

SOFTWARE

Open Access



# proMAD: semiquantitative densitometric measurement of protein microarrays

Anna Jaeschke<sup>1,2†</sup>, Hagen Eckert<sup>3,4\*†</sup>  and Laura J. Bray<sup>1,2</sup>

## Abstract

**Background:** Protein microarrays are a versatile and widely used tool for analyzing complex protein mixtures. Membrane arrays utilize antibodies which are captured on a membrane to specifically immobilize several proteins of interest at once. Using detection antibodies, the bound protein-antibody-complex is converted into visual signals, which can be quantified using densitometry. The reliability of such densitometric assessments depends on a variety of factors, not only sample preparation and the choice of acquisition device but also the selected analysis software and the algorithms used for readout and processing data. Currently available software packages use a single image of a membrane at an optimal exposure time selected for that specific experimental framework. This selection is based on a user's best guess and is subject to inter-user variability or the acquisition device algorithm. With modern image acquisition systems proving the capacity to collect signal development over time, this information can be used to improve densitometric measurements. Here we introduce *proMAD*, a toolkit for protein microarray analysis providing a novel systemic approach for the quantification of membrane arrays based on the kinetics of the analytical reaction.

**Results:** Briefly, our toolkit ensures an exact membrane alignment, utilizing basic computer vision techniques. It also provides a stable method to estimate the background light level. Finally, we model the light production over time, utilizing the knowledge about the reaction kinetics of the underlying horseradish peroxidase-based signal detection method.

**Conclusion:** *proMAD* incorporates the reaction kinetics of the enzyme to model the signal development over time for each membrane creating an individual, self-referencing concept. Variations of membranes within a given experimental set up can be accounted for, allowing for a better comparison of such. While the open-source library can be implemented in existing workflows and used for highly user-tailored analytic setups, the web application, on the other hand, provides easy platform-independent access to the core algorithm to a wide range of researchers. *proMAD*'s inherent flexibility has the potential to cover a wide range of use-cases and enables the automation of data analytic tasks.

**Keywords:** Membrane antibody array, Protein microarray, Densitometry, Web application, Python

\*Correspondence: [hagen.eckert@tu-dresden.de](mailto:hagen.eckert@tu-dresden.de)

<sup>†</sup>Anna Jaeschke and Hagen Eckert contributed equally to this work.

<sup>3</sup>Institute for Materials Science and Max Bergmann Center of Biomaterials, Technische Universität Dresden, Dresden, Germany

<sup>4</sup>Dresden Center for Computational Materials Science (DCMS), Technische Universität Dresden, Dresden, Germany

Full list of author information is available at the end of the article



## Background

Protein microarrays provide a versatile platform for the large-throughput analysis of numerous proteins present in a complex mixture based on the highly specific interactions between antibodies and antigens [1, 2]. The "sandwich assay," a multiplexed format of the Enzyme-linked Immunosorbent Assay (ELISA), is a widely used type of microarray. A variety of specific capture antibodies are immobilized on an array membrane. The sample, a blend of various proteins, is incubated with the membrane, and each target protein is trapped by a matching antibody capturing dozens of targets in parallel. A second antibody, detection or reporter antibody, is then employed to generate a chemiluminescent or fluorescent signal which can be detected using a film or, nowadays more commonly used, a CCD camera. While fluorescent dyes allow for multi-color detection systems, enzyme-based methods provide a significant improvement of sensitivity due to the amplification of the signal [2]. The most frequently used enzyme is horseradish peroxidase, which catalyzes the reaction of luminol with  $H_2O_2$ , thereby generating light [3, 4].

The signals on the array membrane are visible as bright spots generated by the protein-antibody-detection-reagent-complex on a dark background. The signal intensity of a spot and the abundance of the target protein are linked. Besides the quantitative assessment of the spots, presence vs. absence, the optical density of the signals can be quantified by densitometry. This quantification technique is also used in other immunoblot-based assays such as Western Blots [5]. Numerous aspects contribute to the reliability and reproducibility of densitometric assessments. Besides sample preparation techniques and the choice of acquisition device, the algorithms used for data processing are essential factors to consider [5–7]. Several programs are available for the quantification of immunoblots, some are linked to an acquisition software package, and others are specific for certain assay types. The method of densitometric measurement and the approaches for background subtraction vary between applications and details of the algorithms are often not openly accessible. Commonly, the densitometric readout is performed on an image at a particular chosen exposure time. This exposure time is determined either by the researcher or an algorithm in the acquisition software. Usually, some form of background subtraction method is applied. The details of these algorithms are determined by the applied software package. As there are no standardized protocols available, immunoblot densitometry has been described as based on traditions and guesswork [5].

Modern image acquisition systems are improving with regards to sensitivity, user-friendliness, and algorithms finding the optimal exposure time. These machines also provide the ability to capture images over a range of exposure times, allowing for the recording of signal

development over time and capturing dynamic changes. Utilizing this information has the potential to improve densitometric quantification approaches for immunoblots. However, currently available analysis tools are not equipped for the large-throughput tasks required for time-based analytics. At present, such image analysis would require time-intensive manual handling. To our knowledge, there is no analysis tool available which specifically employs the dynamic development of optical signals for protein microarray membrane analysis.

Here we introduce *proMAD*, a toolkit for analyzing protein microarrays and a novel systemic analytical concept for the quantification of optical signals detected on the membrane. Utilizing the information of dynamic signal change over time combined with the reaction kinetics, *proMAD* allows us to model the signal development for each membrane. Thereby, the signal quantification does not rely on the information obtained from a single image acquired at a particular exposure time, which might not be optimal for every membrane in a set of membranes. This individual, self-referencing approach allows for a better comparison of different membranes from the same kit within a given experimental framework.

With the *proMAD* open-source library, the core algorithm can be implemented in highly customizable workflows which includes user-defined membrane layouts. A simple, and platform-independent access to the toolkit is provided by the *proMAD* web application. Thereby, the toolkit is accessible for a wide range of users and use-cases such as various membrane types.

## Implementation

The core algorithm used in *proMAD* consists of three main parts. Firstly, the steps to align the raw image sets are presented. Secondly, an approach to estimate the background light level in a stable manner is described. Finally, the correlation between reaction progression and light production over time is demonstrated. The detailed experimental procedures for obtaining the processed images are described in the [Supplementary information](#).

### Image alignment

Exact alignment of the image sets is imperative to ensure correct signal readout and to generate reproducible results. Multiple causes can affect image alignments such as rotation, warp, or stretching. For example, due to the dampness of the membrane, it can lay skewed on the recording tray. In other cases, the position might not remain constant over time. To correct the membrane position, basic computer vision is employed.

To align the images, the software relies on the ordered structure of the analytic spots on the membrane. The bright "reference spots" on the edge of the membranes are convenient markers for adjustment. Initially, the images

provided by the user either need to be already roughly rotated upright or a rotation parameter needs to be passed to the library. Subsequently, the software finds the anchor points automatically by extracting the contours of the brightest spots. For this purpose, the two-dimensional version of the marching cube algorithm is used [8]. Thereby, the contours are searched at 60% of the maximum gray value found in the raw image and for each contour, a central point  $C$  is calculated (Fig. 1 ●). From the collection of central points, three points are selected to generate the first guess  $G$  (Fig. 1 ◆) of the anchor point positions as outlined in Eqs. 1 to 3.

$$G_1 = [\min(C_X), \min(C_Y)] \quad (1)$$

$$G_2 = [\min(C_X), \max(C_Y)] \quad (2)$$

$$G_3 = [\max(C_X), \min(C_Y)] \quad (3)$$

For each estimated point, the contour center  $C$  with the shortest distance to the initial guess  $G$  is selected as an anchor point (Fig. 1 ▲).

It is important to note that this part of the algorithm would need modification to support membranes that have an alternative layout of “reference spots”. Three references points are required to implement a full alignment.

Three tests are applied to ensure that the selected points are meaningful prior to the modification of the image: First, the distance ratio between the three points is compared to the values expected for the particular membrane type. Second, the angle formed by the points and third, the distance between the initial guess and the reference points are checked. The expected values might differ for each membrane type and are determined in the array configuration file. If the checks are within the given tolerances, the image is warped and trimmed. The overall dimension

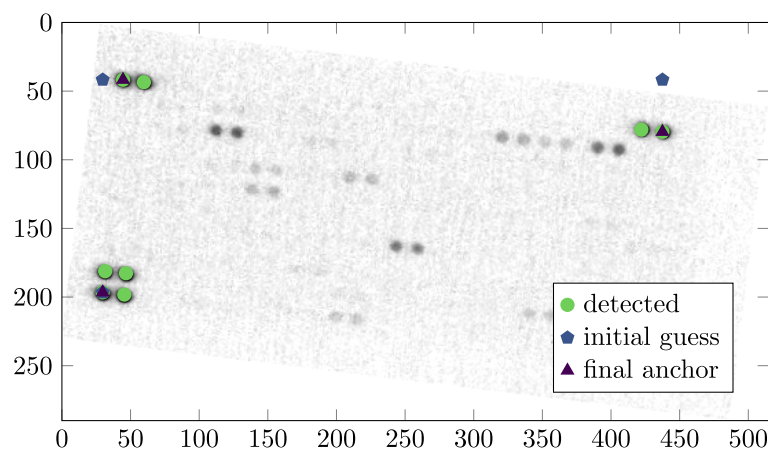
is chosen in such a way, that each spot is centered in a 30-pixel square. This particular value can be changed without altering the results of the presented methods in this paper.

### Background estimation

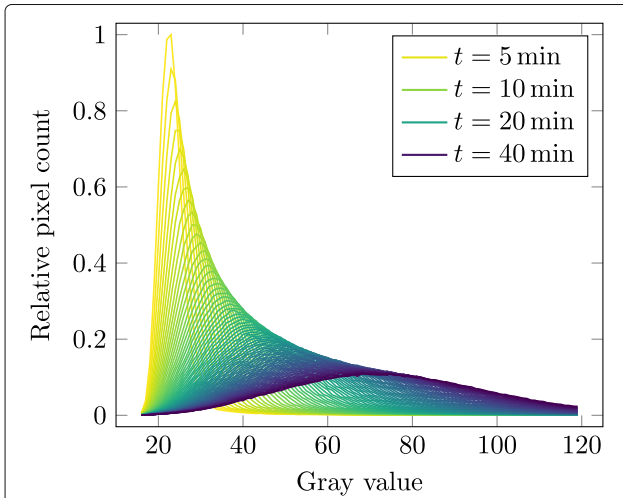
During the recording of the image sets, the sensor collects signals that are not related to the analytical reaction in the corresponding spot. Light originating from an imperfect enclosure or reflections of the whole membrane is recorded over time. Additionally, noise generated by the sensor accumulates as well. Quantifying this background noise is essential to improve the quality of the data extraction and can also be used as a reference point to compare single images with unknown exposure times, that were measured in the same setup.

In this work, an approach based on the normal distribution of the background noise was used. The vast majority of the membranes can be considered dark. Therefore, an extensive sample range is available. To adjust for different input formats, all imported images are internally scaled to be in the range between zero and one. However, the process of generating the background value  $b$  is demonstrated in Figs. 2 and 3 on an unscaled 16 bit example, to visualize the steps more concisely.

Initially, the image data of each time step is treated as a simple collection of values. Next, the dataset is reduced from of 65 536 to 2048 levels of precision to enable better statistical analytics. The highest peak (see Fig. 2) in these 2048 bins represents the first guess of the background gray value  $b_G$ . All values higher than  $2b_G$  are discarded and as such are not considered part of the background noise. This cutoff ensures that the light originating from the chemical reaction does not alter the background level quantification. The remaining values are fitted to a normal distribution. Finally, the mean of the resulting



**Fig. 1** Image alignment. Example membrane image which is not aligned correctly. The detected bright spots are displayed as circles (●). The initial guesses  $G_1, G_2, G_3$  of the anchor points are represented as diamonds (◆), and the final anchors as triangles (▲). The image of the membrane is inverted for clarity



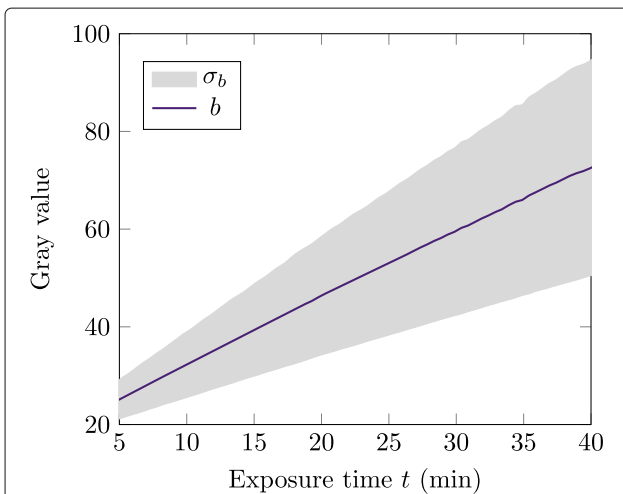
**Fig. 2** Background noise distribution. Evolution of the background noise distribution over time for a single membrane. The grayscale precision is reduced to 2048 steps for histogram representation

distribution is used to define the background parameter  $b$  of a membrane.

As shown in Fig. 3, the background values  $b$  exhibit a linear relationship to the exposure time. Moreover, the deviation of the distribution  $\sigma_b$  grows linear over time.

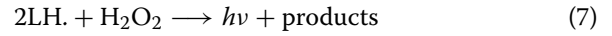
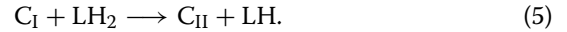
### Reaction kinetic

A commonly used detection method in membrane arrays is based on light generated by the enzyme horseradish peroxidase which reacts with  $\text{H}_2\text{O}_2$  to activate luminol. The kinetics of this reaction were extensively studied by Cormier and Prichard in 1968 [4]. Our approach utilizes this knowledge as the foundation to model the light production. The reaction steps of luminol ( $\text{LH}_2$ ) with  $\text{H}_2\text{O}_2$



**Fig. 3** Background. Background gray value development  $b$  over time  $t$ . The gray area indicates the standard deviation  $\pm\sigma_b$

catalyzed by the horseradish peroxidase enzyme ( $E$ ) are listed in Eqs. 4 to 7.



The reaction rate ( $\nu$ ) can be simplified and written as in Eq. 8. Here, the reaction constants presented by Cormier and Prichard [4] are combined into the parameters  $\alpha$ ,  $\beta$ , and  $\gamma$  for clarity.

$$\nu = \frac{\alpha C_E C_{\text{LH}_2} C_{\text{H}_2\text{O}_2}}{\beta C_{\text{LH}_2} + \gamma C_{\text{H}_2\text{O}_2}} \quad (8)$$

Using the numeric results published by Cormier and Prichard [4] and Chance [9], we can deduct the relation of the simplified parameters to be  $\alpha > \gamma > \beta$ . Additionally, an excess of  $\text{H}_2\text{O}_2$  ( $C_{\text{H}_2\text{O}_2} > C_{\text{LH}_2}$ ) is expected, resulting in a further simplification represented in Eq. 10

$$\nu = \frac{\alpha C_E C_{\text{LH}_2} C_{\text{H}_2\text{O}_2}}{\gamma C_{\text{H}_2\text{O}_2}} \quad (9)$$

$$\nu = k_r C_E C_{\text{LH}_2} \quad (10)$$

The value of the apparent rate constant ( $k_r = \alpha/\gamma$ ) can be approximated to  $k_r \approx 1.4 \times 10^6 (\text{Ms})^{-1}$ . As the apparent rate constant  $k_r$ , and the concentration of the catalytic enzyme  $C_E$  are independent of the reaction time, the reaction rate development is determined by the concentration of one reagent ( $\text{LH}_2$ ). Under these circumstances, the reaction can be described as of first order and Eq. 11 outlines the resulting evolution of the luminol concentration. The reaction rate of the enzyme can be expressed in a time-dependent manner (Eq. 12).

$$C_{\text{LH}_2}(t) = C_{0,\text{LH}_2} \exp(-k_r C_E t) \quad (11)$$

$$\nu(t) = k_r C_E C_{0,\text{LH}_2} \exp(-k_r C_E t) \quad (12)$$

A light flux is produced based on the reaction turnover (Eq. 7) and its intensity is related to the reaction speed  $\nu(t) \propto I(t)$ . By introducing the proportionality factor  $k_i$ , we can describe the light intensity  $I(t)$  in Eq. 13. The variable  $t$  is chosen to represent the exposure time as recorded by the image acquisition device, to enable better comparison with the experimental data. The time between starting the reaction (i.e., pipetting the reagent onto the membrane) and the start of the image recording is included by adding the preparation time  $t_0$  as a constant.

$$I(t) = k_i k_r C_E C_{0,\text{LH}_2} \exp(-k_r C_E (t + t_0)) \quad (13)$$

In the experimental setup, the light flux is measured cumulative over time. Therefore, Eq. 13 is integrated over the exposure time to analyze the data.

$$\int_0^t I(t) dt = k_i C_{0,LH_2} \left[ -\exp(-k_r C_E(t_0 + t)) + \exp(-k_r C_E t_0) \right] \quad (14)$$

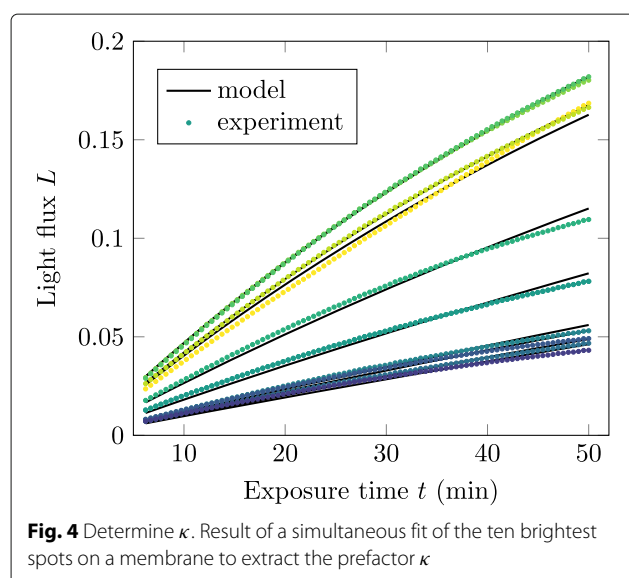
$$L(t) = \kappa \left[ -\exp(-k_r C_E(t_0 + t)) + \exp(-k_r C_E t_0) \right] \quad (15)$$

The solution for this integration in the boundaries is presented in Eq. 14. As the prefactors will be indistinguishable in the analytics of the data, they are combined as  $\kappa = k_i C_{0,LH_2}$  (Eq. 15) to simplify the expression. This factor  $\kappa$  is specific for each image set. It depends on the initial luminol concentration, on the emission efficiency, and the specifics of the recording equipment.

In the implementation of *proMAD*, the light flux  $L(t)$  is measured by the average signal in the square around each spot reduced by the background value  $b$  of that frame. The simplified prefactor  $\kappa$  is defined by minimizing the errors of the brightest spots simultaneous while fitting to Eq. 15 (see Fig. 4). Finally, the concentration of the horseradish peroxidase enzyme  $C_E$  bounded to the proteins is determined with the estimated  $\kappa$  for each spot individually. The target protein content bound on the membrane is directly reflected by the enzyme concentration.

## Results and discussion

The core algorithm of *proMAD* can be accessed through the open-source library or the web application. The Python library can be implemented in user-specific analysis setups and presents a wide range of opportunities for customization. The web tool is suited for users who want



**Fig. 4** Determine  $\kappa$ . Result of a simultaneous fit of the ten brightest spots on a membrane to extract the prefactor  $\kappa$

to run the stable version *proMAD* without the need to install any software.

At present, a range of commonly used image formats can be imported. Raw images (.scn) from ChemiDoc™ MP systems (BioRad, Gladesville, Australia) can be used directly. The core algorithm, the reaction-based model, is available when the exposure times can be accessed in the loaded images. Multiple additional evaluation modes can be selected in cases where the exposure time information is missing. The raw mode returns a list of averaged gray values for the spots on all original images in the stack. Deduction of the histogram-based background value from this list is possible. Moreover, a morphological background detection can be used to separate the foreground and the background of the raw images. The local background approach calculates the mean of the ratios between the average of a spot on the raw images, and the extracted backgrounds. Besides, the linear correlation between the histogram background value and the average foreground value or the raw image over time can be used to evaluate the signals. The results for a selected evaluation method can be summarized in a report. At present, four report modules are available: json, csv, excel, and  $\LaTeX$ . The excel and  $\LaTeX$  report files contain the average values of each analyte, as well as a graphical representation of the samples with the highest signals. An image of the membrane serves as an alignment check. Additionally, information about the software version and user-provided naming of the data set and membranes is included in the report. Currently, *proMAD* supports four different membrane types. However, other membrane array layouts can be easily implemented, given the presence of at least three reference spots for the alignment algorithm.

Fundamental aspects of the toolkit, such as the algorithm to measure the intensity as well as the automatic alignment system and the background analysis can be adapted for other types of densitometric measurements. The main algorithm presented, depends on the reduction of the light-emitting material over time. Therefore, it can be applied to techniques that are also using an enzyme-based detection method with known reaction kinetics.

## Library

The *proMAD* library is accessible through the Python Package Index (PyPI). Alternatively, it can be installed directly using the code available on Github [10]. The user can set up a highly tailored workflow and integration with other analysis protocols.

## Web application

The *proMAD* web application [11] provides an easy-to-use interface to process membrane images in the cloud. This approach is suitable to analyze small sample sets without the need to set up the development environment.

The web interface guides the users through the process step-by-step. Multiple input files can be bundled in a zip or tar container to accelerate the upload. Several membranes recorded in one image stack can be separated and processed. The process can be completed in multiple sessions using the displayed request key. At the end of the image analysis process, various formats for downloading the results are presented.

## Conclusion

Here we present *proMAD*, a novel systemic analytical approach for the quantification of membrane protein arrays. The distinctive, self-referencing concept is based on the evaluation of the dynamic signal development using the knowledge about the kinetics of the underlying chemical reaction. Modern image acquisition devices allow for collecting data about signal development over time for each given membrane. By utilizing this information, deviations between different membranes can be accounted for, enabling better comparison of such within a given experimental setup. While the web application provides an easy and platform-independent access to the *proMAD* core algorithm to a wide range of researchers, the underlying open-source library, on the other hand, allows for highly customized data analysis workflows. The toolkit is relevant to a variety of end-users while keeping the underlying algorithm openly accessible. *proMAD*'s intrinsic flexibility has the potential to apply the presented algorithm to a wide range of additional use-cases.

## Availability and requirements

**Project name:** proMAD

**Project home page:** <https://promad.dev>

**Source code:** <https://github.com/theia-dev/proMAD>

**Operating system(s):** platform-independent; web application

**Programming language:** Python  $\geq 3.6$

**Other requirements:** refer to [requirements.txt](#)

**License:** MIT

**Any restrictions to use by non-academics:** None

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12859-020-3402-4>.

**Additional file 1:** Supplementary information. The supplementary file contains the detailed experimental procedures for obtaining the processed images.

## Abbreviations

*b*: Background level; *b<sub>G</sub>*: Background level guess; *C*: Concentration; *C<sub>0</sub>*: Initial concentration; *C<sub>X,Y</sub>*: Contour center; *E*: Horseradish peroxidase enzyme; *G*: Anchor point guess; *h*: Planck constant; *I*: Light intensity; *k<sub>i</sub>*: Proportionally constant; *k<sub>r</sub>*: Apparent rate constant; *L*: Light flux; *LH<sub>2</sub>*: Luminol; *t*: Time; *t<sub>0</sub>*: Recording start time; *v*: Reaction speed; *I, II*: Intermediate reaction products;

$\alpha, \beta, \gamma$ : Reaction parameter;  $\sigma_b$ : Background level standard deviation;  $\kappa$ : Combined prefactor;  $\nu$ : Photon frequency

## Acknowledgements

Not applicable.

## Authors' contributions

AJ performed the experiments. HE developed the algorithm. AJ and HE analyzed and interpreted the data, developed the website, and wrote the manuscript. AJ, HE and LJB read, edited and approved the final manuscript.

## Funding

We would like to thank Cure Cancer Australia and the Centre in Regenerative Medicine for supporting this project. Hagen Eckert was supported by the Deutsche Forschungsgemeinschaft (FOR 2038: Nanostructured Organic Matrices in Biological Silica Mineralization). Open Access Funding was provided by the Publication Fund of the TU Dresden.

## Availability of data and materials

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Institute of Health and Biomedical Innovation, Queensland University of Technology, 4059 Kelvin Grove, Queensland, Australia. <sup>2</sup>School of Mechanical, Medical and Process Engineering, Science and Engineering Faculty, Queensland University of Technology, 4059 Kelvin Grove, Queensland, Australia. <sup>3</sup>Institute for Materials Science and Max Bergmann Center of Biomaterials, Technische Universität Dresden, Dresden, Germany. <sup>4</sup>Dresden Center for Computational Materials Science (DCMS), Technische Universität Dresden, Dresden, Germany.

Received: 21 September 2019 Accepted: 7 February 2020

Published online: 24 February 2020

## References

- Kopf E, Zharhary D. Antibody arrays - A emerging tool in cancer proteomics. *Int J Biochem Cell Biol.* 39(7):1305–17. <https://doi.org/10.1016/j.biocel.2007.04.029>.
- Sutandy FXR, Qian J, Chen C-S, Zhu H. Overview of Protein Microarrays. *Curr Protocol Protein Sci.* 2013;72(1):27–1127116. <https://doi.org/10.1002/0471140864.ps2701s72>.
- Poetz O, Schwenk JM, Kramer S, Stoll D, Templin MF, Joos TO. Protein microarrays: catching the proteome. *Mech Ageing Dev.* 2005;126(1):161–70. <https://doi.org/10.1016/j.mad.2004.09.030>.
- Cormier MJ, Prichard PM. An investigation of the mechanism of the luminescent peroxidation of luminol by stopped flow techniques. *J Biol Chem.* 1968;243(18):4706–14.
- Gassmann M, Grenacher B, Rohde B, Vogel J. Quantifying Western blots: Pitfalls of densitometry. *Electrophoresis.* 2009;30(11):1845–55. <https://doi.org/10.1002/elps.200800720>.
- Ghosh R, Gilda JE, Gomes AV. The necessity of and strategies for improving confidence in the accuracy of western blots. *Expert Rev Proteomics.* 2014;11(5):549–60. <https://doi.org/10.1586/14789450.2014.939635>.
- Taylor SC, Berkelman T, Yadav G, Hammond M. A Defined Methodology for Reliable Quantification of Western Blot Data. *Mol Biotechnol.* 2013;55(3):217–26. <https://doi.org/10.1007/s12033-013-9672-6>.
- Lorenson WE, Cline HE. Marching cubes: A high resolution 3D surface construction algorithm. In: *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '87*, vol. 21. New York: ACM Press; 1987. p. 163–9. <https://doi.org/10.1145/37401.37422>. <http://portal.acm.org/citation.cfm?doid=37401.37422>.

9. Chance B. The kinetics and stoichiometry of the transition from the primary to the secondary peroxidase peroxide complexes. *Arch Biochem Biophys*. 1952;41(2):416–24. [https://doi.org/10.1016/0003-9861\(52\)90470-0](https://doi.org/10.1016/0003-9861(52)90470-0).
10. Source repository for proMAD. <https://github.com/theia-dev/proMAD>.
11. Web application for proMAD. <https://promad.dev>.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

