published data on potential drug targets.** *Nature Rev Drug Discov* 2011, **10**:712–713.
8. Begley CG, Ellis LM: **Raise standards for preclinical cancer research.** *Nature* 2012, **483**:531–533.
9. Ball CA, Sherlock G, Brazma A: **Funding high-throughput data sharing.** *Nat Biotechnol* 2004, **22**:1179–1183.
10. Cusick ME, Yu H, Smolyar A, Venkatesan K, Carvunis A-R, Simonis N, Rual J-F, Borick H, Braun P, Dreze M, Vandenhoute J, Galli M, Yazaki J, Hill DE, Ecker JR, Roth FP, Vidal M: **Literature-curated protein interaction datasets.** *Nat Methods* 2009, **6**:39–46.
11. Leitner F, Chatr-aryamontri A, Mardis SA, Ceol A, Krallinger M, Licata L, Hirschman L, Cesareni G, Valencia A: **The FEBS Letters/BioCreative II.5 experiment: making biological information accessible.** *Nat Biotechnol* 2010, **28**:897–899.
12. Sansone SA, Rocca-Serra P, Field D, Maguire E, Taylor C, Hofmann O, Fang H, Neumann S, Tong W, Amaral-Zettler L, Begley K, Booth T, Bougueleret L, Burns G, Chapman B, Clark T, Coleman LA, Copeland J, Das S, de Daruvar A, de Matos P, Dix I, Edmunds S, Evelo CT, Forster MJ, Gaudet P, Gilbert J, Goble C, Griffin JL, Jacob D: **Toward interoperable bioscience data.** *Nat Genetics* 2012, **44**:121–126.

13. Kettner C, Field D, Sansone S-A, Taylor C, Aerts J, Binns N, Blake A, Britten CM, de Marco A, Fostel J, Gaudet P, González-Beltrán A, Hardy N, Hellemans J, Hermjakob H, Juty N, Leebens-Mack J, Maguire E, Neumann S, Orchard S, Parkinson H, Piel W, Ranganathan S, Rocca-Serra P, Santarsiero A, Shotton D, Sterk P, Untergasser A, Whetzel PL: **Meeting report from the second “Minimum Information about a Biological or Biomedical Investigation” (MIBBI) workshop.** *Standards Genomic Sci* 2010, **3**:3.
14. Tenenbaum JD, Sansone SA, Haendel M: **A sea of standards for omics data: sink or swim?** *J Am Med Inform Assoc* 2013. doi: 10.1136/amiajnl-2013-002066.
15. Buckle AM, Bate MA, Androulakis S, Cinquanta M, Basquin J, Bonneau F, Chatterjee DK, Cittaro D, Gräslund S, Gruszka A, Page R, Suppmann S, Wheeler JX, Agostini D, Taussig M, Taylor CF, Bottomley SP, Villaverde A, de Marco A: **Recombinant protein quality evaluation: proposal for a Minimal Information standard.** *Standards Genomic Sci* 2011, **5**:195–197.
16. Schroter S, Black N, Evans S, Godlee F, Osorio L, Smith R: **What errors do peer reviewers detect, and does training improve their ability to detect them?** *J R Soc Med* 2008, **101**:507–514.
17. Hrynaszkiewicz I: **A call for BMC Research Notes contributions promoting best practice in data standardization, sharing and publication.** *BMC Res Notes* 2010, **3**:235.
18. Sala E, de Marco A: **Screening optimized protein purification protocols by coupling small-scale expression and mini-size exclusion chromatography.** *Prot Expr Purif* 2010, **74**:231–235.

Additional file

Additional_file_1 as PDF

Additional file 1 Flowchart corresponding to a basic lab-scale protein production protocol. Process evaluation check-list allows for the precise identification of the steps and illustrates for each module the meaningful actions necessary to characterize the proteins used as reagents in biological experiments. Click the links on the image to obtain specifications and instructions. When available, module annotations should be completed according to the guidelines of accepted MI platforms: Biosharing/MIBBI Foundry <http://www.biosharing.org/standards/mibbi>. For instance: protein_purification_chromatography; http://wwwdev.ebi.ac.uk/micheckout/checkout/html?output-type=view_as_html_table&accessions=Column_chromatography. protein_sds_page; http://wwwdev.ebi.ac.uk/micheckout/checkout/html?output-type=view_as_html_table&accessions=Gel_electrophoresis.

Additional files provided with this submission:

Additional file 1: 9941692481242931_add1.pdf, 418K

<http://www.biomedcentral.com/imedia/7064367714114567/supp1.pdf>